



Unraveling the Molecular Evolution of Blood Coagulation Genes in Fishes and Cetaceans

João Pedro V. Mariz and Mariana F. Nery*

Evolutionary Genomics Laboratory, Department of Genetics, Evolution, Microbiology and Immunology, Biology Institute, State University of Campinas, Campinas, Brazil

OPEN ACCESS

Edited by:

Mark S. Springer,
University of California, Riverside,
United States

Reviewed by:

Juan Andrés López,
University of Alaska Fairbanks,
United States
Mariano González-Castro,
Consejo Nacional de Investigaciones
Científicas y Técnicas
(CONICET), Argentina

*Correspondence:

Mariana F. Nery
marinery@unicamp.br

Specialty section:

This article was submitted to
Marine Evolutionary Biology,
Biogeography and Species Diversity,
a section of the journal
Frontiers in Marine Science

Received: 06 August 2020

Accepted: 05 November 2020

Published: 27 November 2020

Citation:

Mariz JPV and Nery MF (2020)
Unraveling the Molecular Evolution of
Blood Coagulation Genes in Fishes
and Cetaceans.
Front. Mar. Sci. 7:592383.
doi: 10.3389/fmars.2020.592383

Among the many physiological changes that occurred during the transition from an aquatic to a terrestrial lifestyle by early vertebrates, blood coagulation has an interesting history. Blood coagulation genes, originally composed of a single pathway in fishes, have expanded to include a second pathway in the evolution of terrestrial vertebrates. However, genes of this second pathway have been secondarily lost in many lineages, as is the case for cetaceans, which returned to the aquatic environment during their evolution. Herein, we investigated the selective pressures on blood coagulation genes in a phylogenetic framework, focusing on fishes and cetaceans. Taking advantage of the availability of the genetic sequences of many vertebrate lineages and using a combination of bioinformatic tools, our results showed a diverse history of gene losses and gains, with different selective pressures acting on different genes of the blood coagulation functional pathway. In addition, there was no evidence of a clear convergent molecular evolution between cetaceans and fishes, highlighting that there are many possible evolutionary mechanisms with regard to a functional pathway that involves many genes.

Keywords: coagulation, genetics, cetaceans, fishes, evolution

INTRODUCTION

Understanding the genetic basis behind ecological transitions, such as the conquest of land by early vertebrates and the return to water by later vertebrate lineages has become a central topic in evolutionary biology. During the evolution of life, these transitions occurred several times and across several taxa, and have required extensive modifications of multiple systems and organs (Uhen, 2007). Accordingly, the emergence and subsequent losses or reversions that occurred in many genes or gene families is related to these habitat transitions (Mirceta et al., 2013; Nikaido et al., 2013). One interesting characteristic that has appeared during early vertebrate evolution, which has changed along with vertebrate evolution, is the blood coagulation system.

Blood coagulation in vertebrates is usually composed of cascades of glycoprotein activation, which promotes the activation of prothrombin to thrombin, and finally fibrinogen to fibrin. The cascades can be arbitrarily divided based on their initial activation process, comprising two pathways: the tissue factor pathway (also called the extrinsic pathway), and the contact pathway (also called the intrinsic pathway). The components of these pathways are known as “factors.” Most of these factors belong to the serine protease family and bear some level of homology (Furie and Furie, 1988).

The tissue factor (TF) pathway is composed of proteases (factor VII, factor IX, factor X, and prothrombin) codified by the genes F7, F9, F10, and F2, respectively, which act with several cofactors (tissue factor, factor V, factor VIII) codified by the genes F3, F5, and F8, respectively. There is evidence that this pathway evolved more than 430 Mya, before the divergence of teleosts in the vertebrate clade (Davidson et al., 2003). On the other hand, the contact pathway (CP) is evolutionarily more recent. Although fish contain a factor of this pathway (a high-molecular-weight kininogen [HMWK] without a CP functional domain), it originated only after the evolution of tetrapods, around 380 Mya (Clack, 2002; Doolittle, 2009). The pathway is composed of the proteases factor XI, factor XII, and plasma prekallikrein (PK), codified by the genes F11, F12, and KLKB1, respectively, in addition to HMWK, as cited above. The factors of both pathways and their relationships are shown in **Figure 1**.

Genes for both the TF and CP seem to have arisen from genetic duplications during vertebrate evolution (Jiang and Doolittle, 2003). Some authors have proposed a round of large-scale duplication events in early chordate history (Holland et al., 1994; Sidow, 1996), and after that, although before amniote divergence, a new duplication event originated the CP. Correspondingly, it has been proposed that F11, PK, and F12 are recent paralogs (Pebusque et al., 1998; Ponczek et al., 2008). Note that genes for blood coagulation factors are denoted by Arabic numerals corresponding to their Roman numeral protein name.

Some groups of tetrapods do not seem to express some or any components present in the CP. Although not extensively researched, birds seem to lack factor XII, and in some cases, factor XI (Tahira et al., 1977; Frost et al., 1999; Nevill, 2009). Nevertheless, most reptiles, except for snakes and sea turtles (Soslau et al., 2004; Nevill, 2009), have a fully functional CP. Among mammals, there is evidence that the sei whale plasma lacks factors XI, XII, and the Fletcher factor (a plasma prekallikrein, codified by KLKB1) (Saito et al., 1976). Later, it was reported that minke whale F12 is a pseudogene (Semba et al., 1998). More recently, a genomic study demonstrated that KLKB1 and F12 were lost in the cetacean stem lineage, and these losses are adaptive as they reduce the risk of thrombus formation during diving (Huelsmann et al., 2019). The loss of both KLKB1 and F12 genes, two key factors that promote thrombosis, do not affect wound sealing in cetaceans. In addition, several blood clotting genes, such as F2, F5, F8, F10, PLAT, and FGB, contain cetacean-specific amino acid changes, and some of them are predicted to cause functional changes (Yim et al., 2014).

Considering that these reports suggest complex molecular dynamics of the blood coagulation genes, we aimed to unravel the molecular evolution of these genes in a phylogenetic framework. By taking advantage of the growing availability of genomic sequences, our study intended to draw a more complete scenario

for the evolution of these complex genetic pathways. Specifically, we investigated the evolutionary dynamics of the gain and loss of blood factor genes in vertebrates and examined whether there was molecular convergence between fish and cetacean CP genes due to similar selective pressures caused by the aquatic habitat.

METHODS

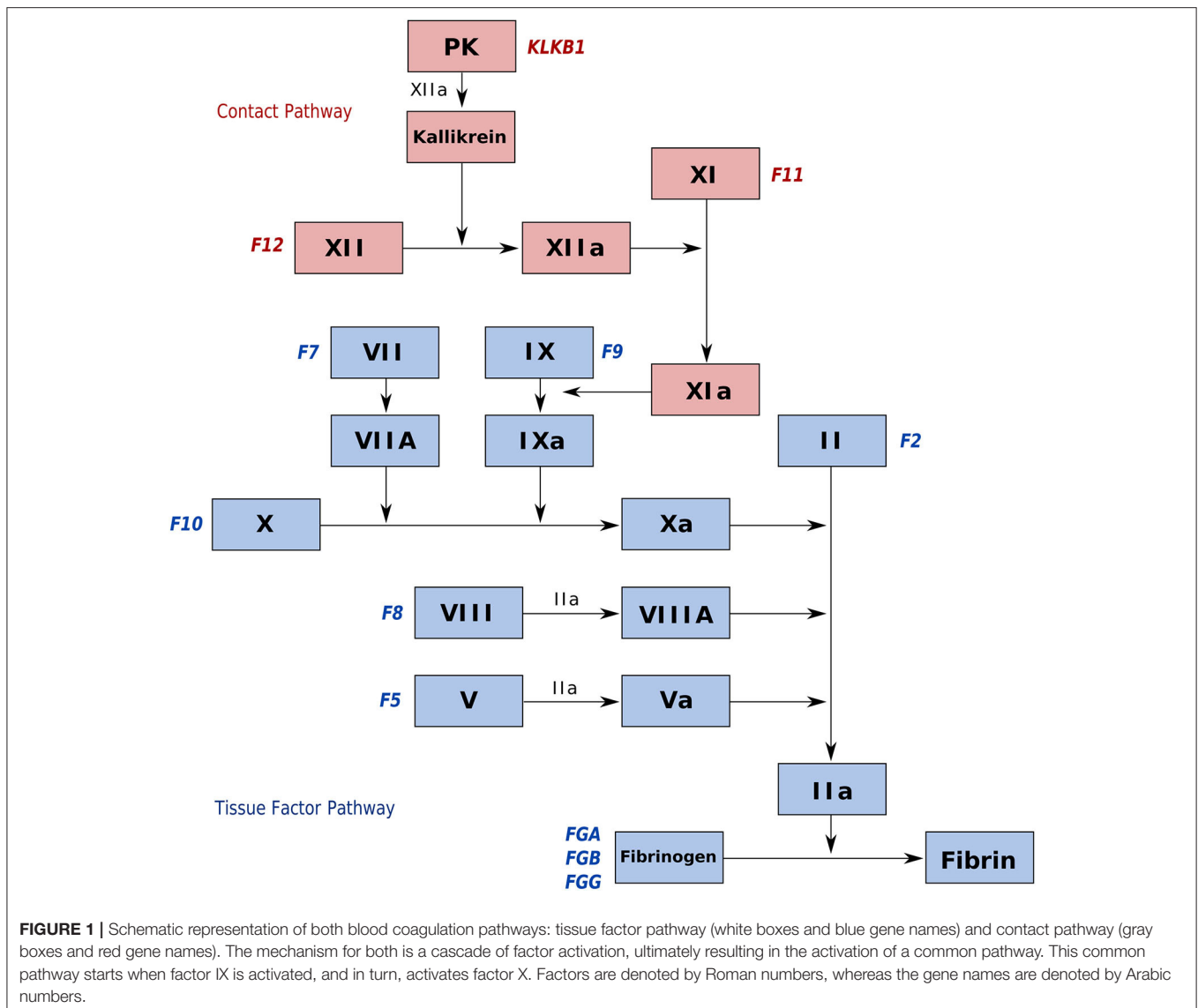
To include as many taxonomic groups as possible, we conducted an exhaustive search using Genbank database (<https://www.ncbi.nlm.nih.gov/genbank/>) for the sequences of 13 different genes related to blood coagulation: PLAT, FGA, FGB, FGG, F2, F5, F7, F8, F9, and F10 from the TF pathway, and F11, F12, and KLKB1 from the CP. To address possible annotation errors or bad quality sequences, we performed the following manual filters after alignment: sequences without start or stop codons were immediately discarded, sequences that did not align satisfactorily with closely related species were also removed, and we included at least two representatives from the main vertebrate orders. After filtering, the coding sequences were aligned using MUSCLE v. 3.8 (Edgar, 2004). We used MEGA7 (Kumar et al., 2016) to convert sequences to protein before aligning and then converted them back to DNA using PAL2NAL (Suyama et al., 2006) to obtain better alignment results.

We used both maximum likelihood and Bayesian methods to estimate phylogenetic relationships for each of our genes. Maximum likelihood trees were built using IQ-TREE v. 1.6 Web Server (Trifinopoulos et al., 2016), using 1,000 bootstrap iterations for support, and Bayesian trees were built using MrBayes's v. 3.2 Markov Chain Monte Carlo algorithm (Ronquist et al., 2012). For the latter, 1,000,000 iterations with six simultaneous chains were used, sampling every 100 generations. We used posterior probability data from those Bayesian trees to support clades established in our maximum likelihood best trees, together with bootstrap support.

To estimate the evolutionary molecular rates of these genes, and the role of natural selection, we used several different algorithms. A maximum-likelihood codon substitution model implemented in the PAML 4.8 package (Yang, 2007) was used to infer the dN/dS (ω) rate (the ratio between non-synonymous and synonymous mutations, which represents selective pressure) for each lineage with two different models. The one-ratio model establishes, *a priori*, a single ω value for all branches, whereas the free-ratio model assumes independence of ω values among branches. We initiated both models using $\omega = 1$ and then compared the results between models through likelihood ratio tests (LRT). For those genes where the free model fitted the data better, we performed a branch-site test, which compared model A with the null model (fixed $\omega = 1$) also implemented in PAML, to identify sites under positive selection in a foreground branch. We ran this algorithm twice for each gene, using only the cetacean clade as the foreground branch and later only in "fish" clades (ray-finned fishes, lobe-finned fishes, and cartilaginous fishes).

Another algorithm used in our analysis was a mixed-effects model of evolution (MEME) implemented in the HyPhy package v.2 on the DataMonkey webserver (Murrell et al., 2012). This

Abbreviations: FXI, Factor XI; FXII, Factor XII; HMWK, High Molecular Weight Kininogen; MEME, Mixed Effects Model of Evolution; PK, Plasma Prekallikrein; TF, Tissue Factor; CP, Contact Pathway.



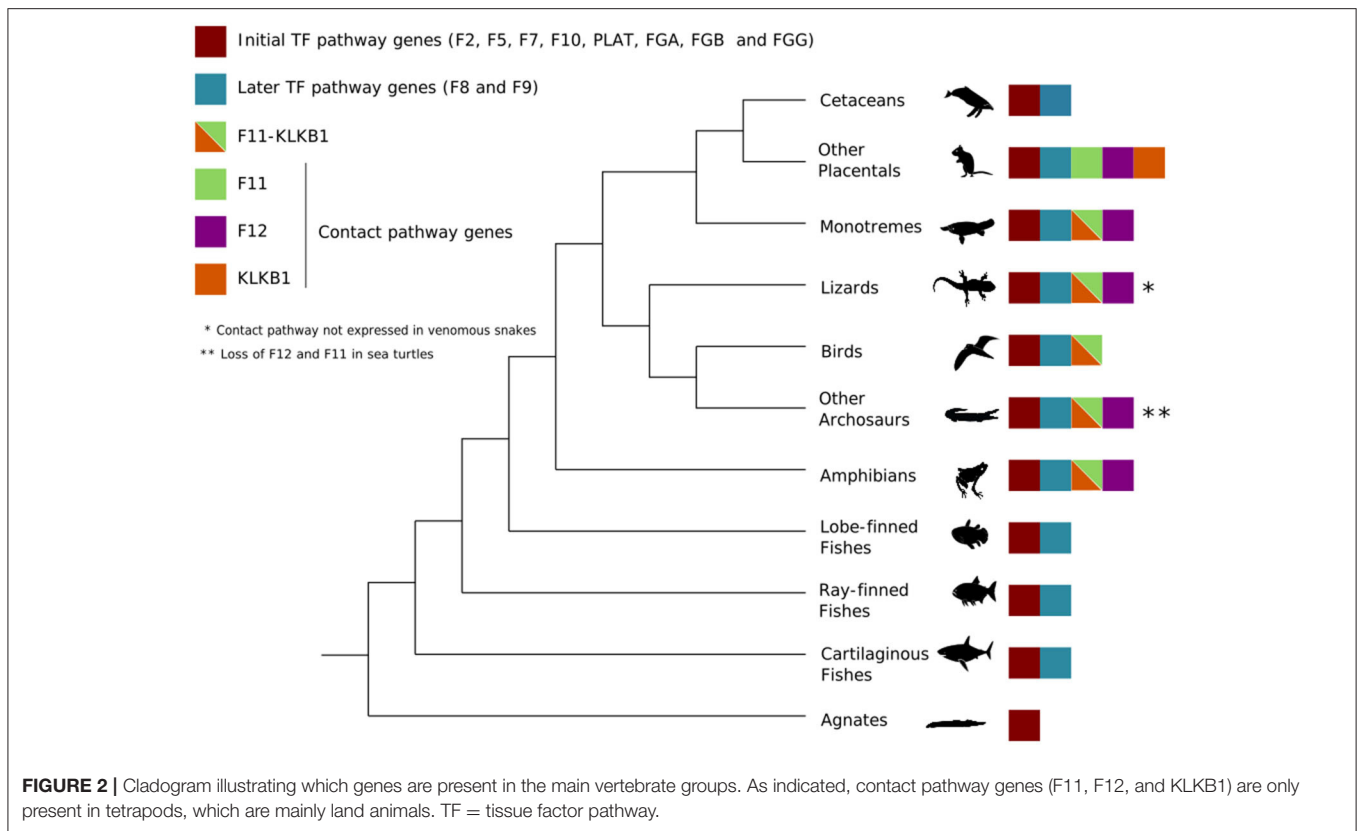
model accounts for heterogeneous selection among sites, while also allowing varying selective pressure among branches, and is able to infer both episodic and pervasive positive selection at a particular site, which is ignored in branch-models. For our data, we used the best-fitting nucleotide substitution model estimated by the automatic model selection tool available on DataMonkey.

We also used the RELAX algorithm (Wertheim et al., 2015) that detects relaxed or intensified selection on a selected branch of the tree. The algorithm builds upon a branch-site model, and is able to detect the relaxation or intensification of the selective strength of a foreground branch relative to a background branch. This is given in terms of an exponential K , which can assume values >1 (indicating intensification) or between 0 and 1 (indicating relaxation). We ran this algorithm using the same foreground branches as those used for the branch-site in PAML (cetaceans and fishes).

RESULTS

We retrieved, on average, 146 sequences for TF pathway genes and 71 for CP genes. Details on species and accession numbers can be found in the **Supplementary Material**. These data gathered from available databases allowed us to draw a phylogenetic scenario that shows the history of those genes over evolutionary time, as depicted in **Figure 2**. As expected, fish clades contained only genes from the TF pathway and in tetrapods, we observed the emergence of the F12 and F11-KLKB1 genes, which belong to the CP. However, as mentioned before, F12 and F11-KLKB1 are lost in some individual lineages, such as sea turtles and venomous snakes.

In mammals, except for monotremes, KLKB1-F11 gives rise to separate KLKB1 and F11 genes. All phylogenetic trees built with both maximum likelihood and Bayesian approaches recovered known species relationships into well-supported



clades, with high bootstrap and posterior probability support (see **Supplementary Material**). Branch lengths varied among different clades in each gene tree and among the genes. Fishes showed very long branch lengths in most trees, except for F10, FGA, and PLAT, which showed short branches in all trees. Mammals had a long root length for all genes, although branches varied from short to very long in specific groups (such as Rodentia or Eulipotyphla in the F2 gene, see **Supplementary Material**). Cetaceans generally showed short branch lengths, except for sperm whales (*Physeter catodon*) for the F7 gene. Fish clades had varying behavior, with long and short branches present, along with different trees.

We tested different modes of selective processes using a range of algorithms and compared the results between fishes and cetaceans. Regarding the variation in omega ratio (ω), we found significant differences throughout the LRT tests between the one model (i.e., assuming the same ω ratio for all lineages) and the free model (i.e., assuming an independent ω ratio for each branch of the tree) for all genes. The better fit of the free model for all genes indicates that ω rates are different among lineages for all 13 genes analyzed. In the 13 genes where the free model fits our data better, we implemented the branch-site model that distinguishes between cetaceans and non-cetaceans as well as branch-site models that distinguish between fishes and non-fishes. The results are summarized in **Table 1** and show that for cetaceans, the F7, F8, and FGA genes show evidence of sites under positive selection, whereas for fishes, only the F7 gene showed sites identified as positively selected.

Among all genes, MEME identified a wide number of codons under positive selection, ranging from 3.38 to 12.12% of the whole gene (see **Supplementary Material**). To identify genes that have undergone relaxed or intensified selection, we implemented the RELAX algorithm, and significant intensified selection was found in fishes for the F2, F5, F8, FGA, FGB, FGG, and PLAT genes. Cetaceans were identified to show intensified selection for the F7, F8, F10, and F11 genes, and relaxed selection for the F2 and F9 genes (**Table 2**).

DISCUSSION

Our objective was to investigate the molecular evolutionary history of blood coagulation genes in vertebrates, focusing mainly on fishes and cetaceans. In addition, we wanted to study the influence of the similar selective pressure exerted by the aquatic environment on the evolutionary history of these genes. Accordingly, we retrieved those genes from genomic databases and performed different analyses to account for the different types of selection that could have occurred during the evolution of these genes. Our findings suggest a dynamic and complex history, showing clades and genes with unique modifications, and not a straightforward convergent pattern of selective pressure in blood coagulation genes in fishes and cetaceans, except for the absence of contact phase genes in both clades, and a few genes with matching patterns of selection.

Regarding the TF pathway genes, the pattern of evolution seen in the phylogenetic trees suggests that, after gene duplication

TABLE 1 | PAML branch-site tests the null model ($\omega = 1$) against the model A (which allows $\omega > 1$ in the foreground branch) and of evolution rates.

	L	np	LRT	P		
Tissue Factor Pathway Genes						
F2	Fishes	Model null	-62885.9348	309	0	1
		Model A	-62885.9348	310		
	Cetaceans	Model null	-62991.0406	309	0	1
		Model A	-62991.0406	310		
F5	Fishes	Model null	-110652.3475	245	0	1
		Model A	-110652.3474	246		
	Cetaceans	Model null	-110912.7091	245	1.076	0.30
		Model A	-110912.1711	246		
F7	Fishes	Model null	-34779.1933	255	1993.335	0
		Model A	-35775.8609	256		
	Cetaceans	Model null	-34771.9592	255	34.733	0
		Model A	-34752.9593	256		
F8	Fishes	Model null	-37851.6447	221	0	1
		Model A	-37851.6447	222		
	Cetaceans	Model null	-37900.4991	221	3.902	0.04
		Model A	-37898.5480	222		
F9	Fishes	Model null	-27466.5827	321	0	1
		Model A	-27466.5827	322		
	Cetaceans	Model null	-27603.1334	321	0	1
		Model A	-27603.1334	322		
F10	Fishes	Model null	-32718.4894	201	0	1
		Model A	-32718.4894	202		
	Cetaceans	Model null	-32793.0455	201	0	1
		Model A	-32793.0455	202		
FGA	Fishes	Model null	-34798.2379	299	0	1
		Model A	-34798.2379	300		
	Cetaceans	Model null	-34860.3565	299	4.354	0.03
		Model A	-34858.1796	300		
FGB	Fishes	Model null	-53601.5199	373	0	1
		Model A	-53601.5199	374		
	Cetaceans	Model null	-53316.2302	373	0	1
		Model A	-53316.2302	374		
FGG	Fishes	Model null	-51206.5939	363	0	1
		Model A	-51206.5939	364		
	Cetaceans	Model null	-51386.9576	363	0	1
		Model A	-51386.9576	364		
PLAT	Fishes	Model null	-52804.8727	311	0	1
		Model A	-52804.8727	312		
	Cetaceans	Model null	-52992.8097	311	1.601	0.21
		Model A	-52992.0091	312		
Contact Pathway Genes						
F11	Fishes	-				
		-				
	Cetaceans	Model null	-28956.58016	137	1.162	0.28
		Model A	-28955.99957	138		

L, likelihood values; np, number of parameters; LRT, likelihood ratio test; P, probability. Significant LRT results only when $P < 0.05$, which are denoted in bold. F12 and KLKB1 are absent from both fishes and cetaceans.

TABLE 2 | Summary of the results from RELAX analyses.

		Relax (K)
Tissue Factor Pathway Genes		
F2	Fishes	1.21
	Cetaceans	0.44
F5	Fishes	1.44
	Cetaceans	Not significant
F7	Fishes	Not significant
	Cetaceans	2.92
F8	Fishes	13.68
	Cetaceans	1.81
F9	Fishes	Not significant
	Cetaceans	0
F10	Fishes	Not significant
	Cetaceans	2
FGA	Fishes	1.35
	Cetaceans	Not significant
FGB	Fishes	1.39
	Cetaceans	Not significant
FGG	Fishes	1.21
	Cetaceans	Not significant
PLAT	Fishes	1.49
	Cetaceans	Not significant
Contact Pathway Genes		
F11	Fishes	-
	Cetaceans	7.94

RELAX indicates relaxed ($K < 1$) or intensified selection ($K > 1$). F12 and KLKB1 are absent from both fishes and cetaceans.

events that preceded tetrapod divergence from fish, blood coagulation genes underwent a quick-change period on the stem branch of the mammalian group. Cetaceans seem to have few specific changes in their tissue pathways after divergence from other mammalian groups. Fishes, on the other hand, underwent fast evolutionary changes in their blood coagulation genes. Fish blood coagulation has been previously described as species-specific with many variations in clotting efficiency within the group (Doolittle and Surgenor, 1962), which is in accordance with our findings. Additionally, it has been noted that the fish molecular evolution rate greatly varies between species, especially among actinopterygian fishes (e.g., Steinke et al., 2006). This seems to be related to the extensive duplication and neofunctionalization of genes in fish clades.

The contact phase genes appeared with the emergence of amphibians during the conquest of land by early vertebrates. As the predecessor F11-KLKB1 present only in non-mammalian tetrapods does not seem to have the same functions as its successors in mammals (Ponczek et al., 2008), it can be hypothesized that after duplication, these recently created CP genes were alleviated from selective pressures and, by positive selection or neutral evolution, they mutated and were under purifying selection again. Accordingly, the contact phase genes also depict a long branch in the stem mammalian lineage in

our phylogenetic trees, followed by short branches in extant mammalian lineages. As with the tissue pathway genes, it likely indicates an acceleration in the evolutionary rate of these genes in early mammals, followed by a restriction of genetic changes in modern mammals, which can be explained by the neofunctionalization process mentioned previously (Zhang, 2003). Several reports in the literature exemplify this process in other genetic systems, such as the red and green opsin genes (Yokoyama and Yokoyama, 1989), or the three types of retinoic acid receptor genes in vertebrate development (Escriva et al., 2006).

The molecular evolution of the genes corroborates the picture drawn by our phylogenetic trees. The enormous diversity in species and habitats of fishes (ray-finned fishes especially) is in accordance with different physiological requirements, and the varying selection regime among fishes is expected. In addition, fishes comprise a group that appeared around 500 Mya, whereas cetaceans appeared more recently in evolutionary history (around 50 Mya); hence, a more homogeneous evolutionary rate among cetacean species is not surprising. In this context, the RELAX results for fishes showed that most fish TF genes are under intensified selection (F2, F5, F8, FGA, FGB, FGG, and PLAT; seven out of 10 genes), whereas, for whales, only F7, F8, F10, and F11 showed evidence of intensified selection (four out of 10). In the cetacean lineage, F2 and F9 were observed with relaxed selection, whereas no gene was identified as being under relaxed selection in fishes. In common between the two clades, only the F8 gene showed signals of intensified selection. On the other hand, the branch-site model identified the F7, F8, and FGA genes as having positively selected sites in the cetacean lineage and only the F7 gene as having positively selected sites in fish lineages. Between fishes and cetaceans, only the F7 gene has sites undoubtedly under positive selection. Taken together, most blood coagulation genes (eight out of 10) from the TF pathway in fishes are under intensified selection or show signals of positive selection, when compared to cetaceans (out of 10 TF pathway genes, four genes are under intensified selection, one has positively selected sites, and two genes are under relaxed selection). Identification of intensified selection and positive selection in most fish blood coagulation genes highlights the dynamic history of these genes to cope differently with the particular and distinct life histories of species in this group. In cetaceans, some of the TF pathway factors could have intensified or relaxed their selection regime due to overall physiological changes in blood coagulation after the loss of CP genes. It has been reported that this loss does not affect wound healing and that TF pathway genes have specific amino acid changes in cetaceans (Yim et al., 2014); therefore, this seems to be a likely scenario.

Using different approaches to assess the selection of panorama acting on these genes allowed us to obtain more complete and detailed information on their evolutionary history. As selective forces can leave many different molecular signatures, depending both on their strength and how this is distributed over evolutionary time, the use of multiple tools extends our capacity to detect multiple forms of these signals and helps to avoid erroneous conclusions by overlapping the results of

different tools. Accordingly, the combination of ML and Bayesian approaches recovering the same trees and being consistent with the current phylogenetic hypothesis shows us that our data is robust and also provides an estimate of the evolutionary rates between branches. PAML and MEME both provide a test for the hypothesis of selection with variation between branches and codons. However, PAML requires a foreground branch to be specified a priori. However, this restricts the number of possible positive selection scenarios (Kosakovsky Pond et al., 2011). Additionally, the PAML likelihood ratio test is more conservative and has relatively lower power (Anisimova et al., 2001). Therefore, MEME was used to support this analysis within a more flexible framework, in order to detect forms of selection not accounted for by PAML. Furthermore, the RELAX algorithm is able to identify a specific type of selection pattern, intensified or relaxed, which are not accounted for by the previous tools.

It is known that during evolution, gene loss can be a mechanism for adaptation and can be an important evolutionary force (Albalat and Cañestro, 2016; Sharma et al., 2018; Huelsmann et al., 2019). A recent study reported the loss of the KLKB1 and F12 genes in the stem cetacean lineage (Huelsmann et al., 2019). One could consider the deficiency in contact phase genes as a common feature of aquatic species, as fishes and cetaceans do not have them. However, the manatee, another aquatic mammal, does have these factors (Barratclough et al., 2016), and there is little information regarding blood clotting genes in pinnipeds, but they also seem to have these factors (Spurling, 1981). Thus, the loss of contact phase genes in cetaceans seems to be more related to an adaptation to diving behavior (Huelsmann et al., 2019) than to the aquatic environment. Sea turtles, which also have deep diving habits, have also been reported to lack factor XII (Soslau et al., 2004). As mentioned earlier, there is information that sei whale plasma also lacks factor XI (Saito et al., 1976). The F11 gene is present in cetacean genomes, and without nonsense mutations, as is the case for the F12 gene. In our analysis, F11 was identified as showing intensified selection in the cetacean lineage, but it remains to be determined whether this intensification is related to neutral evolution, as the other genes of this pathway are absent. In addition, functional studies examining factor XI in the plasma of more cetacean species are essential to clarify whether cetaceans still possess the gene in their genomes, albeit without expression.

The apparent lack of an overall pattern of convergence is expected, as molecular evolutionary processes are highly complex. When considering a set of genes that belong to the same functional pathway, such as blood coagulation, selective pressures act in a different way and with different intensities on each gene, as is the case for other genetic systems (e.g., Huang et al., 2011; Streicker et al., 2012). In other words, to achieve the same functional adaptation, there are a large number of possible molecular modifications. Taken together, our results suggest that cetacean blood coagulation evolutionary history, even though it is also adapted to underwater niches, took a different evolutionary path than that of fishes, although maintaining some similarities, such as the loss of CP genes (F12 and KLKB1) and the same molecular evolution pattern in two genes of the TF pathway (F7 and F8). In addition, this work sheds

light on many aspects of blood coagulation gene evolution among vertebrates, highlighting the complex and dynamic history of gains and losses in this functional pathway among vertebrate lineages and the selection pressures acting upon these genes.

DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**. Alignment data was deposited in Figshare under the doi: 10.6084/m9.figshare.13013156.

AUTHOR CONTRIBUTIONS

JM and MN designed the study, analyzed the data, and wrote the manuscript. All authors contributed to the article and approved the submitted version.

REFERENCES

- Albalat, R., and Cañestro, C. (2016). Evolution by gene loss. *Nat. Rev. Genet.* 17:379. doi: 10.1038/nrg.2016.39
- Anisimova, M., Bielawski, J. P., and Yang, Z. (2001). Accuracy and power of the likelihood ratio test in detecting adaptive molecular evolution. *Mol. Biol. Evol.* 18, 1585–1592. doi: 10.1093/oxfordjournals.molbev.a003945
- Barratclough, A., Floyd, R. F., Conner, B., Reep, R., Ball, R., and Stacy, N. (2016). Normal hemostatic profiles and coagulation factors in healthy free-living Florida manatees (*Trichechus manatus latirostris*). *J. Wildl. Dis.* 52, 907–911. doi: 10.7589/2015-12-328
- Clack, J. A. (2002). Patterns and processes in the early evolution of the tetrapod ear. *J. Neurobiol.* 53, 251–264. doi: 10.1002/neu.10129
- Davidson, C. J., Tuddenham, E. G., and McVey, J. H. (2003). 450 million years of hemostasis. *J. Thromb. Haemost.* 1, 1487–1494. doi: 10.1046/j.1538-7836.2003.00334.x
- Doolittle, R. F. (2009). Step-by-step evolution of vertebrate blood coagulation. *Cold Spring Harbor Symp. Quant. Biol.* 74, 35–40. doi: 10.1101/sqb.2009.74.001
- Doolittle, R. F., and Surgenor, D. M. (1962). Blood coagulation in fish. *Am. J. Physiol.* 203, 964–970. doi: 10.1152/ajplegacy.1962.203.5.964
- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340
- Escriba, H., Bertrand, S., Germain, P., Robinson-Rechavi, M., Umbhauer, M., Cartry, J., et al. (2006). Neofunctionalization in vertebrates: the example of retinoic acid receptors. *PLoS Genet.* 2:e102. doi: 10.1371/journal.pgen.0020102
- Frost, C. L., Naudé, R. J., Oelofsen, W., and Jacobson, B. (1999). Comparative blood coagulation studies in the ostrich. *Immunopharmacology* 45, 75–81. doi: 10.1016/S0162-3109(99)00058-2
- Furie, B., and Furie, B. C. (1988). The molecular basis of blood coagulation. *Cell* 53, 505–518. doi: 10.1016/0092-8674(88)90567-3
- Holland, P. W., Garcia-Fernández, J., Williams, N. A., and Sidow, A. (1994). Gene duplications and the origins of vertebrate development. *Development* 1994, 125–133.
- Huang, Y., Temperley, N. D., Ren, L., Smith, J., Li, N., and Burt, D. W. (2011). Molecular evolution of the vertebrate TLR1 gene family—a complex history of gene duplication, gene conversion, positive selection, and co-evolution. *BMC Evol. Biol.* 11:149. doi: 10.1186/1471-2148-11-149
- Huelsmann, M., Hecker, N., Springer, M. S., Gatesy, J., Sharma, V., and Hiller, M. (2019). Genes lost during the transition from land to water in cetaceans highlight genomic changes associated with aquatic adaptations. *Sci. Adv.* 5:eaaw6671. doi: 10.1126/sciadv.aaw6671
- Jiang, Y., and Doolittle, R. F. (2003). The evolution of vertebrate blood coagulation as viewed from a comparison of pufferfish and sea squirt genomes. *Proc. Natl. Acad. Sci. U.S.A.* 100, 7527–7532. doi: 10.1073/pnas.0932632100
- Kosakovsky Pond, S. L., Murrell, B., Fourment, M., Frost, S. D., Delpont, W., and Scheffler, K. (2011). A random effects branch-site model for detecting episodic diversifying selection. *Mol. Biol. Evol.* 28, 3033–3043. doi: 10.1093/molbev/msr125
- Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054
- Mirceta, S., Signore, A. V., Burns, J. M., Cossins, A. R., Campbell, K. L., and Berenbrink, M. (2013). Evolution of mammalian diving capacity traced by myoglobin net surface charge. *Science* 340:1234192. doi: 10.1126/science.1234192
- Murrell, B., Wertheim, J. O., Moola, S., Weighill, T., Scheffler, K., and Pond, S. L. K. (2012). Detecting individual sites subject to episodic diversifying selection. *PLoS Genet.* 8:e1002764. doi: 10.1371/journal.pgen.1002764
- Nevill, H. (2009). Diagnosis of nontraumatic blood loss in birds and reptiles. *J. Exotic Pet Med.* 18, 140–145. doi: 10.1053/j.jepm.2009.04.011
- Nikaido, M., Noguchi, H., Nishihara, H., Toyoda, A., Suzuki, Y., Kajitani, R., et al. (2013). Coelacanth genomes reveal signatures for evolutionary transition from water to land. *Genome Res.* 23, 1740–1748. doi: 10.1101/gr.158105.113
- Pebusque, M. J., Coulier, F., Birnbaum, D., and Pontarotti, P. (1998). Ancient large-scale genome duplications: phylogenetic and linkage analyses shed light on chordate genome evolution. *Mol. Biol. Evol.* 15, 1145–1159. doi: 10.1093/oxfordjournals.molbev.a026022
- Ponczek, M. B., Gailani, D., and Doolittle, R. F. (2008). Evolution of the contact phase of vertebrate blood coagulation. *J. Thromb. Haemost.* 6, 1876–1883. doi: 10.1111/j.1538-7836.2008.03143.x
- Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D. L., Darling, A., Höhna, S., et al. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61, 539–542. doi: 10.1093/sysbio/sys029
- Saito, H., Poon, M. C., Goldsmith, G. H., Ratnoff, O. D., and Árnason, Ú. (1976). Studies on the blood clotting and fibrinolytic system in the plasma from a sei (baleen) whale. *Proc. Soc. Exp. Biol. Med.* 152, 503–507. doi: 10.3181/00379727-152-39427
- Semba, U., Shibuya, Y., Okabe, H., and Yamamoto, T. (1998). Whale Hageman factor (factor XII): prevented production due to pseudogene conversion. *Thromb. Res.* 90, 31–37. doi: 10.1016/S0049-3848(97)00307-1
- Sharma, V., Hecker, N., Roscito, J. G., Foerster, L., Langer, B. E., and Hiller, M. (2018). A genomics approach reveals insights into the importance

FUNDING

All funding for this research was provided by São Paulo Research Foundation (FAPESP 2015/18269-1).

ACKNOWLEDGMENTS

We would like to thank FAEPEX-UNICAM for the JM grant, Lucas Canesin for helping with server usage and analysis setup, and Érica de Souza for guidance with *in vitro* experiments.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmars.2020.592383/full#supplementary-material>

- of gene losses for mammalian adaptations. *Nat. Commun.* 9:1215. doi: 10.1038/s41467-018-03667-1
- Sidow, A. (1996). Gen(om)e duplications in the evolution of early vertebrates. *Curr. Opin. Genet. Dev.* 6, 715–722. doi: 10.1016/S0959-437X(96)80026-8
- Soslau, G., Wallace, B., Vicente, C., Goldenberg, S. J., Tupis, T., Spotila, J., et al. (2004). Comparison of functional aspects of the coagulation cascade in human and sea turtle plasmas. *Compar. Biochem. Physiol. B Biochem. Mol. Biol.* 138, 399–406. doi: 10.1016/j.cbpc.2004.05.004
- Spurling, N. W. (1981). Comparative physiology of blood clotting. *Compar. Biochem. Physiol. A Physiol.* 68, 541–548. doi: 10.1016/0300-9629(81)90359-5
- Steinke, D., Salzburger, W., Braasch, I., and Meyer, A. (2006). Many genes in fish have species-specific asymmetric rates of molecular evolution. *BMC Genomics* 7:20. doi: 10.1186/1471-2164-7-20
- Streicker, D. G., Altizer, S. M., Velasco-Villa, A., and Rupprecht, C. E. (2012). Variable evolutionary routes to host establishment across repeated rabies virus host shifts among bats. *Proc. Natl. Acad. Sci. U.S.A.* 109, 19715–19720. doi: 10.1073/pnas.1203456109
- Suyama, M., Torrents, D., and Bork, P. (2006). PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res.* 34(suppl_2), W609–W612. doi: 10.1093/nar/gkl315
- Tahira, N., Dube, B., and Agrawal, G. P. (1977). Blood coagulation studies in some wild Indian birds: effect of different tissue thromboplastins. *J. Compar. Pathol.* 87, 451–457. doi: 10.1016/0021-9975(77)90034-2
- Trifinopoulos, J., Nguyen, L. T., von Haeseler, A., and Minh, B. Q. (2016). W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res.* 44, W232–W235. doi: 10.1093/nar/gkw256
- Uhen, M. D. (2007). Evolution of marine mammals: back to the sea after 300 million years. *Anat. Rec.* 290, 514–522. doi: 10.1002/ar.20545
- Wertheim, J. O., Murrell, B., Smith, M. D., Kosakovsky Pond, S. L., and Scheffler, K. (2015). RELAX: detecting relaxed selection in a phylogenetic framework. *Mol. Biol. Evol.* 32, 820–832. doi: 10.1093/molbev/msu400
- Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24, 1586–1591. doi: 10.1093/molbev/msm088
- Yim, H. S., Cho, Y. S., Guang, X., Kang, S. G., Jeong, J. Y., Cha, S. S., et al. (2014). Minke whale genome and aquatic adaptation in cetaceans. *Nat. Genet.* 46, 88–92. doi: 10.1038/ng.2835
- Yokoyama, S., and Yokoyama, R. (1989). Molecular evolution of human visual pigment genes. *Mol. Biol. Evol.* 6, 186–197. doi: 10.1093/oxfordjournals.molbev.a040537
- Zhang, J. (2003). Evolution by gene duplication: an update. *Trends Ecol. Evol.* 18, 292–298. doi: 10.1016/S0169-5347(03)00033-8

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Mariz and Nery. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.