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 [1,2]. 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 [3–20]. 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 [17,21–25]. 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 [26], and GDF3 has been associated with the regulation of adipose tissue homeostasis and energy balance during nutrient overload [27–29].

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 [3,4,13–20,5–12]. 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 [15,17,21]. 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 [30]. 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 [30,31]. Given this scenario, the single copy gene found in amphibians and birds would be a co-ortholog of the mammalian genes [30]. 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

Phylogenetic analyses suggest an independent origin of the GDF1 and GDF3 genes in mammals and amphibians

From an evolutionary perspective, the definition of homologous relationships among GDF1 and GDF3 genes is still a matter of debate [30,31]. The resolution of this homology is important as these genes are involved in fundamental developmental processes which are conserved all across vertebrates [8,32–34]. Thus, if extant species inherited these genes from the vertebrate ancestor, the developmental processes in which GDF1 and GDF3 are involved have a single evolutionary origin and are comparable among species.

Based on evolutionary analyses and the developmental processes in which GDF1 has been shown to participate, it has been suggested that GDF1 in mammals, birds, and amphibians are 1:1 orthologs [15,17,21,30,35]. Given that it has not been possible to identify copies of the GDF3 gene in amphibians and birds, it has been proposed that this gene is an evolutionary innovation of

112 mammals, and the single gene copy found in amphibians and birds is co-
113 ortholog to the mammalian duplicates [30]. However, a second copy of a GDF
114 gene has been annotated in the genome of the western clawed frog (*Xenopus*
115 *tropicalis*) [36]. This newly described gene is located at the 3´ side of the GDF1
116 gene and indicates that amphibians, like mammals, also have a repertoire of
117 two GDF genes (GDF1 (Vg1) and GDF3 (derrière)). This new finding further
118 complicates the resolution of homologous relationships and the origin of
119 genes involved in the early development of all vertebrates [8,32–34].
120 Our maximum likelihood and Bayesian reconstructions revealed that
121 the GDF1 and GDF3 genes of amphibians and mammals were both
122 reciprocally monophyletic (Fig. 1), indicating that these genes originated
123 independently in each of these two lineages (Fig. 1). In agreement with the
124 literature, we found only one gene in sauropsids (e.g. birds, turtles, lizards
125 and allies), suggesting that they retained the ancestral condition of a single
126 gene copy as found in non-tetrapod vertebrates (e.g. coelacanths, bony fish,
127 chondricthyes and cyclostomes)(Fig. 1). Dot-plot comparisons provided
128 further support for the presence of a single gene copy in sauropsids as no
129 traces of an extra GDF gene were present in the syntenic region of
130 representative species of the group (Fig. 2). However, we cannot rule out an
131 alternative scenario of duplication and subsequent gene loss in the ancestor
132 of sauropsids. Thus, our evolutionary analyses suggest that the GDF1 and
133 GDF3 genes present in amphibians and mammals diversified independently
134 in the ancestor of each of these lineages. In other words, genes that
135 participate in fundamental processes during early development have been
136 reinvented two independent times during the evolutionary history of tetrapods.
137 As a consequence of the independent origin of these genes in amphibians
138 and mammals, they are not 1:1 orthologs.
139 To further test our hypothesis of the independent origins of the GDF1
140 and GDF3 genes in amphibians and mammals we performed topology tests.
141 In these analyses we compared our phylogenetic tree (Fig. 1 and Fig. 3B) to
142 the topology predicted from a one-duplication model in which the duplication
143 event that gave rise to the GDF1 and GDF3 genes in amphibians and
144 mammals occurred in the ancestor of tetrapods (Fig. 3A). In the one-
145 duplication model, it is assumed that the ancestor of sauropsids had two gene

146 copies and that subsequently one of these was lost. Thus, according to the
147 one-duplication model (Fig. 3A), the predicted phylogeny would recover a
148 monophyletic group containing GDF1 sequences from mammals, sauropsids,
149 and amphibians sister to a clade containing GDF3 sequences from the same
150 groups (Fig. 3A). Alternatively, the predicted phylogeny from the two-
151 duplication model would retrieve a clade containing GDF1 and GDF3
152 sequences from mammals sister to a clade containing GDF1/3 sequences
153 from sauropsids; additionally a clade containing GDF1 and GDF3 sequences
154 from amphibians would be recovered sister to the mammalian/sauropsid clade
155 (Fig. 3B). Results of the topology tests rejected the phylogeny predicted by
156 the one-duplication model (Weighted Shimodaira and Hasegawa, $p< 10^{-4}$;
157 Weighted Kishino Hasegawa, $p< 10^{-4}$). Thus, this result provided additional
158 support to our hypothesis that the GDF1 and GDF3 genes of amphibians and
159 mammals are the product of lineage independent duplication events in the
160 ancestors of each of these groups.
161 In the literature there are other cases of groups of genes that perform
162 similar biological functions in a diversity of species but that have originated via
163 lineage independent duplication events [37–43]. Among these, the independent
164 origin of the β-globin gene cluster in all main groups of tetrapods (e.g. therian
165 mammals, monotremes, birds, crocodiles, turtles, squamates, amphibians)
166 represents a well-documented phenomenon [41–43]. This case is of particular
167 interest as the two β-globin subunits that come from a gene family that has
168 been reinvented several times during the evolutionary history of tetrapods are
169 assembled in a tetramer with two α-globin subunits that belong to a group of
170 genes that possess a single origin [37,41–43]. Besides this example, the
171 independent origin of gene families makes the task of comparison difficult, as
172 the repertoire of genes linked to physiological processes in different lineages
173 do not have the same evolutionary origin. Additionally, the fact that the
174 evolutionary process can give rise to similar phenotypes following different
175 mutational pathways makes the problem of comparing even more difficult
176 (Natarajan et al., 2016). This is particularly important when extrapolating the
177 results of physiological studies performed in model species to other
178 organisms.
179

180 ***Two translocation events during the evolutionary history of the GDF1***
181 ***and GDF3 genes of tetrapods***
182 Based on the chromosomal distribution of the GDF1 and GDF3 genes and the
183 conservation pattern of flanking genes, we propose that during the
184 evolutionary history of tetrapods, these genes underwent two chromosomal
185 translocation events (Fig. 4). According to our analyses the genomic region
186 that harbors the single gene copy of chondricthyes, bony fish, and
187 coelacanths is conserved, yet this region in these groups differs from the
188 regions in which GDF1 and GDF3 are located in tetrapods. Overall this
189 suggests that the first translocation event occurred in the ancestor of
190 tetrapods (Fig. 4). Interestingly, the genomic region where the single gene
191 copy is located in non-tetrapod vertebrates, which is defined by the presence
192 of upstream (BMP2, HAO1, TMX4 and PLCB1) and downstream genes
193 (FERMT1, LRRN4, CRLS1), is conserved in tetrapods. Given this, it is
194 possible to identify the chromosomal location where the gene was located in
195 the tetrapod ancestor before the first translocation event (Fig. 4). On the other
196 hand, the fact that the mammalian GDF1 and GDF3 genes are located in
197 different chromosomes suggests that the second translocation event occurred
198 in the ancestor of the group (Fig. 4). In humans, the GDF1 gene is located on
199 chromosome 19 while GDF3 is located on chromosome 12. In opossum
200 (*Didelphis virginiana*) the GDF1 and GDF3 genes are located on
201 chromosomes 3 and 8, respectively.
202
203 ***Evolution of GDF1 and GDF3 genes in vertebrates***
204 In this study we present compelling evidence suggesting that genes involved
205 in the formation of the primitive streak, anterior visceral endoderm, mesoderm
206 and the establishment of the left-right identity [3,4,14–17,20,30,5–7,9–13] in amphibians
207 and mammals are the product of independent duplications events. As such,
208 these results indicate that the GDF1 and GDF3 genes of amphibians and
209 mammals are not 1:1 orthologs.
210 Thus, according to our results the last common ancestor of vertebrates
211 had a repertoire of one gene (GDF1/3), and this condition is maintained in
212 actual species of cylostomes, chondricthyes, bony fish, and coelacanths (Fig.
213 5). In the ancestor of tetrapods, the single gene copy was translocated from a

214 chromosomal region defined by the presence of the BMP2, HAO1, TMX4,
215 PLCB1, FERMT1, LRRN4, CRLS1 genes to a chromosomal region defined by
216 the presence of the CERS1, COPE, DDX49 and HOMER3 genes (Fig. 4).
217 After this translocation, the ancestral gene underwent a duplication event in
218 the amphibian ancestor, giving rise to the GDF1 (Vg1) and GDF3 (derrière)
219 genes as they are found in actual species. In the case of the western clawed
220 frog (*Xenopus tropicalis*) these genes are located in tandem on chromosome
221 28 (Fig. 4). Sauropsids, the group that includes birds, crocodiles, turtles,
222 lizards and snakes, inherited the ancestral condition of a single gene copy as
223 is seen in non-tetrapod vertebrates (Fig. 5). Finally, in the ancestor of
224 mammals, the ancestral gene also underwent a duplication event, giving rise
225 to the GDF1 and GDF3 genes found in extant species of mammals. After
226 duplication but before the radiation of the group, a second translocation event
227 occurred in the ancestor of the group (Fig. 5). Thus, all mammals inherited a
228 repertoire of two genes (GDF1 and GDF3) that are located on two
229 chromosomes.
230
231 **Concluding remarks**
232 This study provides a comprehensive evolutionary analysis of the GDF1 and
233 GDF3 genes in representative species of all main groups of vertebrates. The
234 main focus of this study was to unravel the duplicative history of the GDF1
235 and GDF3 genes and to understand homologous relationships among
236 vertebrates. Understanding homology in this case is particularly important as
237 these genes perform fundamental roles during early development that are
238 conserved across vertebrates [8,32–34]. Surprisingly, our results revealed that the
239 GDF1 and GDF3 genes present in amphibians and mammals are the product
240 of independent duplication events in the ancestor of each of these groups.
241 Subsequently, the GDF1 and GDF3 genes of amphibians and mammals are
242 not 1:1 orthologs. Our results also show that all other vertebrate groups - i.e
243 non-tetrapods and sauropsids – maintained the ancestral condition of a single
244 gene copy (GDF1/3).
245 From an evolutionary perspective the independent duplication events
246 that occurred in the ancestors of mammals and amphibians could have
247 resulted in the division of labor, with some degree of redundancy, of the

248 function performed by the ancestral gene. In support of this idea, it has been
249 shown that the GDF1 and GDF3 genes of mammals have partially redundant
250 functions during development where GDF1 can to some degree compensate
251 for the lack of GDF3 [30]. Additionally, it has also been shown that double
252 knockout animals (GDF1$^{-/-}$ and GDF3$^{-/-}$) present more severe phenotypes
253 than those of either single knockout (GDF1$^{-/-}$ or GDF3$^{-/-}$)[30]. Detailed
254 comparisons of the developmental roles of the GDF1 and GDF3 genes
255 between mammals and amphibians will shed light regarding the reinvention of
256 genes that possess fundamental roles during early development. On the other
257 hand, comparing mammals and amphibians with sauropsids will provide
258 useful information regarding the evolutionary fate of duplicated genes. Finally,
259 it would be interesting to study the evolutionary history of genes that
260 cooperate with GDF1 and GDF3 during early development (e.g. nodal) [8,33,44]
261 in order to understand the evolutionary nature of the entire developmental
262 network.
263
264 **Material and Methods**
265 **DNA sequences and phylogenetic analyses**
266 We annotated GDF1 and GDF3 genes in representative species of chordates.
267 Our study included representative species from mammals, birds, reptiles,
268 amphibians, coelacanths, holostean fish, teleost fish, cartilaginous fish,
269 cyclostomes, urochordates and cephalochordates (Supplementary dataset 1
270 and 2). We identified genomic pieces containing GDF 1 and GDF3 genes in
271 the Ensembl database using BLASTN with default settings or NCBI database
272 (refseq_genomes, htgs, and wgs) using tbalstn (Altschul et al., 1990) with
273 default settings. Conserved synteny was also used as a criterion to define the
274 genomic region containing GDF1 and GDF3 genes. Once identified, genomic
275 pieces were extracted including the 5´and 3´ flanking genes. After extraction,
276 we curated the existing annotation by comparing known exon sequences to
277 genomic pieces using the program Blast2seq with default parameters
278 (Tatusova and Madden 1999). Putatively functional genes were characterized
279 by an open intact reading frame with the canonical exon/intron structure
280 typical of vertebrate GDF1 and GDF3 genes. Sequences derived from shorter
281 records based on genomic DNA or cDNA were also included in order to attain

282 a broad and balanced taxonomic coverage. Amino acid sequences were
283 aligned using the L-INS-i strategy from MAFFT v.7 [45] (Supplementary Dataset
284 3). Phylogenetic relationships were estimated using maximum likelihood and
285 Bayesian approaches. We used the proposed model tool from IQ-Tree [46] to
286 select the best-fitting model (JTT+F+R5). Maximum likelihood analysis was
287 also performed in IQ-Tree [46] to obtain the best tree. Node support was
288 assessed with 1000 bootstrap pseudoreplicates using the ultrafast routine.
289 Bayesian searches were conducted in MrBayes v.3.1.2 [47]. Two independent
290 runs of six simultaneous chains for $10x10^6$ generations were set, and every
291 2,500 generations were sampled using default priors. The run was considered
292 to have reached convergence once the likelihood scores formed an
293 asymptote and the average standard deviation of the split frequencies
294 remained < 0.01. We discarded all trees that were sampled before
295 convergence, and we evaluated support for the nodes and parameter
296 estimates from a majority rule consensus of the last 2,000 trees. Sea squirt
297 (*Ciona intestinalis*) and Florida lancelet (*Branchiostoma floridae*) BMP2
298 sequences were used as outgroups.
299
300 **Assessment of conserved synteny**
301 We examined genes found up- and downstream of GDF1 and GDF3 in
302 species representative of vertebrates. Synteny assessment were conducted
303 for human (*Homo sapiens*), chicken (*Gallus gallus*), western-clawed frog
304 (*Xenopus tropicalis*), coelacanth (*Latimeria chalumnae*), spotted gar
305 (*Lepisosteus oculatus*), elephant shark (*Callorhinchus milii*) and sea lamprey
306 (*Petromyzon marinus*). Initial ortholog predictions were derived from the
307 EnsemblCompara database [48] and were visualized using the program
308 Genomicus v91.01[49]. In other cases, the genome data viewer platform from
309 the National Center for Biotechnology information was used.
310
311 **Acknowledgements**
312 This work was supported by the Fondo Nacional de Desarrollo Científico y
313 Tecnológico (FONDECYT 1160627) grant to JCO.
314
315

320 **Figure legends**

321

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

327

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

333

334 **Figure 3.** Schematic representations of alternative hypotheses of the sister
335 group relationships among duplicated GDF genes in tetrapods. A) According
336 to the one-duplication model the predicted phylogeny recovers a monophyletic
337 group containing GDF1 sequences from mammals, sauropsids and
338 amphibians sister to a clade containing GDF3 sequences from the same
339 groups. B) The phylogenetic prediction from the two-duplication model
340 retrieves a clade containing GDF1 and GDF3 sequences from mammals
341 sister to a clade containing GDF1/3 sequences from sauropsids; additionally a
342 clade containing GDF1 and GDF3 sequences from amphibians is recovered
343 sister to the mammalian/sauropsid clade.

344

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

349

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

354 coelacanths and sauropsids. In the ancestor of tetrapods, the single gene
355 copy was translocated to a different chromosomal location. After that, in the
356 amphibian ancestor, the single gene copy underwent a duplication event
357 giving rise to the amphibian GDF1 (Vg1) and GDF3 (derrière) genes. In the
358 ancestor of mammals, the single gene copy also underwent a duplication
359 event, giving rise to the mammalian GDF1 and GDF3 genes. After the
360 duplication, but before the radiation of the group, a second translocation event
361 occurred in the ancestor of the group. Thus, all mammals inherited a
362 repertoire of two genes (GDF1 and GDF3) that were located on two
363 chromosomes.
364

365 **References**
366
367

368 1. Lee, S.-J. Identification of a Novel Member (GDF-1) of the Transforming
369 Growth Factor-beta Superfamily. *Mol. Endocrinol.* **4,** 1034–1040 (1990).
370 2. Jones, C. M., Simon-Chazottes, D., Guenet, J.-L., Hogan, B. L. M. &
371 Pasteur, L. Isolation of Vgr-2, a Novel Member of the Transforming
372 Growth Factor-P- Related Gene Family. *Mol. Endocrinol.* **6,** 1961–1968
373 (1992).
374 3. Lee, S.-J. Expression of growth/differentiation factor 1 in the nervous
375 system: Conservation of a bicistronic structure. *Biochemistry* **88,** 4250–
376 4254 (1991).
377 4. Dale ', L., Matthews, G. & Colman2, A. Secretion and mesoderm-
378 inducing activity of the TGF-j-related domain of Xenopus Vgl. *EMBO J.*
379 **12,** 4471–4480 (1993).
380 5. Chen, C. The Vg1-related protein Gdf3 acts in a Nodal signaling
381 pathway in the pre-gastrulation mouse embryo. *Development* **133,** 319–
382 329 (2006).
383 6. Andersson, O., Reissmann, E., Jörnvall, H. & Ibáñez, C. F. Synergistic
384 interaction between Gdf1 and Nodal during anterior axis development.
385 *Dev. Biol.* **293,** 370–381 (2006).
386 7. Bengtsson, H. *et al.* Generation and characterization of a Gdf1
387 conditional null allele. *Genesis* **46,** 368–372 (2008).
388 8. Komatsu, Y. & Mishina, Y. Establishment of left – right asymmetry in
389 vertebrate development : the node in mouse embryos. *Cell. Mol. Life*
390 *Sci.* **70,** 4659–4666 (2013).
391 9. Arias, C. F., Herrero, M. A., Stern, C. D. & Bertocchini, F. A molecular
392 mechanism of symmetry breaking in the early chick embryo. *Scietific*
393 *Reports* **7,** 1–6 (2017).
394 10. Bisgrove, B. W., Su, Y.-C. & Yost, H. J. Maternal Gdf3 is an obligatory
395 cofactor in Nodal signaling for embryonic axis formation in zebrafish.
396 *Elife* **6,** e28534 (2017).
397 11. Montague, T. G. & Schier, A. F. Vg1-Nodal heterodimers are the
398 endogenous inducers of mesendoderm. *Elife* **6,** e28183 (2017).

399 12. PellicciaJ.L., Jindal, G. A. & Burdine, R. D. Gdf3 is required for robust
400 Nodal signaling during germ layer formation and left-right patterning.
401 *Elife* **6,** e28635 (2017).
402 13. Thomsen, G. H. & Melton, D. A. Processed Vg1 protein is an axial
403 mesoderm inducer in xenopus. *Cell* **74,** 433–441 (1993).
404 14. Hyatt, B. A., Lohr, J. L. & Yost, H. J. Initiation of vertebrate left–right
405 axis formation by maternal Vg1. *Nature* **384,** 62–65 (1996).
406 15. Hyatt, B. A. & Yost, H. J. The left-right coordinator: The role of Vg1 in
407 organizing left-right axis formation. *Cell* **93,** 37–46 (1998).
408 16. Söderström, S. & Ebendal, T. Localized expression of BMP and GDF
409 mRNA in the rodent brain. *J. Neurosci. Res.* **56,** 482–492 (1999).
410 17. Rankin, C. T., Bunton, T., Lawler, A. M. & Lee, S. J. Regulation of left-
411 right patterning in mice by growth/differentiation factor-1. *Nat. Genet.*
412 **24,** 262–265 (2000).
413 18. Wright, C. V. E. Mechanisms of Left-Right Asymmetry : What ' s Right
414 and What ' s Left ? A recent meeting at the Juan March Foundation in.
415 *Cell* **1,** 179–186 (2001).
416 19. Hamada, H., Meno, C., Watanabe, D. & Saijoh, Y. Establishment of
417 vertebrate left-right asymmetry. *Nat. Rev. Genet.* **3,** 103–113 (2002).
418 20. Somi, S., Houweling, A. C., Buffing, A. A. M., Moorman, A. F. M. & Van
419 Den Hoff, M. J. B. Expression of cVg1 mRNA during chicken embryonic
420 development. *Anat. Rec.* **273A,** 603–608 (2003).
421 21. Wall, N. A., Craig, E. J., Labosky, P. A. & Kessler, D. S. Mesendoderm
422 induction and reversal of left-right pattern by mouse Gdf1, a Vg1-related
423 gene. *Dev. Biol.* **227,** 495–509 (2000).
424 22. Karkera, J. D. *et al.* Loss-of-Function Mutations in Growth Differentiation
425 Factor-1 (GDF1) Are Associated with Congenital Heart Defects in
426 Humans. *Am. J. Hum. Genet.* **81,** 987–994 (2007).
427 23. Kaasinen, E. *et al.* Recessively inherited right atrial isomerism caused
428 by mutations in growth/differentiation factor 1 (GDF1). *Hum. Mol. Genet.*
429 **19,** 2747–2753 (2010).
430 24. Bao, M. W. *et al.* Cardioprotective role of growth/differentiation factor 1
431 in post-infarction left ventricular remodelling and dysfunction. *J. Pathol.*
432 **236,** 360–372 (2015).

433 25. Zhang, J. *et al.* Association of GDF1 rs4808863 with fetal congenital
434 heart defects: A case-control study. *BMJ Open* **5,** e009352 (2015).
435 26. Yang, W. *et al.* Epigenetic silencing of GDF1 disrupts SMAD signaling
436 to reinforce gastric cancer development. *Oncogene* **35,** 2133–2144
437 (2016).
438 27. Witthuhn, B. A. & Bernlohr, D. A. Upregulation of bone morphogenetic
439 protein GDF-3/Vgr-2 expression in adipose tissue of FABP4/aP2 null
440 mice. *Cytokine* **14,** 129–135 (2001).
441 28. Wang, W., Yang, Y., Meng, Y. & Shi, Y. GDF-3 is an adipogenic
442 cytokine under high fat dietary condition. *Biochem. Biophys. Res.*
443 *Commun.* **321,** 1024–1031 (2004).
444 29. Andersson, O. *et al.* Growth/differentiation factor 3 signals through
445 ALK7 and regulates accumulation of adipose tissue and diet-induced
446 obesity. *Proc. Natl. Acad. Sci. USA* **105,** 7252–7256 (2008).
447 30. Andersson, O., Bertolino, P. & Ibáñez, C. F. Distinct and cooperative
448 roles of mammalian Vg1 homologs GDF1 and GDF3 during early
449 embryonic development. *Dev. Biol.* **311,** 500–511 (2007).
450 31. Levine, A. J., Brivanlou, A. H., Suzuki, A. & Hemmati-Brivanlou, A.
451 GDF3, a BMP inhibitor, regulates cell fate in stem cells and early
452 embryos. *Development* **133,** 209–16 (2006).
453 32. Mercola, M. & Levin, M. Left-right assymetry determination in
454 vertebrates. *Annu. Rev. Cell Dev. Biol.* **17,** 779–805 (2001).
455 33. Levine, A. J. GDF3, a BMP inhibitor, regulates cell fate in stem cells and
456 early embryos. *Development* **133,** 209–216 (2005).
457 34. Nakamura, T. & Hamada, H. Left-right patterning: conserved and
458 divergent mechanisms. *Development* **139,** 3262 (2012).
459 35. Seleiro, E. A. P., Connolly, D. J. & Cooke, J. Early developmental
460 expression and experimental axis determination by the chicken Vg1
461 gene. *Curr. Biol.* **6,** 1476–1486 (1996).
462 36. Suzuki, A. *et al.* Genome organization of the vg1 and nodal3 gene
463 clusters in the allotetraploid frog Xenopus laevis. *Dev. Biol.* (2017).
464 doi:10.1016/j.ydbio.2016.04.014
465 37. Goodman, M., Czelusniak, J., Koop, B. F., Tagle, D. A. & Slightom, J. L.
466 Globins: a case study in molecular phylogeny. *Cold Spring Harb. Symp.*

467 *Quant. Biol.* **52,** 875–90 (1987).
468 38. Kriener, K., O ’huigin, C. & Klein, J. Independent Origin of Functional
469 MHC Class II Genes in Humans and New World Monkeys. *Hum.*
470 *Immunol.* **62,** 1–14 (2001).
471 39. Wallis, O. C. & Wallis, M. Characterisation of the GH gene cluster in a
472 new-world monkey, the marmoset (Callithrix jacchus). *J. Mol.*
473 *Endocrinol.* **29,** 89–97 (2002).
474 40. Li, Y. *et al.* Independent origin of the growth hormone gene family in
475 New World monkeys and Old World monkeys/hominoids. *J. Mol.*
476 *Endocrinol.* **35,** 399–409 (2005).
477 41. Opazo, J. C., Hoffmann, F. G. & Storz, J. F. Genomic evidence for
478 independent origins of NL -like globin genes in monotremes and therian
479 mammals. *Proc. Natl. Acad. Sci. USA* **105,** 1590–1595 (2008).
480 42. Hoffmann, F. G., Storz, J. F., Gorr, T. A. & Opazo, J. C. Lineage-
481 specific patterns of functional diversification in the α-and β-globin gene
482 families of tetrapod vertebrates. *Mol. Biol. Evol.* **27,** 1126–1138 (2010).
483 43. Hoffmann, F. G., Vandewege, M. W., Storz, J. F. & Opazo, J. C. Gene
484 Turnover and Diversification of the α- and β-Globin Gene Families in
485 Sauropsid Vertebrates. *Genome Biol. Evol.* **10,** 344–358 (2018).
486 44. Ramsdell, A. F. & Yost, H. J. Molecular mechanisms of vertebrate left–
487 right development. *Trends Genet.* **14,** 459–465 (1998).
488 45. Katoh, K. & Standley, D. M. MAFFT Multiple Sequence Alignment
489 Software Version 7: Improvements in Performance and Usability. *Mol.*
490 *Biol. Evol.* **30,** 772–780 (2013).
491 46. Trifinopoulos, J., Nguyen, L.-T., von Haeseler, A. & Minh, B. Q. W-IQ-
492 TREE: a fast online phylogenetic tool for maximum likelihood analysis.
493 *Nucleic Acids Res.* **44,** W232–W235 (2016).
494 47. Ronquist, F. & Huelsenbeck, J. P. MrBayes 3: Bayesian phylogenetic
495 inference under mixed models. *Bioinformatics* **19,** 1572–4 (2003).
496 48. Herrero, J. *et al.* Ensembl comparative genomics resources. *Database*
497 **2016,** baw053 (2016).
498 49. Louis, A., Nguyen, N. T. T., Muffato, M. & Roest Crollius, H. Genomicus
499 update 2015: KaryoView and MatrixView provide a genome-wide

perspective to multispecies comparative genomics. *Nucleic Acids Res.* **43,** D682–D689 (2015).

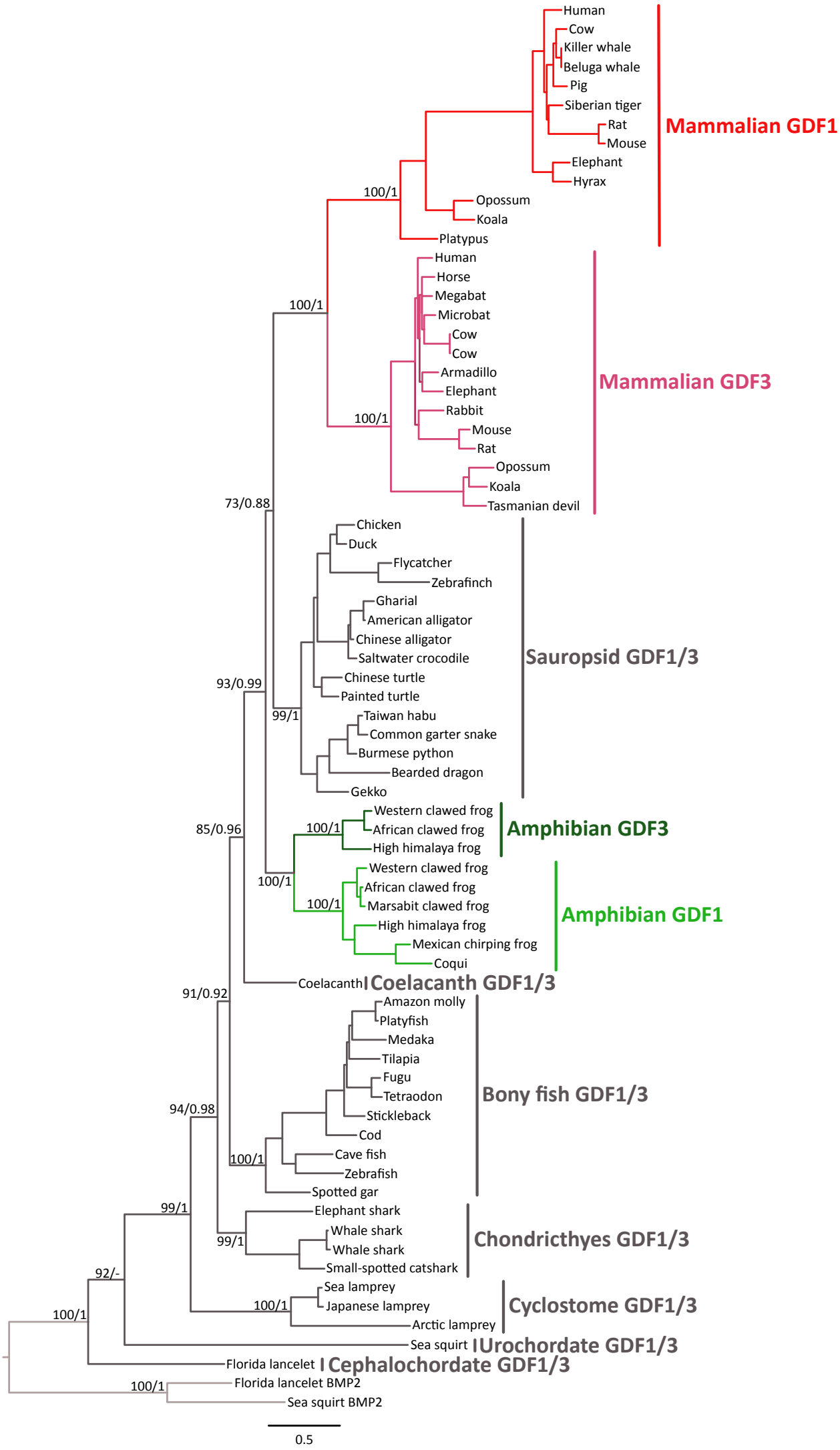
Contributions

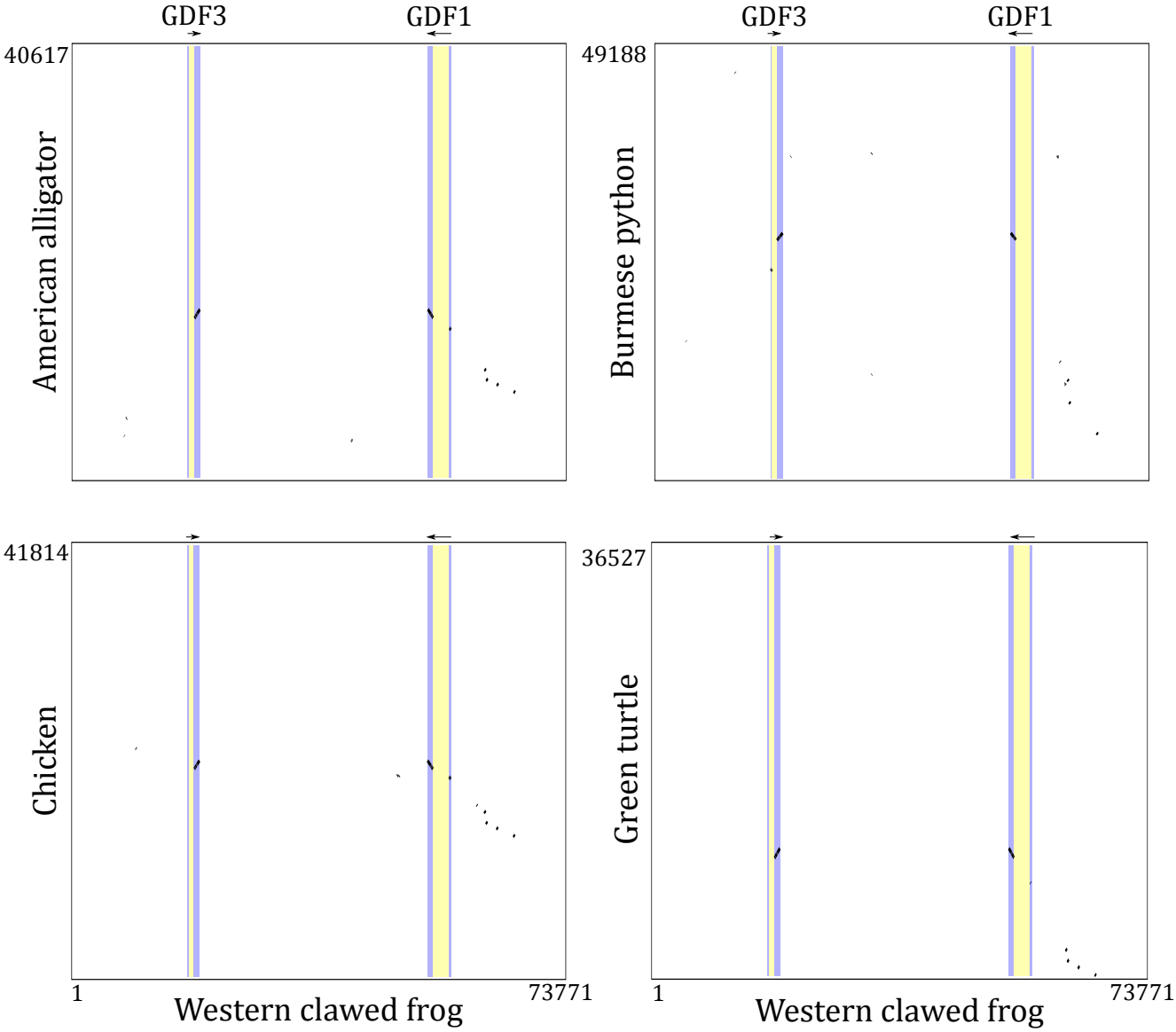
Designed the research: JCO; carried out the research: JCO and KZ; contributed materials/reagents/analysis tools: JCO; wrote the paper: JCO

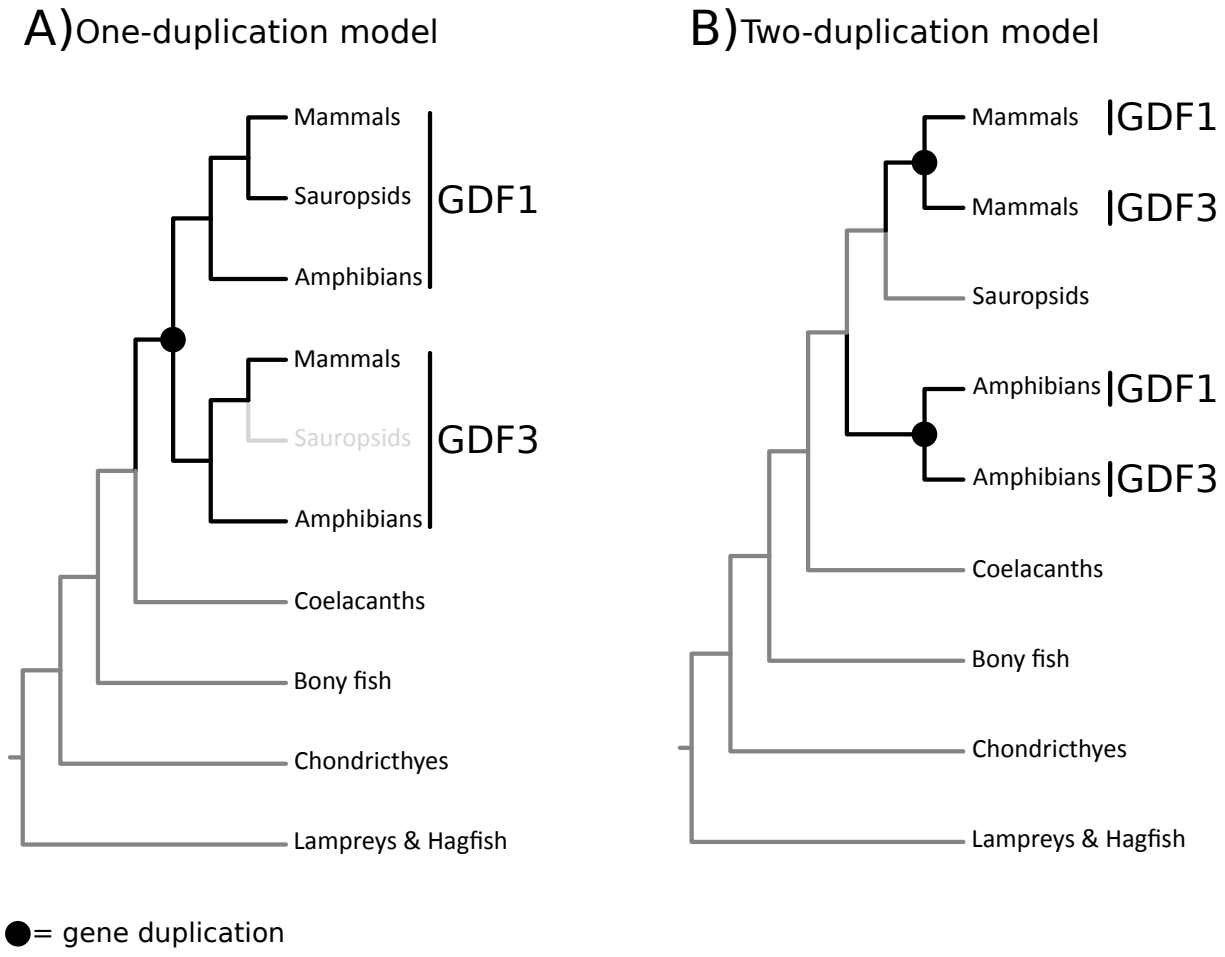
Additional information

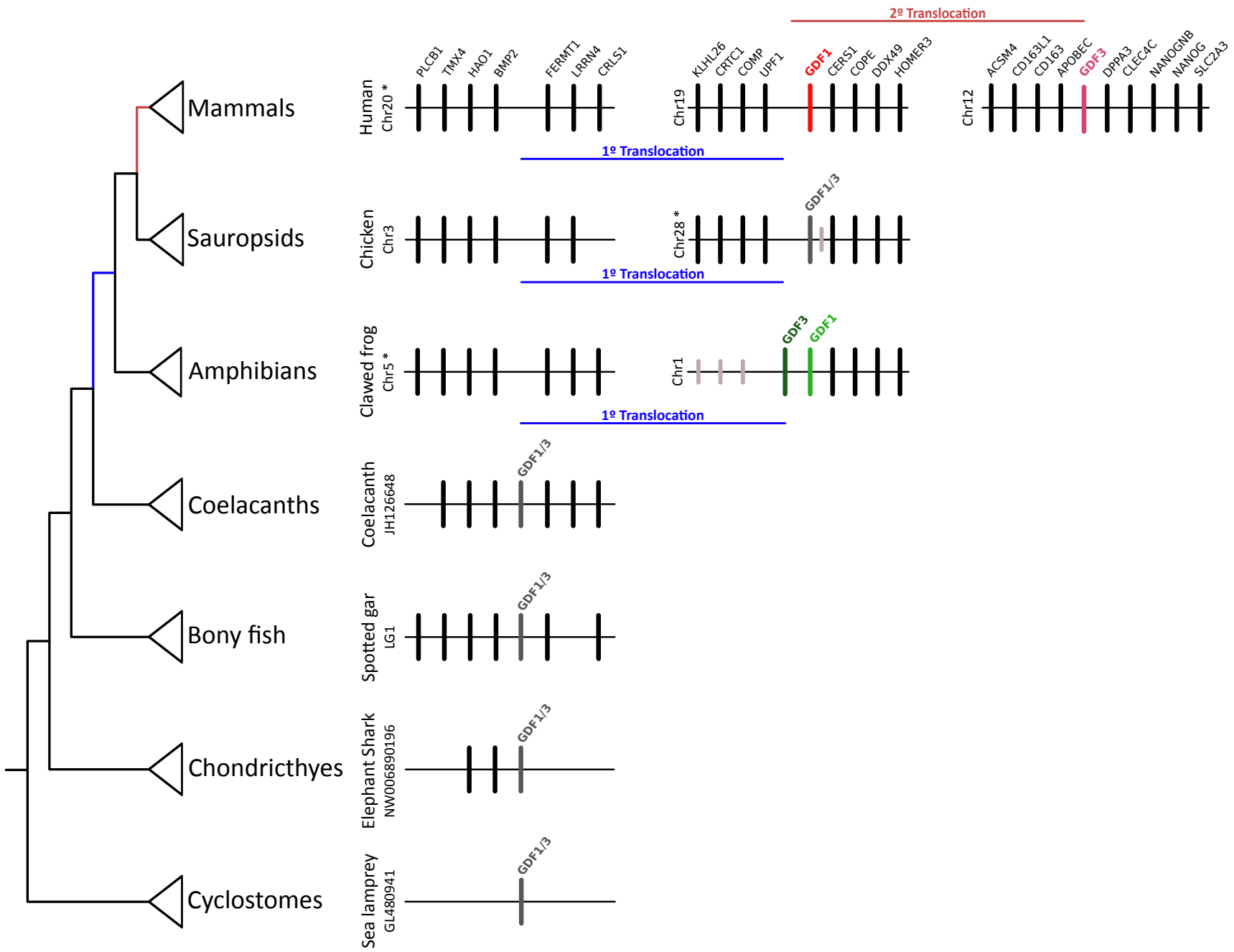
Competing interests

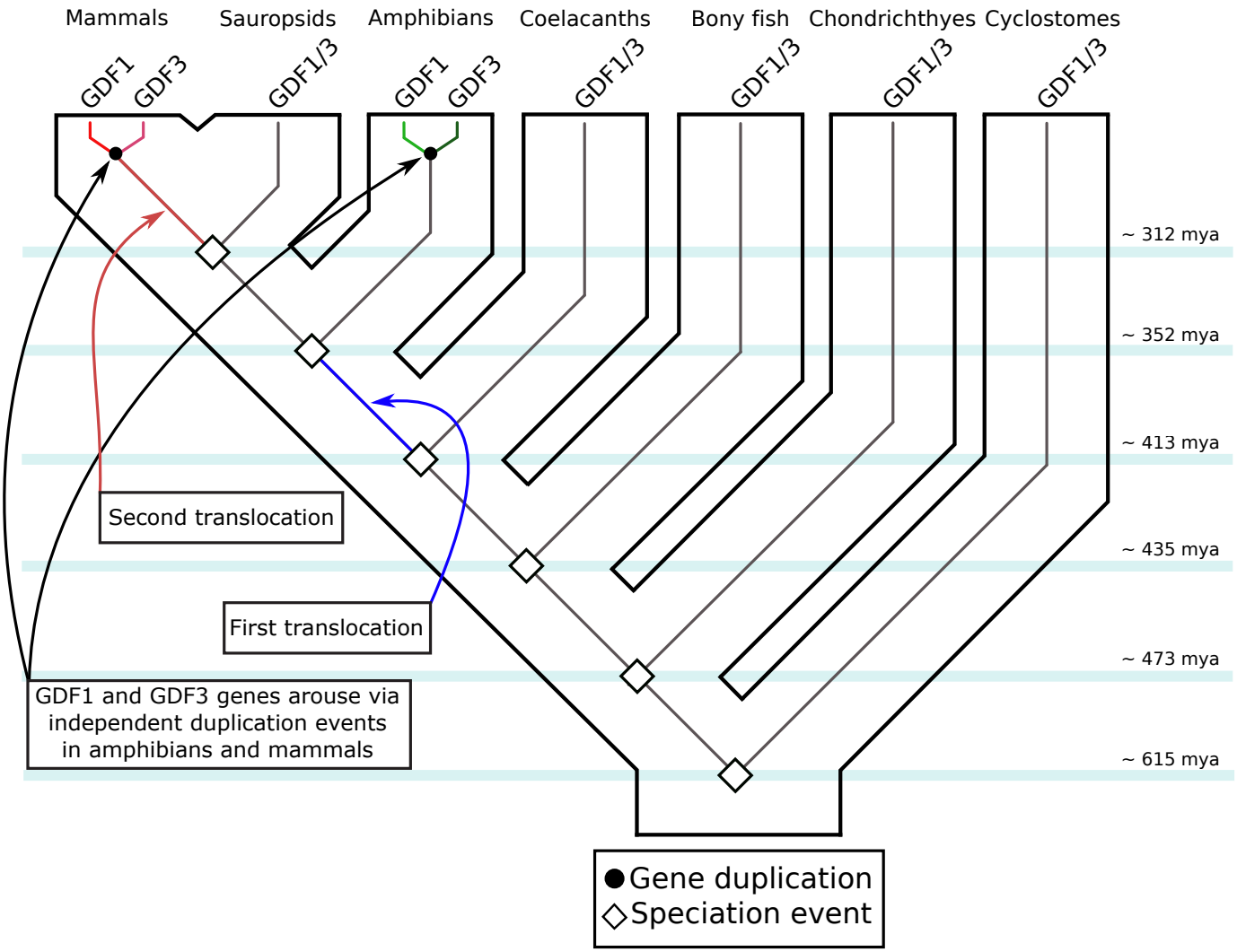
The authors declare no competing interests.


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