Article

Phylogenetic evidence for independent origins of GDF1 and GDF3 genes in amphibians and mammals

Juan C. Opazo & Kattina Zavala
Instituto de Ciencias Ambientales y Evolutivas, Facultad de Ciencias,
Universidad Austral de Chile, Valdivia, Chile.

Corresponding author
Juan C. Opazo
Instituto de Ciencias Ambientales y Evolutivas
Facultad de Ciencias
Universidad Austral de Chile
Phone: +56 63 2221674
Email: jopazo@gmail.com

Abstract

Growth differentiation factors 1 (GDF1) and 3 (GDF3) are members of the transforming growth factor superfamily (TGF-β) that is involved in fundamental early-developmental processes that are conserved across vertebrates. The evolutionary history of these genes is still under debate due to ambiguous definitions of homologous relationships among vertebrates. Thus, the goal of this study was to unravel the evolution of the GDF1 and GDF3 genes of vertebrates, emphasizing the understanding of homologous relationships and their evolutionary origin. Surprisingly, our results revealed that the GDF1 and GDF3 genes found in amphibians and mammals are the products of independent duplication events of an ancestral gene in the ancestor of each of these lineages**.** The main implication of this result is that the GDF1 and GDF3 genes of amphibians and mammals are not 1:1 orthologs. In other words, genes that participate in fundamental processes during early development have been reinvented two independent times during the evolutionary history of tetrapods.

Keywords: gene duplication; gene family evolution; tetrapods; left-right identity; development

Introduction

Growth differentiation factors 1 (GDF1) and 3 (GDF3) are members of the transforming growth factor superfamily (TGF-β) that were originally isolated from mouse embryonic libraries (Lee 1990; Jones et al., 1992). They perform fundamental roles during early development, GDF1 has been mainly associated to the regulation of the left-right patterning, whereas GDF3 is mainly involved in the formation of the anterior visceral endoderm, mesoderm and the establishment of anterior-posterior identity of the body (Lee 1991; Dale ' et al., 1993; Thomsen and Melton 1993; Hyatt et al., 1996; Hyatt and Yost 1998; Söderström and Ebendal 1999; Rankin et al., 2000; Wright 2001; Hamada et al., 2002; Somi et al., 2003; Andersson et al., 2006; Chen 2006; Bengtsson et al., 2008; Komatsu and Mishina 2013; Arias et al., 2017; Bisgrove et al., 2017; Montague and Schier 2017). Deficiencies in GDF1/GDF3 give rise to a broad spectrum of defects including right pulmonary isomerism, visceral *situs inversus*, transposition of the great arteries, and cardiac anomalies among others (Rankin et al., 2000; Wall et al., 2000; Karkera et al., 2007; Kaasinen et al., 2010; Bao et al., 2015; Zhang et al., 2015). In addition to their developmental roles, GDF1 has been described as a tumor suppressor gene in gastric cells, counteracting tumorogenesis by stimulating the SMAD signaling pathway (Yang et al., 2016), and GDF3 has been associated with the regulation of adipose tissue homeostasis and energy balance during nutrient overload (Witthuhn and Bernlohr 2001; Wang et al., 2004; Andersson et al., 2008).

The evolutionary history of the GDF1 and GDF3 genes is still a matter of debate due to the unclear definition of homologous relationships. Understanding homology is a fundamental aspect of biology as it allows us to comprehend the degree of relatedness between genes that are associated to a given phenotype in a group of organisms. This is particularly important for GDF1 and GDF3 as these genes perform biological functions during early stages of development that define key aspects of the body plan of all vertebrates (Lee 1991; Dale ' et al., 1993; Thomsen and Melton 1993; Hyatt et al., 1996; Hyatt and Yost 1998; Söderström and Ebendal 1999; Rankin et al., 2000; Wright 2001; Hamada et al., 2002; Somi et al., 2003; Andersson et al., 2006; Chen 2006; Bengtsson et al., 2008; Komatsu and Mishina 2013; Arias

et al., 2017; Bisgrove et al., 2017; Montague and Schier 2017). Until now, most of the inferences of the homology relationships of these genes have been based on functional information. For example, based on the developmental processes that these genes regulate, it has been suggested that the mammalian GDF1 gene is the true ortholog of the Vg1 (GDF1) gene found in amphibians (Hyatt and Yost 1998; Rankin et al., 2000; Wall et al., 2000). Furthermore, phylogenetic analyses performed by Andersson et al., (2007) were not able to define orthologous relationships between the GDF1 and GDF3 genes among vertebrates. However after performing genomic comparisons, these authors did propose that the GDF1 gene present in mammals is the true ortholog of the Vg1 (GDF1) gene present in amphibians (Andersson et al., 2007). It was also suggested that the GDF3 gene could be an evolutionary innovation of mammals, as Andersson et al., (2007) did not find GDF3 sequences in the amphibian and bird genomes (Levine et al., 2006; Andersson et al., 2007). Given this scenario, the single copy gene found in amphibians and birds would be a co-ortholog of the mammalian genes (Andersson et al., 2007). More recently, a second copy of a GDF gene (derrière) has been annotated at the 3´ side of the Vg1 (GDF1) in the genome of the western clawed frog (*Xenopus tropicalis*), thus further complicating the definition of homologous relationships and the origin of genes that accomplish fundamental roles during early development in vertebrates.

With emphasis on understanding homologous relationships and their evolutionary origin, the goal of this study was to unravel the evolution of the GDF1 and GDF3 genes of vertebrates. Surprisingly, our phylogenetic analyses revealed that the GDF1 and GDF3 genes of amphibians and mammals are the products of independent duplication events of an ancestral gene in the ancestor of each of these groups. We also found the signature of two chromosomal translocations, the first occurred in the ancestor of tetrapods whereas the second was found in the ancestor of mammals. Thus, our results support the hypothesis that in amphibians and mammals, descendent copies of the same ancestral gene (GDF1/3) have independently subfunctionalized to perform key developmental functions in vertebrates.

Results and Discussion

112 ***Phylogenetic analyses suggest an independent origin of the GDF1 and***
113 ***GDF3 genes in mammals and amphibians***
114 From an evolutionary perspective, the definition of homologous relationships
115 among GDF1 and GDF3 genes is still a matter of debate (Levine et al., 2006;
116 Andersson et al., 2007). The resolution of this homology is important as these
117 genes are involved in fundamental developmental processes which are
118 conserved all across vertebrates (Mercola and Levin 2001; Levine 2005;
119 Nakamura and Hamada 2012; Komatsu and Mishina 2013). Thus, if extant
120 species inherited these genes from the vertebrate ancestor, the
121 developmental processes in which GDF1 and GDF3 are involved have a
122 single evolutionary origin and are comparable among species.
123 Based on evolutionary analyses and the developmental processes in
124 which GDF1 has been shown to participate, it has been suggested that GDF1
125 in mammals, birds, and amphibians are 1:1 orthologs (Seleiro et al., 1996;
126 Hyatt and Yost 1998; Rankin et al., 2000; Wall et al., 2000; Andersson et al.,
127 2007). Given that it has not been possible to identify copies of the GDF3 gene
128 in amphibians and birds, it has been proposed that this gene is an
129 evolutionary innovation of mammals, and the single gene copy found in
130 amphibians and birds is co-ortholog to the mammalian duplicates (Andersson
131 et al., 2007). However, a second copy of a GDF gene has been annotated in
132 the genome of the western clawed frog (*Xenopus tropicalis*) (Suzuki et al.,
133 2017). This newly described gene is located at the 3´ side of the GDF1 gene
134 and indicates that amphibians, like mammals, also have a repertoire of two
135 GDF genes (GDF1 (Vg1) and GDF3 (derrière)). This new finding further
136 complicates the resolution of homologous relationships and the origin of
137 genes involved in the early development of all vertebrates (Mercola and Levin
138 2001; Levine 2005; Nakamura and Hamada 2012; Komatsu and Mishina
139 2013).
140 Our maximum likelihood and Bayesian reconstructions revealed that
141 the GDF1 and GDF3 genes of amphibians and mammals were both
142 reciprocally monophyletic (Fig. 1), indicating that these genes originated
143 independently in each of these two lineages (Fig. 1). In agreement with the
144 literature, we found only one gene in sauropsids (e.g. birds, turtles, lizards
145 and allies), suggesting that they retained the ancestral condition of a single

146 gene copy as found in non-tetrapod vertebrates (e.g. coelacanths, bony fish,
147 chondricthyes and cyclostomes)(Fig. 1). Dot-plot comparisons provided
148 further support for the presence of a single gene copy in sauropsids as no
149 traces of an extra GDF gene were present in the syntenic region of
150 representative species of the group (Fig. 2). However, we cannot rule out an
151 alternative scenario of duplication and subsequent gene loss in the ancestor
152 of sauropsids. Thus, our evolutionary analyses suggest that the GDF1 and
153 GDF3 genes present in amphibians and mammals diversified independently
154 in the ancestor of each of these lineages. In other words, genes that
155 participate in fundamental processes during early development have been
156 reinvented two independent times during the evolutionary history of tetrapods.
157 As a consequence of the independent origin of these genes in amphibians
158 and mammals, they are not 1:1 orthologs.
159 To further test our hypothesis of the independent origins of the GDF1
160 and GDF3 genes in amphibians and mammals we performed topology tests.
161 In these analyses we compared our phylogenetic tree (Fig. 1 and Fig. 3B) to
162 the topology predicted from a one-duplication model in which the duplication
163 event that gave rise to the GDF1 and GDF3 genes in amphibians and
164 mammals occurred in the ancestor of tetrapods (Fig. 3A). In the one-
165 duplication model, it is assumed that the ancestor of sauropsids had two gene
166 copies and that subsequently one of these was lost. Thus, according to the
167 one-duplication model (Fig. 3A), the predicted phylogeny would recover a
168 monophyletic group containing GDF1 sequences from mammals, sauropsids,
169 and amphibians sister to a clade containing GDF3 sequences from the same
170 groups (Fig. 3A). Alternatively, the predicted phylogeny from the two-
171 duplication model would retrieve a clade containing GDF1 and GDF3
172 sequences from mammals sister to a clade containing GDF1/3 sequences
173 from sauropsids; additionally a clade containing GDF1 and GDF3 sequences
174 from amphibians would be recovered sister to the mammalian/sauropsid clade
175 (Fig. 3B). Results of the topology tests rejected the phylogeny predicted by
176 the one-duplication model (Weighted Shimodaira and Hasegawa, $p< 10^{-4}$;
177 Weighted Kishino Hasegawa, $p< 10^{-4}$). Thus, this result provided additional
178 support to our hypothesis that the GDF1 and GDF3 genes of amphibians and

179 mammals are the product of lineage independent duplication events in the
180 ancestors of each of these groups.
181 In the literature there are other cases of groups of genes that perform
182 similar biological functions in a diversity of species but that have originated via
183 lineage independent duplication events (Goodman et al., 1987; Kriener et al.,
184 2001; Wallis and Wallis 2002; Li et al., 2005; Opazo et al., 2008; Hoffmann et
185 al., 2010; Hoffmann et al., 2018). Among these, the independent origin of the
186 β-globin gene cluster in all main groups of tetrapods (e.g. therian mammals,
187 monotremes, birds, crocodiles, turtles, squamates, amphibians) represents a
188 well-documented phenomenon (Opazo et al., 2008; Hoffmann et al., 2010;
189 Hoffmann et al., 2018). This case is of particular interest as the two β-globin
190 subunits that come from a gene family that has been reinvented several times
191 during the evolutionary history of tetrapods are assembled in a tetramer with
192 two α-globin subunits that belong to a group of genes that possess a single
193 origin (Goodman et al., 1987; Opazo et al., 2008; Hoffmann et al., 2010;
194 Hoffmann et al., 2018). Besides this example, the independent origin of gene
195 families makes the task of comparison difficult, as the repertoire of genes
196 linked to physiological processes in different lineages do not have the same
197 evolutionary origin. Additionally, the fact that the evolutionary process can
198 give rise to similar phenotypes following different mutational pathways makes
199 the problem of comparing even more difficult (Natarajan et al., 2016). This is
200 particularly important when extrapolating the results of physiological studies
201 performed in model species to other organisms.
202
203 ***Two translocation events during the evolutionary history of the GDF1***
204 ***and GDF3 genes of tetrapods***
205 Based on the chromosomal distribution of the GDF1 and GDF3 genes and the
206 conservation pattern of flanking genes, we propose that during the
207 evolutionary history of tetrapods, these genes underwent two chromosomal
208 translocation events (Fig. 4). According to our analyses the genomic region
209 that harbors the single gene copy of chondricthyes, bony fish, and
210 coelacanths is conserved, yet this region in these groups differs from the
211 regions in which GDF1 and GDF3 are located in tetrapods. Overall this
212 suggests that the first translocation event occurred in the ancestor of

213 tetrapods (Fig. 4). Interestingly, the genomic region where the single gene
214 copy is located in non-tetrapod vertebrates, which is defined by the presence
215 of upstream (BMP2, HAO1, TMX4 and PLCB1) and downstream genes
216 (FERMT1, LRRN4, CRLS1), is conserved in tetrapods. Given this, it is
217 possible to identify the chromosomal location where the gene was located in
218 the tetrapod ancestor before the first translocation event (Fig. 4). On the other
219 hand, the fact that the mammalian GDF1 and GDF3 genes are located in
220 different chromosomes suggests that the second translocation event occurred
221 in the ancestor of the group (Fig. 4). In humans, the GDF1 gene is located on
222 chromosome 19 while GDF3 is located on chromosome 12. In opossum
223 (*Didelphis virginiana*) the GDF1 and GDF3 genes are located on
224 chromosomes 3 and 8, respectively.
225
226 ***Evolution of GDF1 and GDF3 genes in vertebrates***
227 In this study we present compelling evidence suggesting that genes involved
228 in the formation of the primitive streak, anterior visceral endoderm, mesoderm
229 and the establishment of the left-right identity (Lee 1991; Dale ' et al., 1993;
230 Thomsen and Melton 1993; Hyatt et al., 1996; Hyatt and Yost 1998;
231 Söderström and Ebendal 1999; Rankin et al., 2000; Somi et al., 2003;
232 Andersson et al., 2006; Chen 2006; Andersson et al., 2007; Bengtsson et al.,
233 2008; Arias et al., 2017; Bisgrove et al., 2017; Montague and Schier 2017) in
234 amphibians and mammals are the product of independent duplications
235 events. As such, these results indicate that the GDF1 and GDF3 genes of
236 amphibians and mammals are not 1:1 orthologs.
237 Thus, according to our results the last common ancestor of vertebrates
238 had a repertoire of one gene (GDF1/3), and this condition is maintained in
239 actual species of cylostomes, chondricthyes, bony fish, and coelacanths (Fig.
240 5). In the ancestor of tetrapods, the single gene copy was translocated from a
241 chromosomal region defined by the presence of the BMP2, HAO1, TMX4,
242 PLCB1, FERMT1, LRRN4, CRLS1 genes to a chromosomal region defined by
243 the presence of the CERS1, COPE, DDX49 and HOMER3 genes (Fig. 4).
244 After this translocation, the ancestral gene underwent a duplication event in
245 the amphibian ancestor, giving rise to the GDF1 (Vg1) and GDF3 (derrière)
246 genes as they are found in actual species. In the case of the western clawed

frog (*Xenopus tropicalis*) these genes are located in tandem on chromosome 28 (Fig. 4). Sauropsids, the group that includes birds, crocodiles, turtles, lizards and snakes, inherited the ancestral condition of a single gene copy as is seen in non-tetrapod vertebrates (Fig. 5). Finally, in the ancestor of mammals, the ancestral gene also underwent a duplication event, giving rise to the GDF1 and GDF3 genes found in extant species of mammals. After duplication but before the radiation of the group, a second translocation event occurred in the ancestor of the group (Fig. 5). Thus, all mammals inherited a repertoire of two genes (GDF1 and GDF3) that are located on two chromosomes.

Concluding remarks

This study provides a comprehensive evolutionary analysis of the GDF1 and GDF3 genes in representative species of all main groups of vertebrates. The main focus of this study was to unravel the duplicative history of the GDF1 and GDF3 genes and to understand homologous relationships among vertebrates. Understanding homology in this case is particularly important as these genes perform fundamental roles during early development that are conserved across vertebrates (Mercola and Levin 2001; Levine 2005; Nakamura and Hamada 2012; Komatsu and Mishina 2013). Surprisingly, our results revealed that the GDF1 and GDF3 genes present in amphibians and mammals are the product of independent duplication events in the ancestor of each of these groups. Subsequently, the GDF1 and GDF3 genes of amphibians and mammals are not 1:1 orthologs. Our results also show that all other vertebrate groups - i.e non-tetrapods and sauropsids – maintained the ancestral condition of a single gene copy (GDF1/3).

From an evolutionary perspective the independent duplication events that occurred in the ancestors of mammals and amphibians could have resulted in the division of labor, with some degree of redundancy, of the function performed by the ancestral gene. In support of this idea, it has been shown that the GDF1 and GDF3 genes of mammals have partially redundant functions during development where GDF1 can to some degree compensate for the lack of GDF3 (Andersson et al., 2007). Additionally, it has also been shown that double knockout animals ($GDF1^{-/-}$ and $GDF3^{-/-}$) present more

281 severe phenotypes than those of either single knockout (GDF1$^{-/-}$ or GDF3$^{-/-}$
282)(Andersson et al., 2007). Detailed comparisons of the developmental roles of
283 the GDF1 and GDF3 genes between mammals and amphibians will shed light
284 regarding the reinvention of genes that possess fundamental roles during
285 early development. On the other hand, comparing mammals and amphibians
286 with sauropsids will provide useful information regarding the evolutionary fate
287 of duplicated genes. Finally, it would be interesting to study the evolutionary
288 history of genes that cooperate with GDF1 and GDF3 during early
289 development (e.g. nodal) (Ramsdell and Yost 1998; Levine 2005; Komatsu
290 and Mishina 2013) in order to understand the evolutionary nature of the entire
291 developmental network.
292
293 **Material and Methods**
294 **DNA sequences and phylogenetic analyses**
295 We annotated GDF1 and GDF3 genes in representative species of chordates.
296 Our study included representative species from mammals, birds, reptiles,
297 amphibians, coelacanths, holostean fish, teleost fish, cartilaginous fish,
298 cyclostomes, urochordates and cephalochordates (Supplementary Table S1).
299 We identified genomic pieces containing GDF 1 and GDF3 genes in the
300 Ensembl database using BLASTN with default settings or NCBI database
301 (refseq_genomes, htgs, and wgs) using tbalstn (Altschul et al., 1990) with
302 default settings. Conserved synteny was also used as a criterion to define the
303 genomic region containing GDF1 and GDF3 genes. Once identified, genomic
304 pieces were extracted including the 5´and 3´ flanking genes. After extraction,
305 we curated the existing annotation by comparing known exon sequences to
306 genomic pieces using the program Blast2seq with default parameters
307 (Tatusova and Madden 1999). Putatively functional genes were characterized
308 by an open intact reading frame with the canonical exon/intron structure
309 typical of vertebrate GDF1 and GDF3 genes. Sequences derived from shorter
310 records based on genomic DNA or cDNA were also included in order to attain
311 a broad and balanced taxonomic coverage. Amino acid sequences were
312 aligned using the L-INS-i strategy from MAFFT v.7 (Katoh and Standley
313 2013). Phylogenetic relationships were estimated using maximum likelihood
314 and Bayesian approaches. We used the proposed model tool from IQ-Tree

315 (Trifinopoulos et al., 2016) to select the best-fitting model (JTT+F+R5).
316 Maximum likelihood analysis was also performed in IQ-Tree (Trifinopoulos et
317 al., 2016) to obtain the best tree. Node support was assessed with 1000
318 bootstrap pseudoreplicates using the ultrafast routine. Bayesian searches
319 were conducted in MrBayes v.3.1.2 (Ronquist and Huelsenbeck 2003). Two
320 independent runs of six simultaneous chains for $10x10^6$ generations were set,
321 and every 2,500 generations were sampled using default priors. The run was
322 considered to have reached convergence once the likelihood scores formed
323 an asymptote and the average standard deviation of the split frequencies
324 remained < 0.01. We discarded all trees that were sampled before
325 convergence, and we evaluated support for the nodes and parameter
326 estimates from a majority rule consensus of the last 2,000 trees. Sea squirt
327 (*Ciona intestinalis*) and Florida lancelet (*Branchiostoma floridae*) BMP2
328 sequences were used as outgroups.
329
330 **Assessment of conserved synteny**
331 We examined genes found up- and downstream of GDF1 and GDF3 in
332 species representative of vertebrates. Synteny assessment were conducted
333 for human (*Homo sapiens*), chicken (*Gallus gallus*), western-clawed frog
334 (*Xenopus tropicalis*), coelacanth (*Latimeria chalumnae*), spotted gar
335 (*Lepisosteus oculatus*), elephant shark (*Callorhinchus milii*) and sea lamprey
336 (*Petromyzon marinus*). Initial ortholog predictions were derived from the
337 EnsemblCompara database (Herrero et al., 2016) and were visualized using
338 the program Genomicus v91.01(Louis et al., 2015). In other cases, the
339 genome data viewer platform from the National Center for Biotechnology
340 information was used.
341
342 **Acknowledgements**
343 This work was supported by the Fondo Nacional de Desarrollo Científico y
344 Tecnológico (FONDECYT 1160627) grant to JCO.
345
346
347
348

351 **Figure legends**

352

353 **Figure 1.** Maximum likelihood tree depicting evolutionary relationships among
354 the GDF1 and GDF3 genes of chordates. Numbers on the nodes represent
355 maximum likelihood ultrafast bootstrap and Bayesian posterior probability
356 support values. BMP2 sequences from urochordates (*Ciona intestinalis*) and
357 cephalochordates (*Branchiostoma floridae*) were used as outgroups.

358

359 **Figure 2.** Dot-plots of pairwise sequence similarity between the GDF1 and
360 GDF3 genes of the western clawed frog (*Xenopus tropicalis*) and the
361 corresponding syntenic region in the American alligator (*Alligator*
362 *mississippiensis*), Burmese python (*Python bivittatus*), chicken (*Gallus gallus*)
363 and green turtle (*Chelonia mydas*).

364

365 **Figure 3.** Schematic representations of alternative hypotheses of the sister
366 group relationships among duplicated GDF genes in tetrapods. A) According
367 to the one-duplication model the predicted phylogeny recovers a monophyletic
368 group containing GDF1 sequences from mammals, sauropsids and
369 amphibians sister to a clade containing GDF3 sequences from the same
370 groups. B) The phylogenetic prediction from the two-duplication model
371 retrieves a clade containing GDF1 and GDF3 sequences from mammals
372 sister to a clade containing GDF1/3 sequences from sauropsids; additionally a
373 clade containing GDF1 and GDF3 sequences from amphibians is recovered
374 sister to the mammalian/sauropsid clade.

375

376 **Figure 4.** Structure of the chromosomal region containing the GDF1 and
377 GDF3 genes of vertebrates. Asterisks denote that the orientation of the
378 genomic piece is from 3' to 5', gray lines represent intervening genes that do
379 not contribute to conserved synteny.

380

381 **Figure 5.** An evolutionary hypothesis of the evolution of the GDF1 and GDF3
382 genes in vertebrates. According to this model the last common ancestor of
383 vertebrates had a repertoire of one gene (GDF1/3), a condition that has been
384 maintained in actual species of cylostomes, chondrichthyes, bony fish,

coelacanths and sauropsids. In the ancestor of tetrapods, the single gene copy was translocated to a different chromosomal location. After that, in the amphibian ancestor, the single gene copy underwent a duplication event giving rise to the amphibian GDF1 (Vg1) and GDF3 (derrière) genes. In the ancestor of mammals, the single gene copy also underwent a duplication event, giving rise to the mammalian GDF1 and GDF3 genes. After the duplication, but before the radiation of the group, a second translocation event occurred in the ancestor of the group. Thus, all mammals inherited a repertoire of two genes (GDF1 and GDF3) that were located on two chromosomes.

397 **References**
398

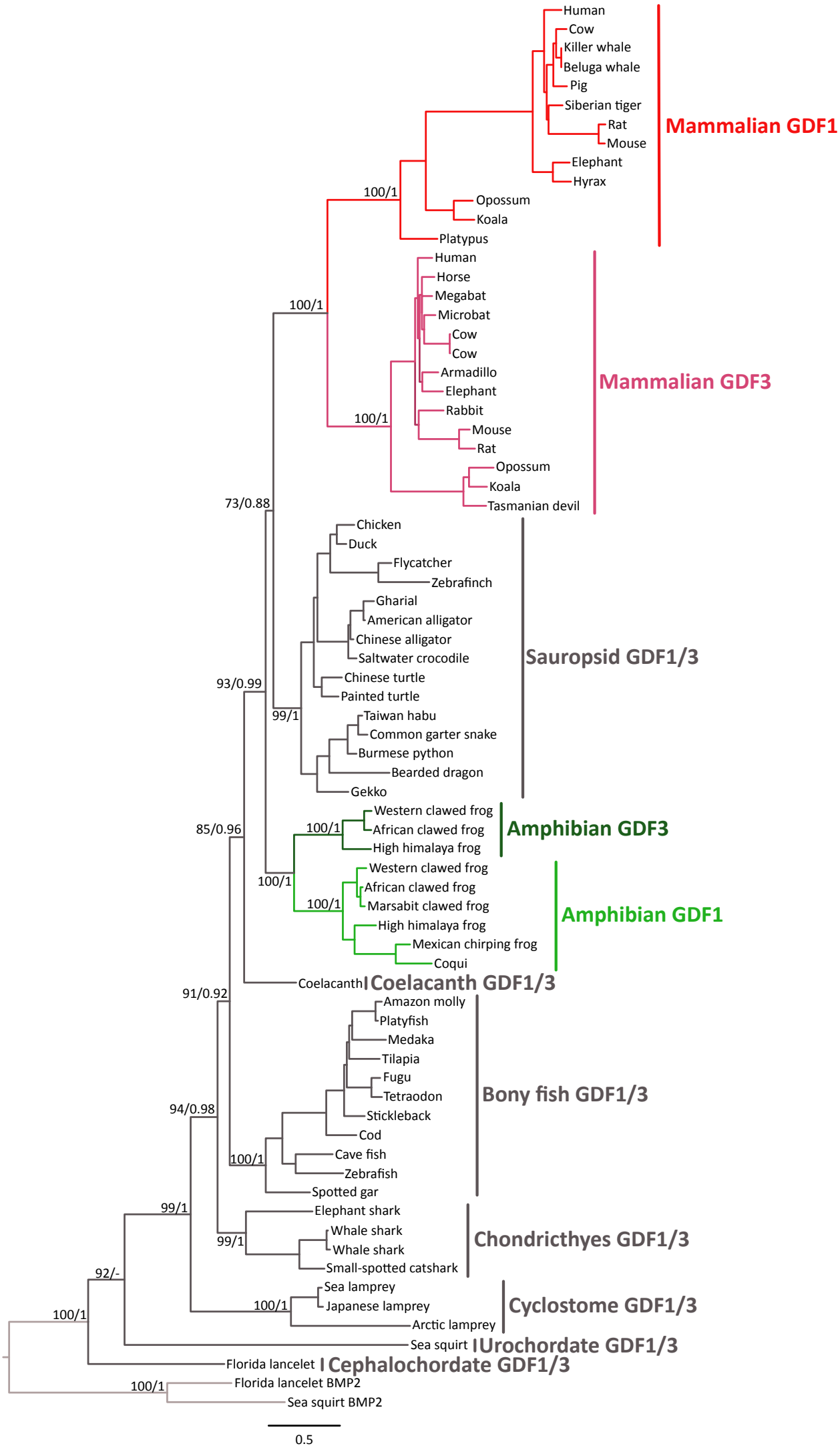
399 Andersson O, Bertolino P, Ibáñez CF. 2007. Distinct and cooperative roles of
400 mammalian Vg1 homologs GDF1 and GDF3 during early embryonic
401 development. Dev. Biol. 311:500–511.
402 Andersson O, Korach-Andre M, Reissmann E, Ibanez CF, Bertolino P,
403 Lefkowitz RJ. 2008. Growth/differentiation factor 3 signals through ALK7
404 and regulates accumulation of adipose tissue and diet-induced obesity.
405 Proc. Natl. Acad. Sci. USA 105:7252–7256.
406 Andersson O, Reissmann E, Jörnvall H, Ibáñez CF. 2006. Synergistic
407 interaction between Gdf1 and Nodal during anterior axis development.
408 Dev. Biol. 293:370–381.
409 Arias CF, Herrero MA, Stern CD, Bertocchini F. 2017. A molecular
410 mechanism of symmetry breaking in the early chick embryo. Scietific
411 Reports 7:1–6.
412 Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local
413 alignment search tool. Journal of Molecular Biology 215:403–410.
414 Bao MW, Zhang XJ, Li L, Cai Z, Liu X, Wan N, Hu G, Wan F, Zhang R, Zhu X,
415 et al., 2015. Cardioprotective role of growth/differentiation factor 1 in post-
416 infarction left ventricular remodelling and dysfunction. J. Pathol. 236:360–
417 372.
418 Bengtsson H, Epifantseva I, Åbrink M, Kylberg A, Kullander K, Ebendal T,
419 Usoskin D. 2008. Generation and characterization of a Gdf1 conditional
420 null allele. Genesis 46:368–372.
421 Bisgrove BW, Su Y-C, Yost HJ. 2017. Maternal Gdf3 is an obligatory cofactor
422 in Nodal signaling for embryonic axis formation in zebrafish. Elife
423 6:e28534.
424 Chen C. 2006. The Vg1-related protein Gdf3 acts in a Nodal signaling
425 pathway in the pre-gastrulation mouse embryo. Development 133:319–
426 329.
427 Dale ' L, Matthews G, Colman2 A. 1993. Secretion and mesoderm-inducing
428 activity of the TGF-j-related domain of Xenopus Vgl. EMBO J. 12:4471–
429 4480.
430 Goodman M, Czelusniak J, Koop BF, Tagle DA, Slightom JL. 1987. Globins: a

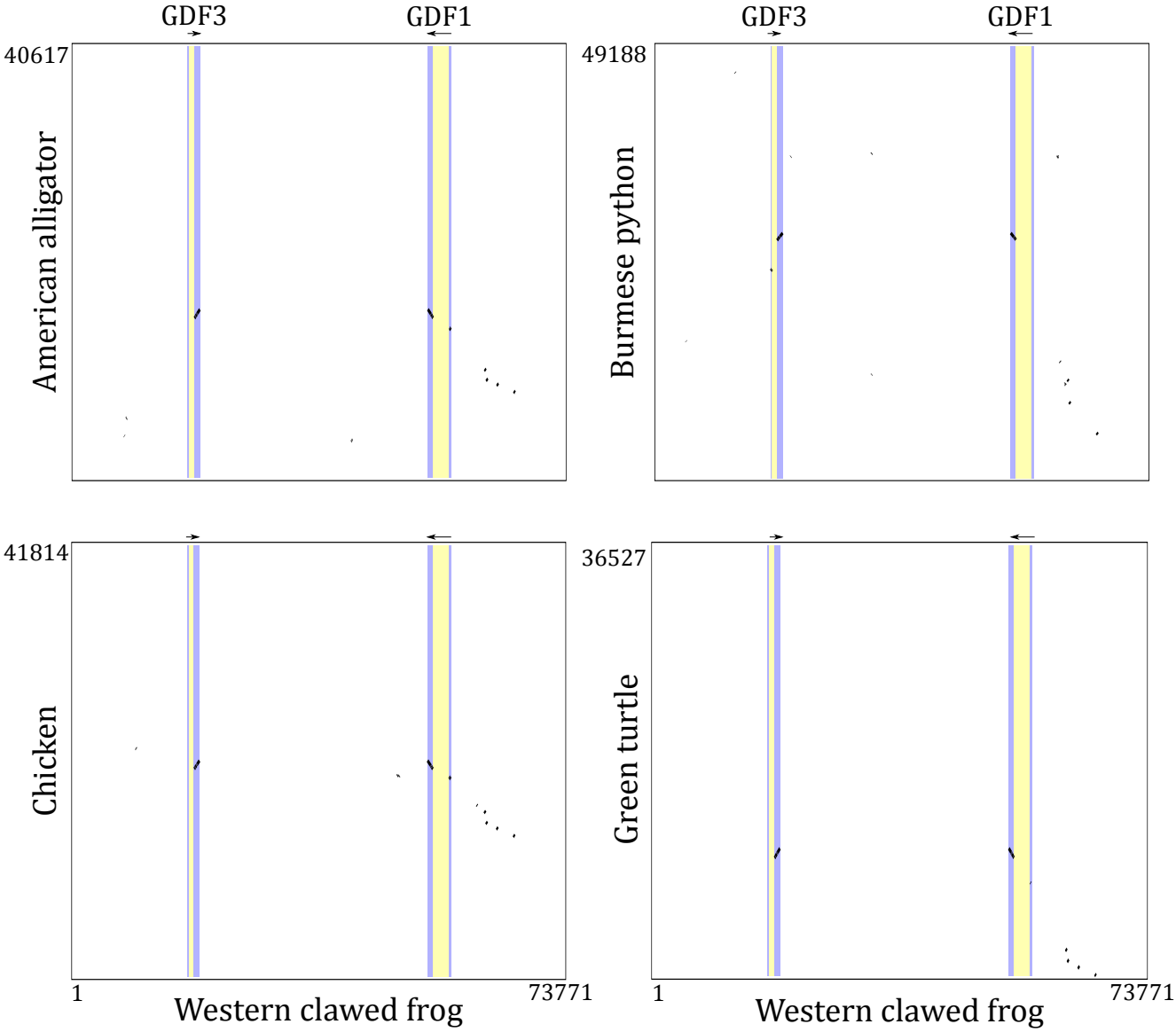
431 case study in molecular phylogeny. Cold Spring Harb. Symp. Quant. Biol.
432 52:875–890.
433 Hamada H, Meno C, Watanabe D, Saijoh Y. 2002. Establishment of
434 vertebrate left-right asymmetry. Nat. Rev. Genet. 3:103–113.
435 Herrero J, Muffato M, Beal K, Fitzgerald S, Gordon L, Pignatelli M, Vilella AJ,
436 Searle SMJ, Amode R, Brent S, et al., 2016. Ensembl comparative
437 genomics resources. Database 2016:baw053.
438 Hoffmann FG, Storz JF, Gorr TA, Opazo JC. 2010. Lineage-specific patterns
439 of functional diversification in the α-and β-globin gene families of tetrapod
440 vertebrates. Mol. Biol. Evol. 27:1126–1138.
441 Hoffmann FG, Vandewege MW, Storz JF, Opazo JC. 2018. Gene Turnover
442 and Diversification of the α- and β-Globin Gene Families in Sauropsid
443 Vertebrates. Genome Biol. Evol. 10:344–358.
444 Hyatt BA, Lohr JL, Yost HJ. 1996. Initiation of vertebrate left–right axis
445 formation by maternal Vg1. Nature 384:62–65.
446 Hyatt BA, Yost HJ. 1998. The left-right coordinator: The role of Vg1 in
447 organizing left-right axis formation. Cell 93:37–46.
448 Jones CM, Simon-Chazottes D, Guenet J-L, Hogan BLM, Pasteur L. 1992.
449 Isolation of Vgr-2, a Novel Member of the Transforming Growth Factor-P-
450 Related Gene Family. Mol. Endocrinol. 6:1961–1968.
451 Kaasinen E, Aittomäki K, Eronen M, Vahteristo P, Karhu A, Mecklin JP,
452 Kajantie E, Aaltonen LA, Lehtonen R. 2010. Recessively inherited right
453 atrial isomerism caused by mutations in growth/differentiation factor 1
454 (GDF1). Hum. Mol. Genet. 19:2747–2753.
455 Karkera JD, Lee JS, Roessler E, Banerjee-Basu S, Ouspenskaia MV, Mez J,
456 Goldmuntz E, Bowers P, Towbin J, Belmont JW, et al., 2007. Loss-of-
457 Function Mutations in Growth Differentiation Factor-1 (GDF1) Are
458 Associated with Congenital Heart Defects in Humans. Am. J. Hum.
459 Genet. 81:987–994.
460 Katoh K, Standley DM. 2013. MAFFT Multiple Sequence Alignment Software
461 Version 7: Improvements in Performance and Usability. Mol. Biol. Evol.
462 30:772–780.
463 Komatsu Y, Mishina Y. 2013. Establishment of left – right asymmetry in
464 vertebrate development : the node in mouse embryos. Cell. Mol. Life Sci.

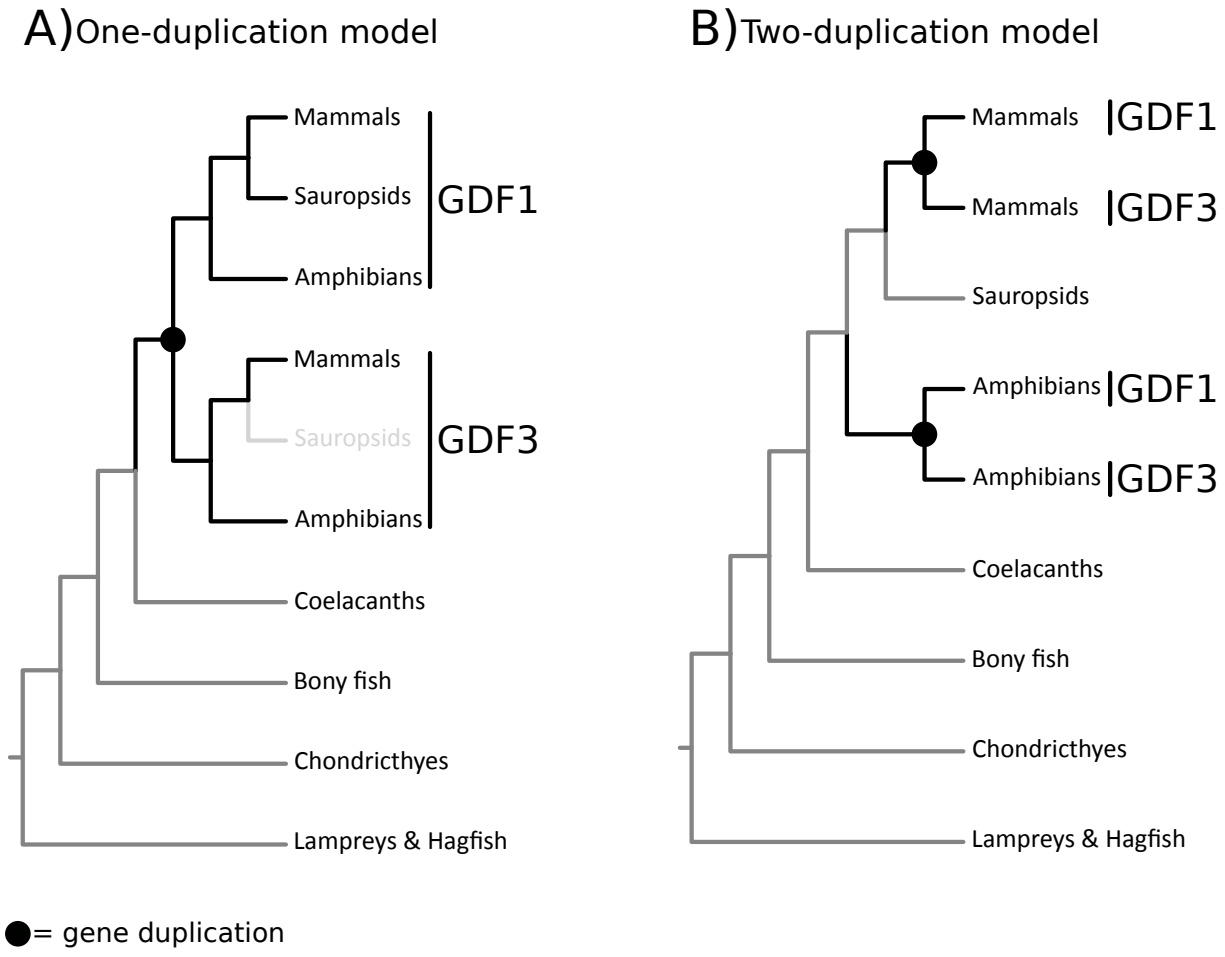
465 70:4659–4666.
466 Kriener K, O ’huigin C, Klein J. 2001. Independent Origin of Functional MHC
467 Class II Genes in Humans and New World Monkeys. Hum. Immunol.
468 62:1–14.
469 Lee S-J. 1990. Identification of a Novel Member (GDF-1) of the Transforming
470 Growth Factor-beta Superfamily. Mol. Endocrinol. 4:1034–1040.
471 Lee S-J. 1991. Expression of growth/differentiation factor 1 in the nervous
472 system: Conservation of a bicistronic structure. Biochemistry 88:4250–
473 4254.
474 Levine AJ. 2005. GDF3, a BMP inhibitor, regulates cell fate in stem cells and
475 early embryos. Development 133:209–216.
476 Levine AJ, Brivanlou AH, Suzuki A, Hemmati-Brivanlou A. 2006. GDF3, a
477 BMP inhibitor, regulates cell fate in stem cells and early embryos.
478 Development 133:209–216.
479 Li Y, Ye C, Shi P, Zou X-J, Xiao R, Gong Y-Y, Zhang Y-P. 2005. Independent
480 origin of the growth hormone gene family in New World monkeys and Old
481 World monkeys/hominoids. J. Mol. Endocrinol. 35:399–409.
482 Louis A, Nguyen NTT, Muffato M, Roest Crollius H. 2015. Genomicus update
483 2015: KaryoView and MatrixView provide a genome-wide perspective to
484 multispecies comparative genomics. Nucleic Acids Res. 43:D682–D689.
485 Mercola M, Levin M. 2001. Left-right assymetry determination in vertebrates.
486 Annu. Rev. Cell Dev. Biol. 17:779–805.
487 Montague TG, Schier AF. 2017. Vg1-Nodal heterodimers are the endogenous
488 inducers of mesendoderm. Elife 6:e28183.
489 Nakamura T, Hamada H. 2012. Left-right patterning: conserved and divergent
490 mechanisms. Development 139:3262.
491 Opazo JC, Hoffmann FG, Storz JF. 2008. Genomic evidence for independent
492 origins of β-like globin genes in monotremes and therian mammals. Proc.
493 Natl. Acad. Sci. USA 105:1590–1595.
494 Ramsdell AF, Yost HJ. 1998. Molecular mechanisms of vertebrate left–right
495 development. Trends Genet. 14:459–465.
496 Rankin CT, Bunton T, Lawler AM, Lee SJ. 2000. Regulation of left-right
497 patterning in mice by growth/differentiation factor-1. Nat. Genet. 24:262–
498 265.

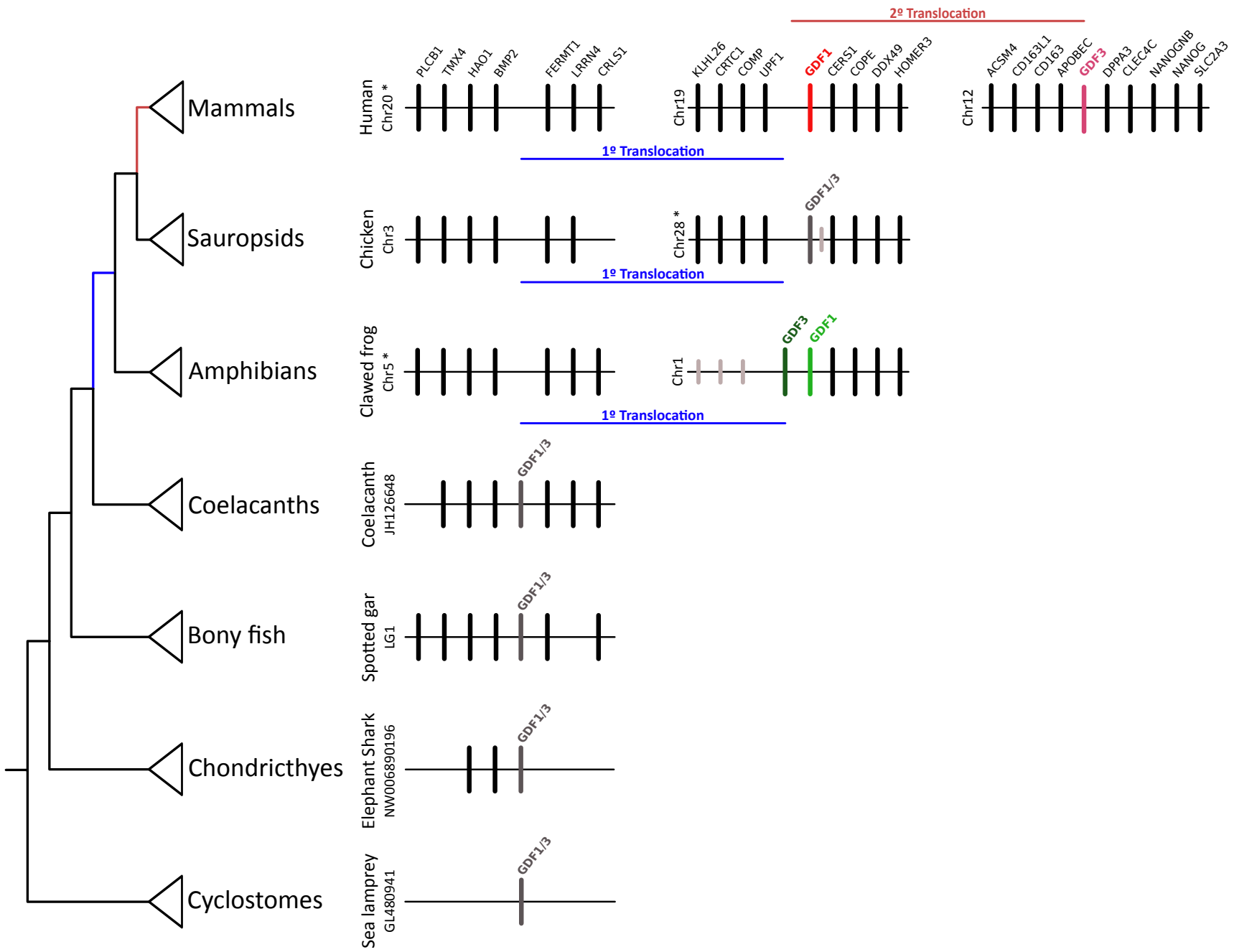
499 Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic
500 inference under mixed models. Bioinformatics 19:1572–1574.
501 Seleiro EAP, Connolly DJ, Cooke J. 1996. Early developmental expression
502 and experimental axis determination by the chicken Vg1 gene. Curr. Biol.
503 6:1476–1486.
504 Söderström S, Ebendal T. 1999. Localized expression of BMP and GDF
505 mRNA in the rodent brain. J. Neurosci. Res. 56:482–492.
506 Somi S, Houweling AC, Buffing AAM, Moorman AFM, Van Den Hoff MJB.
507 2003. Expression of cVg1 mRNA during chicken embryonic development.
508 Anat. Rec. 273A:603–608.
509 Suzuki A, Uno Y, Takahashi S, Grimwood J, Schmutz J, Mawaribuchi S,
510 Yoshida H, Takebayashi-Suzuki K, Ito M, Matsuda Y, et al., 2017.
511 Genome organization of the vg1 and nodal3 gene clusters in the
512 allotetraploid frog *Xenopus laevis*. Dev. Biol. 426: 236-244.
513 Tatusova, TA, Madden TL. 1999. BLAST 2 sequences, a new tool for
514 comparing protein and nucleotide sequences. FEMS Microbiology Letters
515 174:247-250.
516 Thomsen GH, Melton DA. 1993. Processed Vg1 protein is an axial mesoderm
517 inducer in xenopus. Cell 74:433–441.
518 Trifinopoulos J, Nguyen L-T, von Haeseler A, Minh BQ. 2016. W-IQ-TREE: a
519 fast online phylogenetic tool for maximum likelihood analysis. Nucleic
520 Acids Res. 44:W232–W235.
521 Wall NA, Craig EJ, Labosky PA, Kessler DS. 2000. Mesendoderm induction
522 and reversal of left-right pattern by mouse Gdf1, a Vg1-related gene.
523 Dev. Biol. 227:495–509.
524 Wallis OC, Wallis M. 2002. Characterisation of the GH gene cluster in a new-
525 world monkey, the marmoset (*Callithrix jacchus*). J. Mol. Endocrinol.
526 29:89–97.
527 Wang W, Yang Y, Meng Y, Shi Y. 2004. GDF-3 is an adipogenic cytokine
528 under high fat dietary condition. Biochem. Biophys. Res. Commun.
529 321:1024–1031.
530 Witthuhn BA, Bernlohr DA. 2001. Upregulation of bone morphogenetic protein
531 GDF-3/Vgr-2 expression in adipose tissue of FABP4/aP2 null mice.
532 Cytokine 14:129–135.

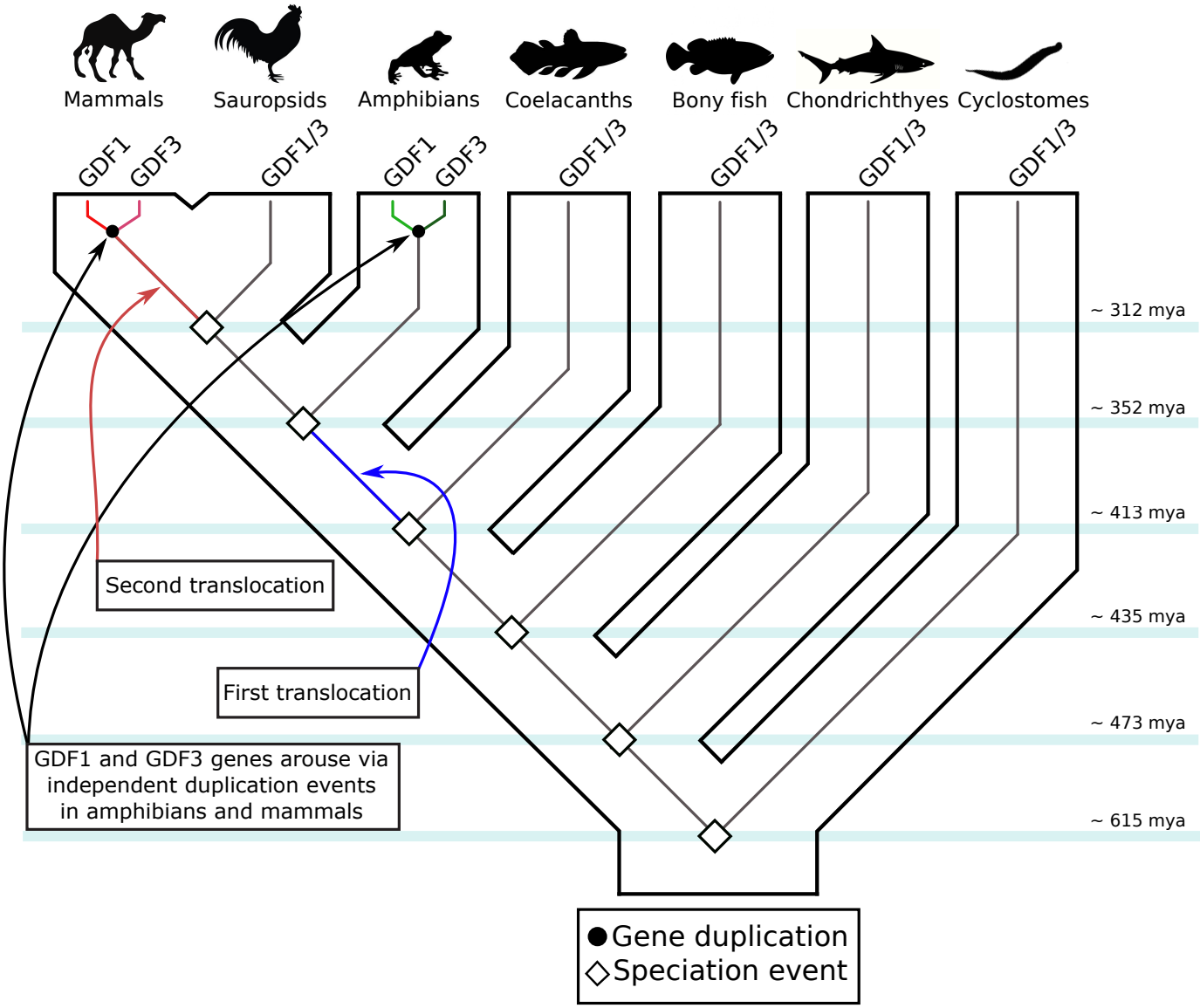
533 Wright CVE. 2001. Mechanisms of Left-Right Asymmetry : What ' s Right and
534 What ' s Left ? Cell 1:179–186.
535 Yang W, Mok MTS, Li MSM, Kang W, Wang H, Chan AW, Chou JL, Chen J,
536 Ng EKW, To KF, et al., 2016. Epigenetic silencing of GDF1 disrupts
537 SMAD signaling to reinforce gastric cancer development. Oncogene
538 35:2133–2144.
539 Zhang J, Wu Q, Wang L, Li X, Ma Y, Yao L. 2015. Association of GDF1
540 rs4808863 with fetal congenital heart defects: A case-control study. BMJ
541 Open 5:e009352.
542
543
544
545
546
547


Human
Cow
Killer whale
Beluga whale
Pig
Siberian tiger
Rat
Mouse
Elephant
Hyrax
Opossum
Koala
Platypus
Mammalian GDF1
100/1
Human
Horse
Megabat
Microbat
Cow
Cow
Armadillo
Elephant
Rabbit
Mouse
Rat
Opossum
Koala
Tasmanian devil
Mammalian GDF3
100/1
100/1
73/0.88
Chicken
Duck
Flycatcher
Zebrafinch
Gharial
American alligator
Chinese alligator
Saltwater crocodile
Chinese turtle
Painted turtle
Taiwan habu
Common garter snake
Burmese python
Bearded dragon
Gekko
Sauropsid GDF1/3
93/0.99
99/1
Western clawed frog
African clawed frog
High himalaya frog
Amphibian GDF3
85/0.96
100/1
100/1
Western clawed frog
African clawed frog
Marsabit clawed frog
High himalaya frog
Mexican chirping frog
Coqui
Amphibian GDF1
100/1
Coelacanth
Coelacanth GDF1/3
91/0.92
Amazon molly
Platyfish
Medaka
Tilapia
Fugu
Tetraodon
Stickleback
Cod
Cave fish
Zebrafish
Spotted gar
Bony fish GDF1/3
94/0.98
100/1
99/1
Elephant shark
Whale shark
Whale shark
Small-spotted catshark
Chondricthyes GDF1/3
99/1
92/-
Sea lamprey
Japanese lamprey
Arctic lamprey
Cyclostome GDF1/3
100/1
100/1
Sea squirt
Urochordate GDF1/3
Florida lancelet
Cephalochordate GDF1/3
100/1
Florida lancelet BMP2
Sea squirt BMP2
0.5


GDF3
GDF1
40617
American alligator
GDF3
GDF1
49188
Burmese python
41814
Chicken
1
73771
Western clawed frog
36527
Green turtle
1
73771
Western clawed frog


A) One-duplication model
Mammals
Sauropsids
Amphibians
GDF1
Mammals
Sauropsids
Amphibians
GDF3
Coelacanths
Bony fish
Chondricthyes
Lampreys & Hagfish
B) Two-duplication model
Mammals
GDF1
Mammals
GDF3
Sauropsids
Amphibians
GDF1
Amphibians
GDF3
Coelacanths
Bony fish
Chondricthyes
Lampreys & Hagfish
●= gene duplication


Mammals
Sauropsids
Amphibians
Coelacanths
Bony fish
Chondricthyes
Cyclostomes
Human
Chr20 *
PLCB1
TMX4
HAO1
BMP2
FERMT1
LRRN4
CRLS1
Chr19
KLHL26
CRTC1
COMP
UPF1
GDF1
CERS1
COPE
DDX49
HOMER3
Chr12
ACSM4
CD163L1
CD163
APOBEC
GDF3
DPPA3
CLEC4C
NANOGNB
NANOG
SLC2A3
2º Translocation
1º Translocation
Chicken
Chr3
Chr28 *
GDF1/3
1º Translocation
Clawed frog
Chr5 *
Chr1
GDF3
GDF1
1º Translocation
Coelacanth
JH126648
GDF1/3
Spotted gar
LG1
GDF1/3
Elephant Shark
NW006890196
GDF1/3
Sea lamprey
GL480941
GDF1/3


Mammals
Sauropsids
Amphibians
Coelacanths
Bony fish
Chondrichthyes
Cyclostomes
GDF1
GDF3
GDF1/3
GDF1
GDF3
GDF1/3
GDF1/3
GDF1/3
GDF1/3
~ 312 mya
~ 352 mya
~ 413 mya
~ 435 mya
~ 473 mya
~ 615 mya
Second translocation
First translocation
GDF1 and GDF3 genes arouse via independent duplication events in amphibians and mammals
Gene duplication
Speciation event