U.PORTO

MESTRADO

CIENCIAS DO MAR E RECURSOS MARINHOS

Chondrichthyan genomes and the evolution of physiology traits in vertebrate ancestry

Inês Catarina Vasconcelos Ribeiro

M

2023

Inês Catarina V. Ribeiro.Chondrichthyan genomes and the evolution of physiology traits in vertebrate ancestry

U.PORTO

Chondrichthyan genomes and the evolution of physiology traits in vertebrate ancestry

Inês Catarina Vasconcelos Ribeiro

INSTITUTO DE CIÊNCIAS BIOMÉDICAS ABEL SALAZAR

M.ICBAS.2023

Inês Catarina Vasconcelos Ribeiro

Chondrichthyan genomes and the evolution of physiology traits in vertebrate ancestry

Dissertação de Candidatura ao grau de Mestre em Ciências do Mar – Recursos Marinhos submetida ao Instituto de Ciências Biomédicas de Abel Salazar da Universidade do Porto.

Orientador – Doutor Filipe Castro

Categoria – Professor auxiliar; Investigador

Afiliação – Faculdade de Ciências da Universidade do Porto; Centro Interdisciplinar de Investigação Marinha e Ambiental

Coorientador- Doutora Raquel Ruivo

Categoria– Investigadora

Afiliação– Centro Interdisciplinar de Investigação Marinha e Ambiental

Declaração de honra

Declaro que o presente trabalho é de minha autoria e não foi utilizado previamente noutrocurso ou unidade curricular, desta ou de outra instituição. As referências a outros autores (afirmações, ideias, pensamentos) respeitam escrupulosamente as regras da atribuição, e encontram-se devidamente indicadas no texto e nas referências bibliográficas, de acordo com as normas de referenciação. Tenho consciência de que a prática de plágio e auto- plágio constitui um ilícito académico.

Inês Vasconcelos Ribeiro

Agradecimentos

Este trabalho é o culminar de uma das etapas mais importantes do meu percurso académico. Ao longo deste projeto contei com o apoio de várias pessoas, a nível académico e também pessoal, e não podia deixar de reconhecer e agradecer a todos aqueles que enriqueceram esta experiência de uma forma ou de outra.

Em primeiro lugar gostaria de agradecer ao meu orientador, Dr. Filipe Castro, pela oportunidade que me deu de integrar a sua equipa e pela confiança depositada em mim para pôr em prática este projeto.

À minha coorientadora, Dra. Raquel Ruivo, agradeço por todo o apoio concedido ao longo deste processo, quer a nivel académico, quer a nivel pessoal. Agradeço também pela motivação e por sempre me incentivar a alargar os meus horizontes no laboratório e despertar o meu espírito crítico.

Aos meus colegas de laboratório agradeço a companhia, a ajuda e o apoio que neles encontrei durante esta minha passagem pelo CIIMAR. Ao João, ao Diogo, à Rita, à Elza e à Marta um agradecimento especial por terem sido pessoas excecionais, sempre com boa disposição e uma palavra amiga nos momentos mais desafiantes.

À minha irmã e sobrinhos por todo o carinho e pela força extra nos dias mais difíceis. Ao Miguel que sempre me incentivou a seguir em frente e a encarar todos os desafios, por maiores que eles fossem.

E, por fim, o agradecimento mais importante de todos: aos meus pais. Obrigada por sempre acreditarem em mim e por me permitirem seguir os meus sonhos. Foram e vão continuar a ser sempre a minha maior fonte de inspiração e motivação.

O presente estudo foi realizado com o apoio do projeto OCEAN3R - Reduce pressures, restore and regenerate the NW-Portuguese ocean and waters, com a referência NORTE-01-0145-FEDER-000064, cofinanciado pelo Programa Operacional Regional do Norte (NORTE 2020), através do Portugal 2020 e do Fundo Europeu de Desenvolvimento Regional (FEDER).

Table of Contents

Table of Figures

1. Abstract

The availability of genomic information has significantly increased in recent decades, at an accelerated pace, becoming one of the most relevant tools for evolutionary biology studies. Fundamentally, the improvement and velocity in genome sequencing techniques has been remarkable, allowing researchers to explore gene composition across different lineages and to get a better understanding of the associations between phenotypes and genotypes. The critical impacts of these genomic approaches can be seen in the study of the vertebrate evolutionary radiation. The rise of vertebrates was marked by the successful colonization of multiple and ecologically diverse habitats. Through this evolutionary process, novel phenotypic traits appeared, accompanied by genomic changes. Being a hallmark in gnathostomes evolution, the stomach comes as a novel phenotypic trait. Therefore, it is of utter importance to understand what led to its appearance and the underlying mechanisms behind the gain and loss of this organ that is unique to jawed vertebrates.

Within gnathostomes, the Chondrichthyes were the first group to emerge around 420 million years ago, comprising all cartilaginous fish including rays, sharks, and chimaeras. Due to their key phylogenetic position among vertebrates, studying this group of animals is highly relevant and allows researchers to get a better understanding of how early vertebrate evolution occurred.

Taking advantage of genomes and genomic data available, the present thesis intends to understand how processes of gene loss and gene duplication affected the appearance of the stomach in jawed vertebrates as well as comprehend the mechanisms that underline the loss of this organ in some gnathostome lineages, with a special focus on cartilaginous fish.

Keywords: stomach; Chondrichthyes; gene loss; gene duplication; pepsinogens; proton pump

2. Resumo

A quantidade de informação genómica disponível tem aumentado significativamente nas últimas décadas, sendo atualmente uma das ferramentas mais relevantes para estudos de biologia evolutiva. Fundamentalmente, o desenvolvimento verificado nas técnicas de sequenciamento de genomas tem sido notável, permitindo aos investigadores explorar a composição génica em diferentes linhagens e obter uma melhor compreensão das associações entre fenótipos e genótipos. Os impactos destas abordagens genómicas podem ser vistos no estudo da radiação evolutiva dos vertebrados. A ascensão dos vertebrados foi marcada por uma colonização bem-sucedida de vários habitats, bastante diversificados do ponto de vista ecológico. Com este processo evolutivo, surgiram novos traços fenotípicos, acompanhados por mudanças genómicas. Sendo uma das adaptações-chave na evolução dos Gnathostomes, o estômago surge como um novo traço fenotípico. Portanto, é de extrema importância entender o que levou ao seu aparecimento e os mecanismos subjacentes à perda e ganho deste órgão, que é encontrado exclusivamente nos vertebrados com mandíbula.

Dentro dos Gnathostomes, os Chondrichthyes foram o primeiro grupo a surgir à cerca de 420 milhões de anos atrás. Este grupo compreende todos os peixes cartilaginosos, incluindo raias, tubarões e quimeras. Devido à sua posição filogenética entre os vertebrados, o estudo destes animais torna-se altamente relevante e permite aos investigadores obter uma melhor compreensão de como ocorreu a evolução inicial dos vertebrados.

Aproveitando os dados genómicos disponíveis, a presente tese pretende compreender como os processos de perda e duplicação de genes afetaram o aparecimento do estômago em vertebrados com mandíbulas, bem como compreender os mecanismos que conduzem à perda deste órgão em algumas linhagens, com especial foco nas espécies de Chondrichthyes.

Palavras-chave: estômago; Condríctios; perda de genes; duplicação de genes; pepsinogénios; bomba de protões

3. Introduction

3.1. The "*dawn*" of vertebrate radiation

Vertebrates are a diverse group of organisms exhibiting a broad range in shape and size, and displaying an astonishing variety of adaptations (Fig.1): from the small New Guinea frog (*Paedophryne amanuensis*), reaching 7.9 mm, to the blue whale (*Balaenoptera musculus*), with its impressive average size of 28.5 m. These unique animals have adapted and colonized water, land, and air [6].

*Figure 1 - **Vertebrate diversity**. The vertebrate group has evolved a variety of shapes, sizes and adaptations over time making it a remarkable group of animals. Source: Canva Pro*

From a taxonomic standpoint, the Vertebrata is a subphylum included in the Chordate phylum, and is characterized by the presence of vertebrae and a spinal column that surrounds the notochord [7]. The vertebrate group comprises a total of seven classes of animals: Cyclostomi, Chondrichthyes, Osteichthyes, Amphibia, Reptilia, Aves and Mammalia (Fig.2). Cyclostomi are extant jawless vertebrates that include animals such as lamprey and hagfishes. Chondrichthyes, representing the cartilaginous fishes (sharks, rays and chimaeras), mark the first group of jawed vertebrates, and will be the specific object of the current work [8].

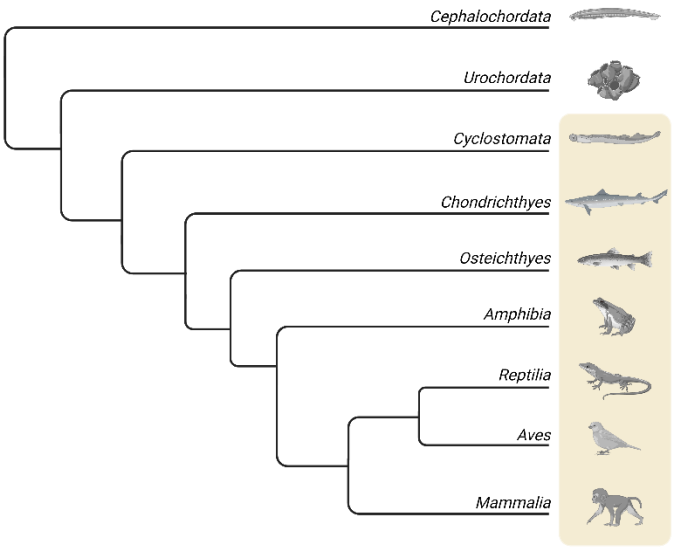

*Figure 2 - **Phylogenetic tree of the chordates**. Vertebrates are divided into two groups: Cyclostomata also known as jawless vertebrates, which include lampreys and hagfish, and the Gnathostomata which comprises all other vertebrate classes of jawed vertebrates. Adapted from [4].*

3.1.1. The origin: a chordate's view

To better understand the rise and radiation of vertebrates it is necessary to look closer to the phylum Chordata.

Chordates are deuterostomes, as is the case for both echinoderms and hemichordates, with which they share a common ancestor. The phylum Chordata is composed of three subphyla (Cephalochordata, Urochordata and Vertebrata) [9].

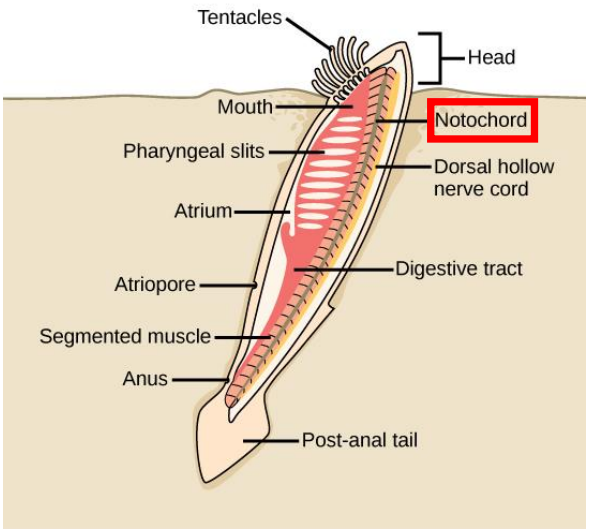

*Figure 3 - **Amphioxus**. Diagram of an amphioxus (Branchiostoma sp.) illustrating its internal anatomy. The notochord, running along its body, was crucial for its burrowing and swimming activities. Retrieved from https://bio.libretexts.org/Bookshelves/Introductory_and_General_Biology/Book%3A_General_Biology_(Boundless)/29%3A_Vertebrates/29.01%3A_Chordates/29.1B%3A_Chordates_and_the_Evolution_of_Vertebrates*

The appearance of the notochord—a key feature of chordates—is believed to be associated with changes in the swimming activity. Non-chordate marine invertebrates display a ciliary motion in larval stages. In contrast, chordate larvae display a fish-like movement, beating their tails side to side to swim. Such action requires the

contraction of the tail's muscles. Thus, it is possible that the notochord, that provides support throughout chordates bodies, evolved in order to grant the necessary stiffness that allows chordate larvae tail movements to be efficient in their swimming activities [9].

With that, it is believed that the first chordate most likely shared a similar body plan to an amphioxus (*Branchiostoma* sp., cephalochordate), with the notochord running along its body, offering support for both burrowing and swimming (Fig.3) [10].

From this point on, several changes occurred at the origin of the vertebrates' lineage. Once again, a higher level of activity and a more demanding metabolism were put forward as drivers of such evolutionary process [10]. With the emergence of vertebrates, several unique traits such as the neural crest, the placodes, a more complex brain, and mineralized tissues (cartilage, bones, and teeth) appeared [10, 11].

Interestingly, and despite the apparent morphological differences, Urochordata (i.e. ascidians and sea squirts) and not Cephalochordata (amphioxus) was shown to be the closest living relative of vertebrates, from a phylogenetic standpoint [12]; with subsequent massive morphological simplifications, and underlying gene loss events, shaping Urochordata development and body plan [13-15].

3.1.2. Key features: *what makes a vertebrate*?

The most notorious of all vertebrate innovations is the skeleton. With vertebrates came the appearance of mineralized tissue both in the form of cartilage and in the form of bone. Vertebrate likely emerged with two types of skeleton tissues: the endoskeleton and the exoskeleton (Fig.4). Endoskeleton bones are typically preceded by cartilage and are associated with muscles within deeper layers of the body. The limb skeleton and the axial skeleton are two examples of endoskeleton found in vertebrates. On the other hand, the exoskeleton is composed of dermal bones, such as teleost scales, and usually occupies a more superficial area of the body. Within the skeleton, it is important to point out vertebrae, the feature that gives this group of animals its name [5, 16].

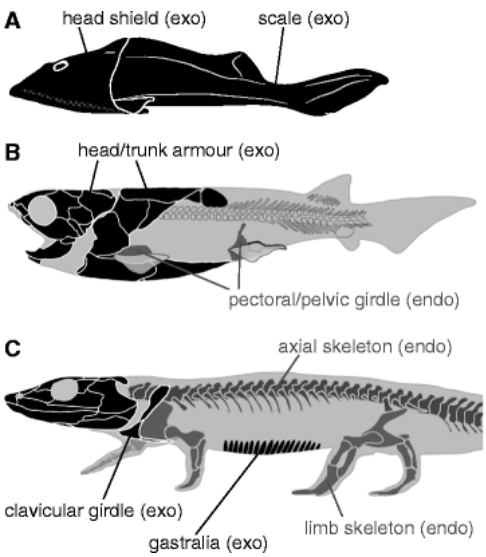

Figure 4 - ***Distribution of endoskeletons (endo) and exoskeletons (exo) in the vertebrate body.*** *Retrieved from [5].*

Yet, besides the skeleton, other vertebrate synapomorphies can be highlighted, notably related with embryonic development. The neural crest cells (Fig.5), unique to vertebrates, represent a group of pluripotent cells that is usually formed between the neural epithelium and the epidermis of an embryo. This cell population is highly migratory during embryogenesis, latter settling and differentiating into numerous cell types such as neurons, glia cells belonging to the nervous system, fibroblasts, or adipocytes [17].

Placodes are thickenings of ectoderm that occur in vertebrate's heads during embryogenesis (Fig.5). Just like neural crest cells, the placodes give origin to different cell types, including components of the paired sense organs and neurons. The placodes are therefore crucial for the cranial sensory nervous system. There are different types of placodes, that form in specific regions of the head and give rise to different cells. The different types of placodes include trigeminal, lens, olfactory, adenohypophyseal, epibranchial, hypobranchial, lateral line and otic placodes: originating a myriad of structures found in vertebrate species such as sense organs, glands, or even epidermal structures (i.e. scales, feathers, hair) and teeth [18-20].

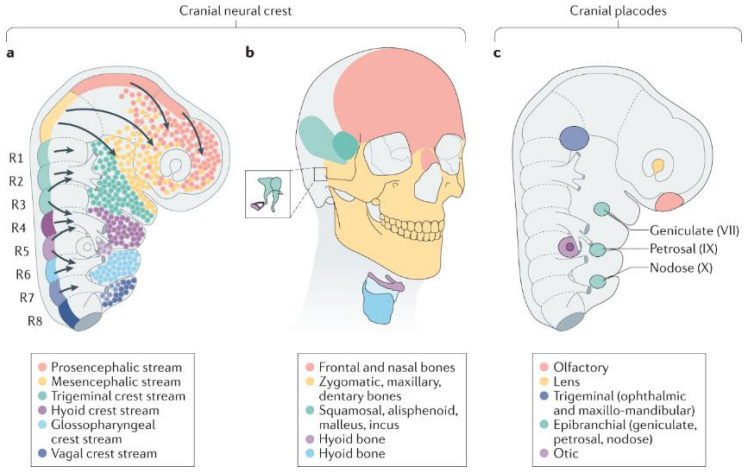

*Figure 5 - **The neural crest and cranial placodes.** The appearance of the neural crest and the cranial placodes was fundamental for the complex evolution of the vertebrate head. It Is believed that these elements were also relevant in the transition from filter-feeding chordates to highly predatory vertebrates. Retrieved from [21].*

Even though the brain itself existed prior to the appearance of vertebrates, it is within this group that significant modifications are found regarding this distinctive organ. With the emergence of chordates, the central nervous system (CNS) evolved into the control center of an animals' body. Several theories have been suggested over the years for the appearance of a polarized nervous system. Yet, despite significant advances in our understanding, it is still unclear whether the CNS had an ancestral origin, followed by episodes of loss, or if it evolved more than once in different groups (Fig.6). The single

origin theory states that the nervous system evolved in a common ancestor of Metazoa and was latter lost through independent events in both Placozoa and Porifera. On the other hand, the multiple origin theory defends that the origin of the nervous system is associated with convergent evolution episodes in Ctenophora and also in Cnidaria+Bilateria [22]. Regarding the neural anatomies, neurons and nerve nets can be found in Ctenophora, Cnidaria and Bilateria. However, brain cephalization is present only in Bilateria [22].

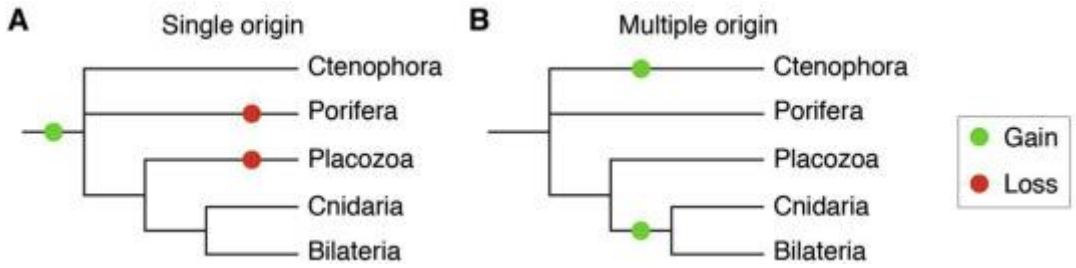

*Figure 6 - **Nervous System origin theories**. (A) Single origin theory that describes two independent loss events in Placozoa and Porifera. (B) Multiple origin theory defending a convergent evolution between Ctenophora and Cnidaria+Bilateria. Retrieved from [22].*

3.2. Early-branching vertebrates: the Chondrichthyes

Cyclostomi and gnathostomes are the two main lineages into which extant vertebrates are split. The jawless vertebrates, or Cyclostomi group, include lampreys and hagfishes. All other vertebrate animals are referred to as Gnathostomes since they exhibit jaws. The cartilaginous fish, also known as Chondrichthyans, are the earliest-branching group of jawed vertebrates. The Chondrichthyans arose 420 million years ago [23]. This group's cartilaginous skeleton serves as a distinguishing feature. The Chondrichthyans are divided into two sub-classes: the Holocephali (chimaeras) and the Elasmobranchii (sharks, skates and rays) (Fig. 7) [2]. As a group, cartilaginous fish can occupy a huge

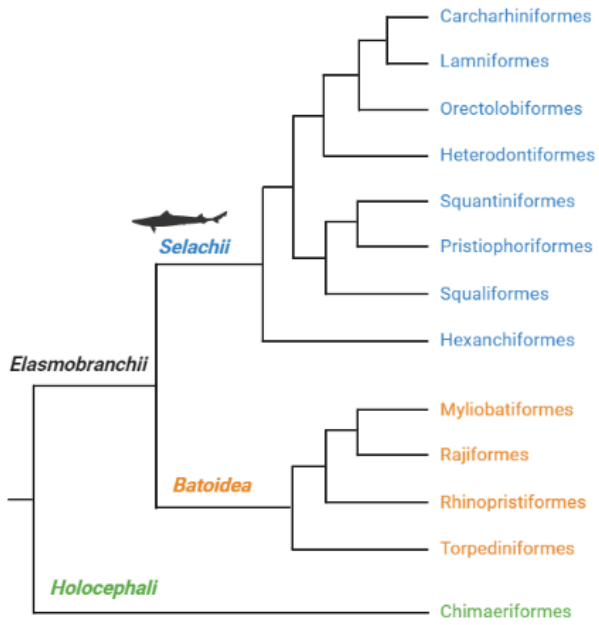

*Figure 7 - **Phylogenic tree of the Chondrichthyes**. Chondrichthyes comprise Elasmobranchii and Holocephali. Elasmobranchii include Selachii (sharks) and Batoidea (rays). Holocephali subclass is composed exclusively of chimaeras. Adapted from [2].*

variety of ecological niches, across the world: from cold to tropical waters, near the surface or within deep sea habitats, as well as in rivers, lakes, and mangroves [2].

Regarding their reproduction all cartilaginous fish have internal fertilization. However, a variety of reproductive strategies can be found in this group, making it the most diverse of vertebrates reproduction-wise [24]. Although the majority of cartilaginous fish are oviparous, strategies of placental viviparity, ovoviviparity or strict lecithotrophic oviparity can be found [24]. In addition to the sexual reproduction, asexual parthenogenesis has been described in captive Chondrichthyans, such as hammerhead sharks (family Sphyrnidae, order Carcharhiniformes) [25]. These species can produce from 1 to 10 pups per litter as is the case of the electric ray (*Torpedo torpedo*), up to 300 pups per litter as described in the whale shark (*Rhincodon typus*) [26, 27].

The lifespan of Chondrichthyan species is also very flexible, with the shortest being of around 10 years (sharpnose sharks, *Rhizoprionodon terraenovae*) and the longest reaching 272 years (Greenland sharks, *Somniosus microcephalus*) [28, 29].

3.2.1. Holocephali

Holocephalans emerged around 420 million years ago [2]. This group has only one extant order, the chimaeriformes (Fig.8), which contains around 39 species classified into three families (Callorhinchidae, Chimaeridae, and Rhinochimaeridae) and six genera (*Callorhinchus*, *Chimaera*, *Hydrolagus*, *Harriotta*, *Neoharriotta*, and *Rhinochimaera*) [30]. These animals are typically deep-sea dwellers, with a preference for depths of 500 m and deeper. However, a few species such as *Hydrolagus colliei* can be found in shallower coastal waters [31]. Chimaeras exhibit sexual and age segregation and may migrate inshore for spawning and mating on a seasonal basis, a behavioral trait that can be observed in several Chondrichthyes species. Regarding their anatomical characteristics, unlike their sister group of elasmobranchs, Holocephali have only one gill opening on each side, which is located next to their pectoral fin [32]. These animals present three pairs of tooth plates, preying small teleosts and benthic invertebrates [33]. One remarkable feature found in most chimaeras is a venomous spine, located near the first dorsal fin [34]. Importantly, chimaeras lack an epigonal organ, Leydig's organ and stomach—all found in their sister group, the elasmobranchs [2].

*Figure 8 - **Chimaera.** Chimaeras are part of the holocephalii subclass. Holocephalans split around 400 million years ago. This group has only one extant order, the chimaeriformes, which contains three families (Callorhinchidae, Chimaeridae, and Rhinochimaera)*

3.2.2. Elasmobranchii

Elasmobranchs diverged around 350 million years ago with this subclass then dividing into two subgroups: Selachii, which includes all shark species and Batoidea, a group that comprises all rays and skates [2].

3.2.2.1. Selachii

Shark species are spread throughout eight orders including Carcharhiniformes, Lamniformes, Orectolobiformes, Heterodontiformes, Squantiniformes, Pristiophoriformes, Squaliformes and Hexanchiformes [2]. Sharks can be found in a variety of habitats, with most species being exclusively marine. They can be found all the way from the surface of the ocean down to depths of around 3000 m [35]. From an anatomical perspective, these animals have between 5 to 7 pairs of gills, that are located on each side of the body (Fig. 9). Their mouths are located on their ventral side and they possess several rows of teeth that constantly grow and are replaced with new ones [36]. While most sharks are carnivorous, some species such as the whale shark (*Rhincodon typus)* have specialized in filter feeding [37].

*Figure 9 - **Shark**. Sharks are part of the elasmobranchii subclass and are part of the selachii group. There are eight known orders including Carchariniformes, Lamniformes, Orectolobiformes, Heterodontiformes, Squantiniformes, Pristiophoriformes, Squaliformes and Hexanchiformes. Source: Canva Pro*

3.2.2.2. Batoidea

The batoidea group is made up of 4 orders that include Mylioformes, Rajiformes, Rhinopristiformes and Torpediniformes [2]. Unlike sharks and chimaeras, rays have a flat body, and their gill slits are located on their ventral side (Fig.10). These animals do not have teeth but instead have flat dental plates that allow them to crush their prey [36]. These marine animals usually live near the ocean floors, with their depths ranging from near-surface areas to depths of around 3000 m [35]. They are spread throughout the world, with a soft preference for tropical and subtropical habitats, although some species prefer temperate and cold environments [2]. Some species have a venomous stinger, present in their tails [34].

*Figure 10 - **Stingray.** Rays and skates compose the batoidea group, within the elasmobranchi subclass. This group is composed of four orders including Mylioformes, Rajiformes, Rhinopristiformes and Torpediniformes. Source: Canva Pro*

3.3. Genomes and Evolution

The genome comprises an organism's complete gene sequence catalogue, as well as regulatory and noncoding regions, making it temporally and structurally dynamic. Such genomic composition arises from processes of gene loss and duplication that translate into evolutionary novelties potentially leading to changes in morphological or physiological traits.

3.3.1. Evolution by gene duplication

Gene duplication has long been acknowledged as a powerful source of evolutionary innovation. Gene duplication can occur through different mechanism, yielding distinct duplication scales (e.g., whole-genome duplication, segmental duplication, and tandem gene duplication); and can be derived from DNA-based or RNA-based sources (e.g., RNA-based transposition or retrotransposition) [38].

Tandem duplications occur at a small scale and correspond to a local duplication event in which a copy of a gene is generated. This seems to be caused by unequal crossing over resulting from homologous recombination—a combination of identical or nearly identical sequences of a chromosome. This duplication tends to create clusters of duplicated genes, that can represent gene families [38]. Tandem duplications have been extensively reported across species. For instance, in *Drosophila melanogaster*, with the appearance of 3 tandemly duplicated genes of the X chromosome that seem to have acquired novel transcription patterns and potentially

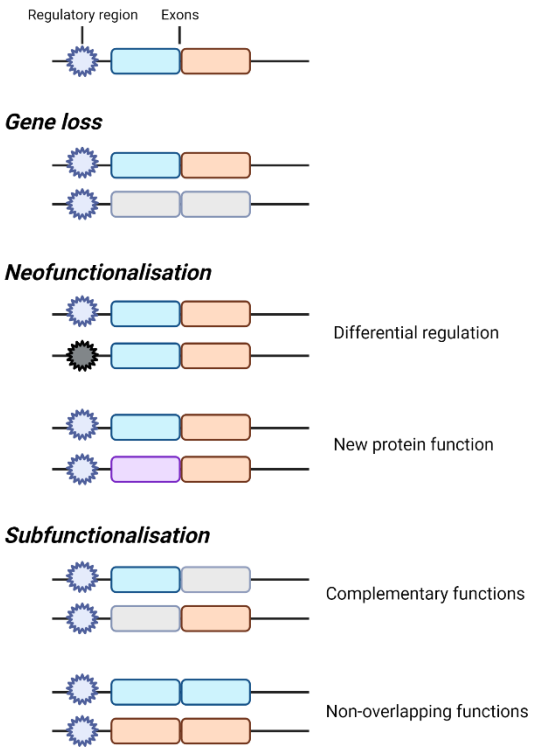

*Figure 11 - **Possible outcomes for duplicated genes.** The most common fate for duplicated genes appears to be gene loss. However, processes such as neofunctionalization and subfunctionalization may also take place. Adapted from [1].*

developed different functions [39]. In teleosts, such as the zebrafish, tandem duplications of CC chemokines clusters seem to represent novel immune adaptations [40]. The retention of tandem duplicates is also well studied in plants, as is the case of the *FMO* gene cluster in *Arabidopsis* lineage [41].

Gene duplicates can also be achieved by the action of transposable elements, in the form of RNA (retroposition) or DNA (transduplication) [38]. Retroposition occurs when reverse transcription takes places, converting mRNA of the original gene into a cDNA that is then enzymatically inserted into the organism's genome. Retrocopies are found in a wide range of species, with special focus on mammals, including the humans (e.g. *ARF6,* ADP Ribosylation Factor 6) [42, 43]. DNA transduplication mechanism has yet to be clearly described and was particularly documented in plants [38].

In addition to single gene duplication, segmental duplications may also occur, leading to the duplication of longer stretches of DNA—i.e. including multiple gene *loci* [44].

Besides local gene duplications, whole genome duplications have been shown to greatly promote evolutionary change and innovation [45]. In fact, vertebrates and/or gnathostomes experienced two whole genome duplications, and these are believed to be the source of some of the novel characteristics found within this group. These include the endoskeleton, placodes, neural crest, and their more complex brain. In addition to the 1R and 2R genome duplications, teleost fish have experienced a third duplication (3R), unique to this lineage [46]. Within teleost, salmonids have a fourth whole genome duplication, that took place in the common ancestor of this group, 80 million years after they diverged from Esociformes [47].

Although gene loss seems to be the most common fate for duplicated genes, several other outcomes have been described (Fig. 11). The duplicated genes can undergo processes of neofunctionalization or subfunctionalization. The first process leads to the appearance of new functions in one of the duplicates. This can happen through changes in the amino acid sequence or via a mutation in the regulatory region. In the case of subfunctionalization, each of the duplicates retains distinct ancestral functions leading to non-overlapping functions of the two duplicated genes [1, 45].

3.3.2. Evolution by secondary gene loss

Although it was previously dismissed as a mechanism associated with the removal of redundant gene duplicates from a genome, gene loss has recently been given a fresh

perspective with the emergence of new genomic data, which has demonstrated its prevalent role as an evolutionary driver [4].

Two different mechanisms are believed to trigger gene loss in an organism (Fig.12). The first process comes as a result of a slow event of accumulation of deleterious mutations, following an initial mutation on the original gene leading to loss-of-function. As for the second cause of gene loss, it consist on the total or partial removal of the gene, related with a rapid mutational event that could be triggered by an uneven crossing over during meiosis or the physical removal of the gene via a viral element [4].

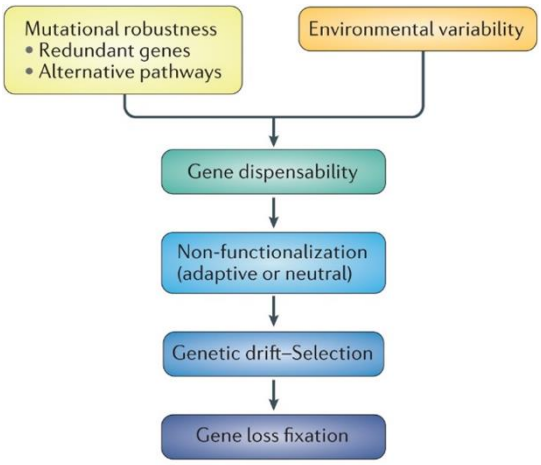

*Figure 12 - **Gene loss mechanism.** Gene loss is closely related to the level of dispensability of the gene, taking into consideration the effect of the loss on the organism's fitness. When several mutations take place, the appearance of redundant genes leads to a higher level of dispensability of the gene, accelerating the gene loss process itself. Environmental changes also play a significant role when it comes to altering gene dispensability. Retrieved from [4].*

To comprehend the true impact of gene loss on the evolution of species, one must examine the probability of non-functionalization mutations being either adaptive or neutral. With the "*less is more*" theory, it is suggested that non-functionalization represents an evolutionary adaptive response that is highlighted when populations face different selective pressures due to changes in environmental conditions [45]. On the other hand, the idea of "*regressive evolution*" states that there is a progressive loss of unnecessary characteristics over time. In these cases gene loss takes on a neutral role [4, 48].

3.4. A *specialized* section of the gut: the Stomach

The stomach is one of the remarkable innovations found exclusively in gnathostomes. It is an enlargement of the gastrointestinal tract which is marked by the presence of hydrochloric acid and pepsinogen [49].

Estimates suggest that gastric acid first appeared around 350 million years ago, with the first elasmobranchs [49]. Within the jawed vertebrates all groups of animals have stomachs. However, few exceptions are found such as in some teleosts, lungfish,

Holocephali and even egg-laying mammals that do not have acid secretion or a gastric phenotype (Fig. 13) [50]. In non-mammal vertebrates, gastric acid and pepsinogen is secreted by specialized cells known as oxynticopeptic cells. These cells are typically found in the gastric mucosa, predominantly in the anterior stomach. In the case of mammals, two different types of cells are responsible for the gastric acid and pepsinogen secretion. Parietal cells are in charge of the gastric acid secretion and chief cells secrete pepsinogen [51].

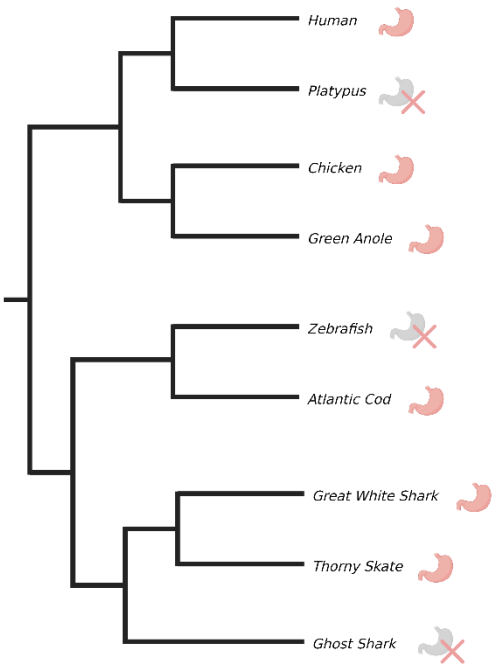

Figure 13 - ***Phylogenetic relations of jawed vertebrates displaying the presence/absence of stomach across lineages.*** *Although the stomach represents a hallmark for Gnathostomata, some lineages including some teleosts, lungfish, chimaeras and egg laying mammals don't have this organ.*

The existence of an acidic environment within the stomach seems to be an advantage as it prevents infections and stimulates pepsinogen activity that in turn triggers protein digestion. Besides, gastric acid is also believed to improve the incorporation of iron, calcium, and cobalamin, essential for the maintenance of brain functions and blood cell production [49].

Several genes associated with the stomach play critical roles in maintaining its integrity and acidic environment. This study focuses on a number of those stomach-specific genes, the most important of which are ATPase H+/K+ Transporting Alpha Subunit (*ATP4A*) and ATPase H+/K+ Transporting Subunit Beta (*ATP4B*), pepsinogens, and Claudin 18 (*CLDN18*).

3.4.1. Stomach acidification: ATP4A-ATP4B

The H+/K+ ATPase is a P-type cation transporter found in the surface of cell membranes present in the kidney and stomach. The gastric proton pump is responsible for the acidification of stomach contents and achieves so by using the energy, in the form of ATP hydrolysis, to exchange the ions. They are located in parietal cells which are specialized cells present in the gastric mucosa. This proton pump is a heterodimer composed of two subunits: the α subunit (*ATP4A*) and the β subunit (*ATP4B*) (Fig.14). When activated the H+/K+ ATPase induces parietal cells to secrete HCl [52]. The loss of *ATP4A* and *ATP4B* genes has been previously highlighted, with a clear prominence in the species that lack a stomach [53].

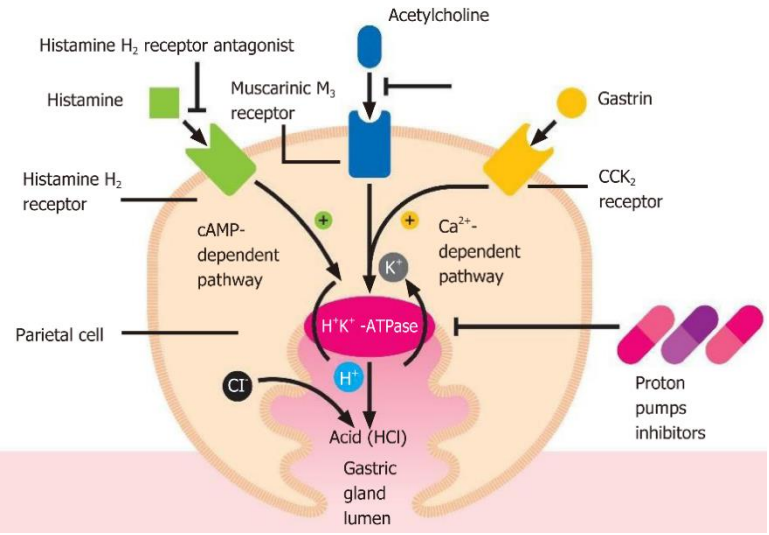

*Figure 14 - **Gastric proton pump.** The H+/K+ ATPase is a cation transporter found in the stomach. It has two subunits: α (ATP4A) and β (ATP4B). When activated it induces HCl secretion from the parietal cells. Retrieved from [48].*

3.4.2. Stomach cellular integrity: *Cldn18*

Cldn18 is one of the many genes that compose the Claudin gene family. These genes are typically associated with tight junctions and have a significant role in maintaining cell polarity and regulating cellular permeability [54]. The claudin gene family has similar sequence and structure among all genes, with the main difference between them being their tissue distribution. In the particular case of *CLDN18*, there are two known isoforms: one is expressed in the lung (*CLDN18.1*) while the other is found in stomach tissue (*CLDN18.2*). The two isoforms differ only in one exon, with an alternative first exon found in each isoform (Fig.15) [55].

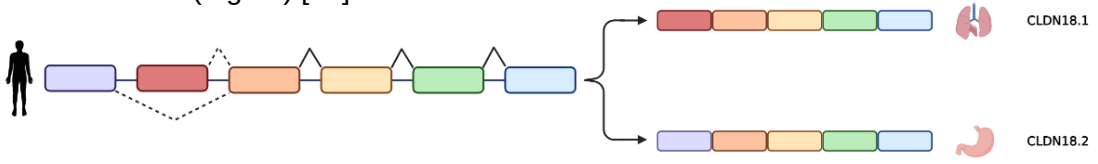

*Figure 15 - **Claudin 18 isoforms.** CLDN18 gene has two known isoforms. One is expressed in the lung (CLDN18.1) while the other is present in the stomach (CLDN18.2). They differ by the first exon only as shown above.*

3.4.3. Protein digestion: Pepsinogens

Pepsinogens represent aspartic proteases with an important role in the specific hydrolysis of peptides. Fundamental in protein digestion, pepsinogens are found in the stomach and are produced in the gastric mucosa. Once they come in contact with the gastric acid, they are activated by cleavage into pepsins (Fig.16) [56]. Generally, a variety of pepsinogen genes is described. These are divided into five groups including pepsinogen A (*PGA*), B (*PGB*), C or progastricsin (*PGC*), F (*PGF*) and prochymosin (*CYM*) [57]. However, the number of genes across pepsinogen groups varies between species and gene family [58]. While pepsinogen A has been reported with multiple copies across species (3 copies of *PGA* are found in human), the *PGC* gene family was long believed to have a single copy. With the study of Castro et al. (2012), it was brought to light that the evolution of the *PGC* gene family faced episodes of expansion, loss and retention [58]. This gene loss evidence could indicate a potential link between the loss of pepsinogen gene families and the appearance of stomach-less lineages [53]. Similarly, processes of loss and inactivation have also been highlighted for the prochymosin gene family [59].

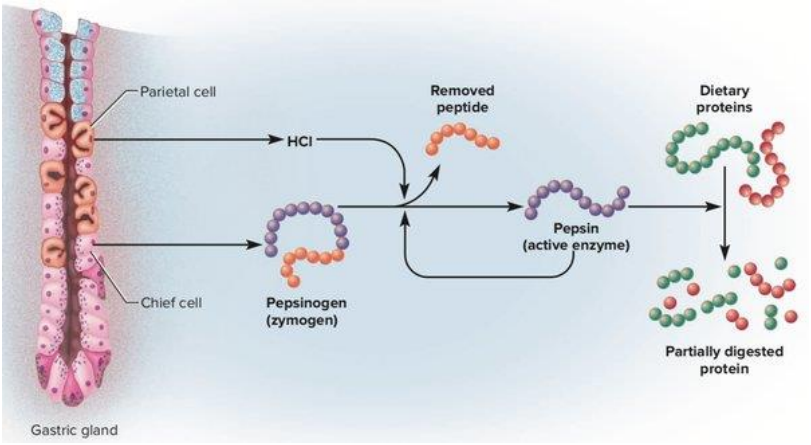

*Figure 16 - **Pepsinogens.** These aspartic proteases are crucial in protein digestion. When in contact with gastric acid they are activated into pepsins. There are 5 types of pepsinogens described: PGA, PGB, PGC, PGF and CYM. Retrieved from https://www.quora.com/What-function-do-pepsinogens-serve*

3.5. Goals

Genomic sequencing has proven a useful tool for studying the origin and preservation of biodiversity in recent years (Fig.17). We can now perform comparative genomic analysis and identify the underlying differences between species at the molecular level thanks to

advances in genomic technology. Such improvements allow us to establish a link between various phenotypes and genetic mechanisms occurring within organisms.

Using novel genomes generated with recent advances in sequencing technologies, the purpose of this study is to investigate the birth and death of a specific physiological/anatomical trait, the stomach, across specific vertebrate lineages.

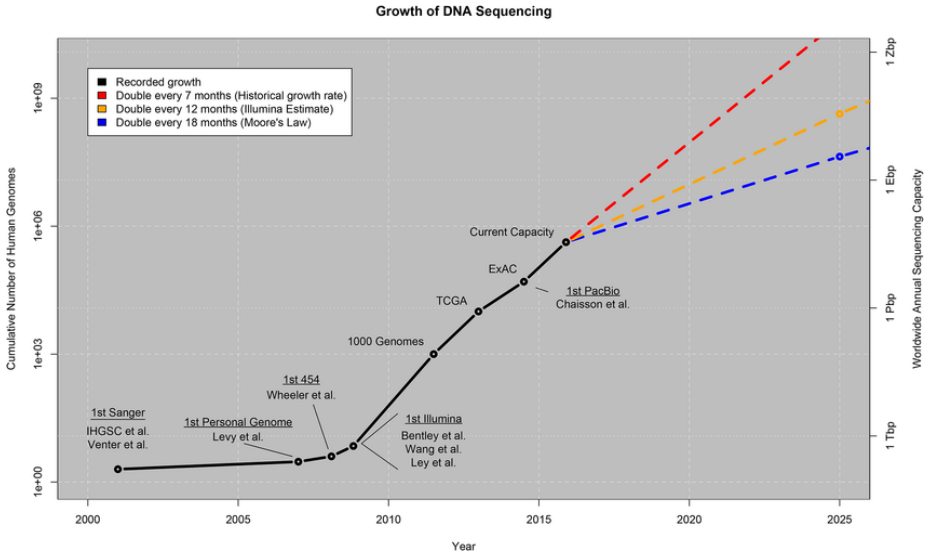

Figure 17 - ***Growth of DNA sequencing.*** *In recent years there has been a significant increase in the number of genomic data available. Genomic technologies have experienced an impressive revolution and are now a tool for the study of evolutionary biology. Retrieved from [3].*

Furthermore, this work plan aims to focus primarily on the Chondrichthyes lineages, for which research has been limited. Because they are among the early-branching vertebrates, investigating these species could provide significant information on early vertebrate evolution.

Focusing on the case study of stomach loss, which has been previously reported among other vertebrates [53], this study expects to verify a possible connection between the loss of gastric phenotype and the loss of specific genes in Chondrichthyes. With that in mind phylogenetic and/or syntenic analysis were performed for the proton pump associated genes (*ATP4A* and *ATP4B*), pepsinogens A and C (*PGA* and *PGC*), Claudin 18 gene (*CLDN18*), and other relevant stomach-related genes.

4. Methods

4.1. Sequence retrieval

Using both Ensembl and NCBI databases, annotated sequences for the genes of interest were retrieved for human (*Homo sapiens*), chimpanzee (*Pan troglodytes*), northern white-cheeked gibbon (*Nomascus leucogenys)*, house mouse (*Mus musculus*), chicken (*Gallus gallus*), turkey (*Meleagris gallopavo*), hooded crow (*Corvus cornix*), zebra finch (*Taeniopygia guttata)*, western painted turtle (*Chrysemys picta bellii)*, common snapping turtle (*Chelydra serpentina)*, green anole (*Anolis carolinensis*), common toad (*Bufo bufo)*, African clawed frog (*Xenopus laevis)*, Gaboon caecilian (*Geotrypetes seraphini)*, Senegal bichir (*Polypterus senegalus*), spotted gar (*Lepisosteus oculatus*), Mississippi paddlefish (*Polyodon spathula*), European eel (*Anguilla anguilla*), zebra fish (*Danio rerio*), tiger barb (*Puntigrus tetrazona*), Asian arowana (*Scleropages formosus*), Atlantic cod (*Gadus morhua),* Japanese rice fish (*Oryzias latipes),* southern platyfish (*Xiphophorus maculatus)*, Nile tilapia (*Oreochromis niloticus*), three-spined stickleback (*Gasterosteus aculeatus),* Japanese puffer (*Takifugu rubripes),* spotted green pufferfish (*Tetraodon nigroviridis)*, Atlantic herring (*Clupea harengus),* electric eel (*Electrophorus electricus*), channel catfish (*Ictalurus punctatus),* northern pike (*Esox lucius)*, Atlantic salmon (*Salmo salar),* pinecone soldierfish (*Myripristis murdjan),* small-spotted catshark *(Scyliorhinus canicula),* great white shark (*Carcharodon carcharias),* thorny skate (*Amblyraja radiata)*, ghost shark (*Callorhinchus milii*), whale shark (*Rhincodon typus)*, zebra shark (*Stegostoma fasciatum),* whitespotted bamboo shark *(*Chiloscyllium plagiosum). The accession numbers of the genome versions can be found in Table 1. Accession numbers of studied genes can be found in the Supplementary Material. For annotated genomes, target genes were retrieved using gene symbols; when no annotated gene was found, BLAST (Basic Local Alignment Search Tool) was used, using as query the target genes from Human (*H. sapiens*).

4.2. Synteny analysis

For the synteny analyses full sequences of the human proteins for the genes of interest were used as query for BLAST (Basic Local Alignment Search Tool) searches. The genomic location of the gene in study as well as four flanking genes on each side was identified for all species of interest (human *H. sapiens*, chicken *G. gallus*, zebrafish *D. rerio*, small-spotted catshark *S. canicula,* great white shark *C. carcharias,* Thorny skate

A. radiata and ghost shark *C. milii)*. If the gene of interest was not found in a particular species following the procedure described above, BLAST searches, using as query conserved neighbouring genes, were performed.

Table 1 - Accession numbers of the genome assemblies used in synteny analysis.

Species	Assembly
Homo sapiens	GCF_000001405.40
Gallus gallus	GCF_016699485.2
Danio rerio	GCF_000002035.6
Scyliorhinus canicular	GCF_902713615.1
Carcharodon carcharias	GCF_017639515.1
Amblyraja radiata	GCF_010909765.2
Callorhinchus milii	GCF_018977255.1

4.3. Phylogenetic analysis

The previously retrieved sequences from Ensembl and NCBI (National Center for Biotechnology Information) were aligned using MAFFT (Multiple Alignment using Fast Fourier Transform) online software using default parameters (https://www.ebi.ac.uk/Tools/msa/mafft/) [60]. Then, Gap Streeze online program (https://www.hiv.lanl.gov/content/sequence/GAPSTREEZE/gap.html?sample_input=1) was used in order to create three files with 0%, 50% and 100% gaps. The resulting files were converted into PHYLIP format via ALTER (Alignment Transformation Environment) software (http://www.sing-group.org/ALTER/). Finally, the PHYLIP amino acid files were uploaded to PhyML online software [61] in which phylogenetic trees were generated using the default parameters and a standard bootstrap of 1000. All phylogenetic trees were later visualized and edited with FigTree (v.1.4.4) program. The described process was applied to all genes of interest.

5. Results and Discussion

The stomach is a key hallmark of jawed vertebrates' evolution. However, some lineages including the Holocephali, do not possess a canonical gastric phenotype [53]. Interestingly, previous studies have found a clear correlation between the loss of stomach and the loss of stomach-specific genes [53], notably related with stomach acidification and protein digestion (proton pump genes and pepsinogens, respectively).

5.1. Patterns of gene duplication and loss underlie the variable stomach phenotype of Elasmobranchi and Holocephalii

5.1.1. ATP4A and ATP4B

In order to understand the mechanisms that underlie the lack of a gastric phenotype in one of the Chondrichthyes subclasses we first analyzed the gastric proton pump associated genes *ATP4A* and *ATP4B*. The gastric pump is a membrane-bound P-type cation transporter crucial for the secretion of HCl, important for food breakdown, enzyme activation and defense against bacterial infections [62].

The genome sequences of 30 species including selected mammals, reptiles, birds, amphibians, teleost, and cartilaginous fish were analyzed.

Sequences with similarity to *ATP4A* (Fig. 18) and *ATP4B* (Fig. 19) were found across all lineages, except for the stomachless chimaera *C. milii* and some teleost species that lack a stomach, such as the zebrafish *D. rerio*, as previously reported [53]. Regarding *ATP4A*, in addition to *C. millii* and *D. rerio*, no *ATP4A* sequence was retrieved from the thorny skate genome (*A. radiata*) (Fig. 18 and 20). Similar to what was described for *ATP4A*, sequences for *ATP4B* gene were found in mammals, birds, reptiles, amphibians, teleosts and cartilaginous fish. As expected, no sequence was retrieved for the stomachless species in study (Fig. 19 and 21). The phylogenetic analysis of both gene families shows well supported branching with the earlier-branching Chondrichthyes sequences clustering together and out-grouping the remaining vertebrate sequences, as expected.

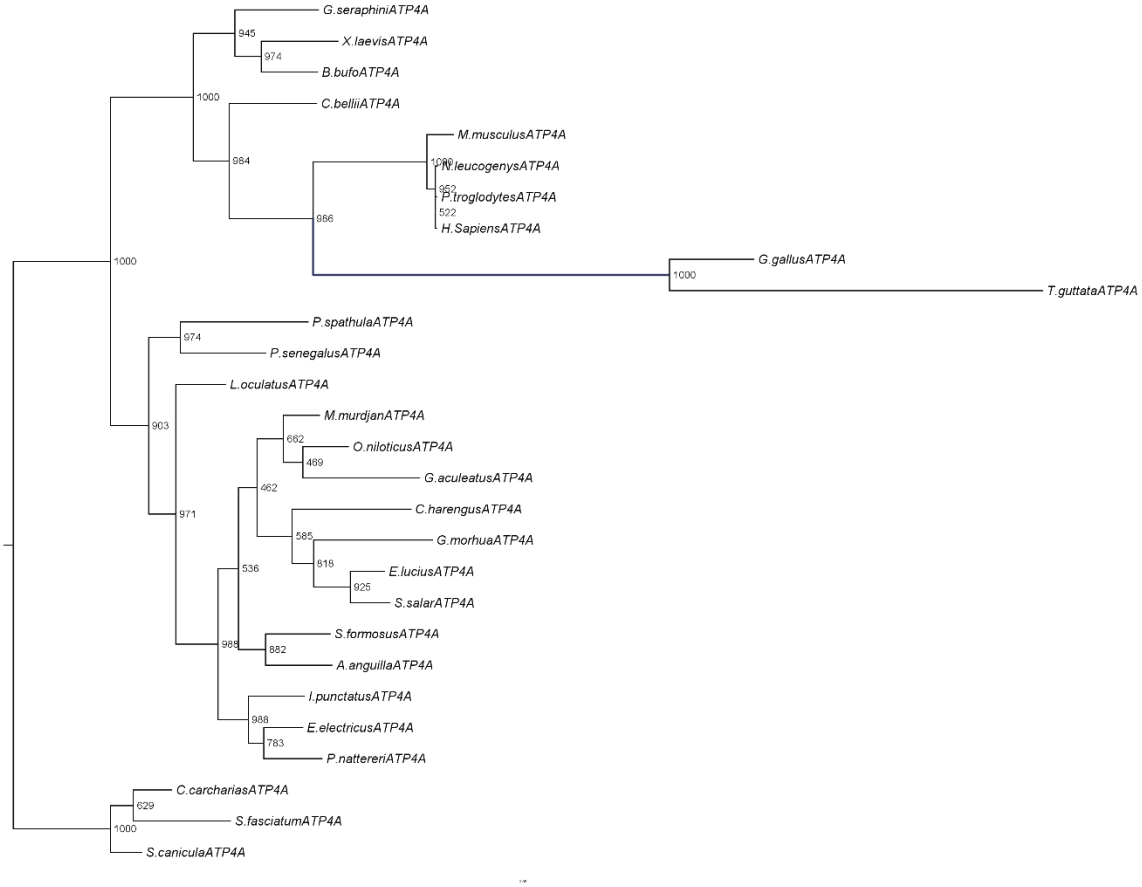

*Figure 18 - **Maximum-likelihood tree of alpha H+/K+-ATPase subunit** genes in vertebrate species with bootstrap values (1000 replicates) shown at each node. The alignment file used contained 28 taxa of 1073 characters. 100% gaps file was used. Human (Homo sapiens), Chimpanzee (Pan troglodytes), Northern white-cheeked gibbon (Nomascus leucogenys), house mouse (Mus musculus), chicken (Gallus gallus), Zebra finch (Taeniopygia guttata), western painted turtle (Chrysemys picta bellii), Common toad (Bufo bufo), African clawed frog (Xenopus laevis), Gaboon caecilian (Geotrypetes seraphini), Senegal bichir (Polypterus senegalus), Spotted gar (Lepisosteus oculatus), Mississippi paddlefish (Polyodon spathula), European eel (Anguilla anguilla), Asian arowana (Scleropages formosus), Atlantic cod (Gadus morhua), Nile tilapia (Oreochromis niloticus), Three-spined stickleback (Gasterosteus aculeatus), Atlantic herring (Clupea harengus), electric eel (Electrophorus electricus), Channel catfish (Ictalurus punctatus), Northern pike (Esox lucius), Atlantic salmon (Salmo salar), red-bellied piranha (Pygocentrus nattereri), Pinecone soldierfish (Myripristis murdjan), small-spotted catshark (Scyliorhinus canicula), great white shark (Carcharodon carcharias), zebra shark (Stegostoma fasciatum. Each phylogeny was rooted with Chondrichthyans sequences.*

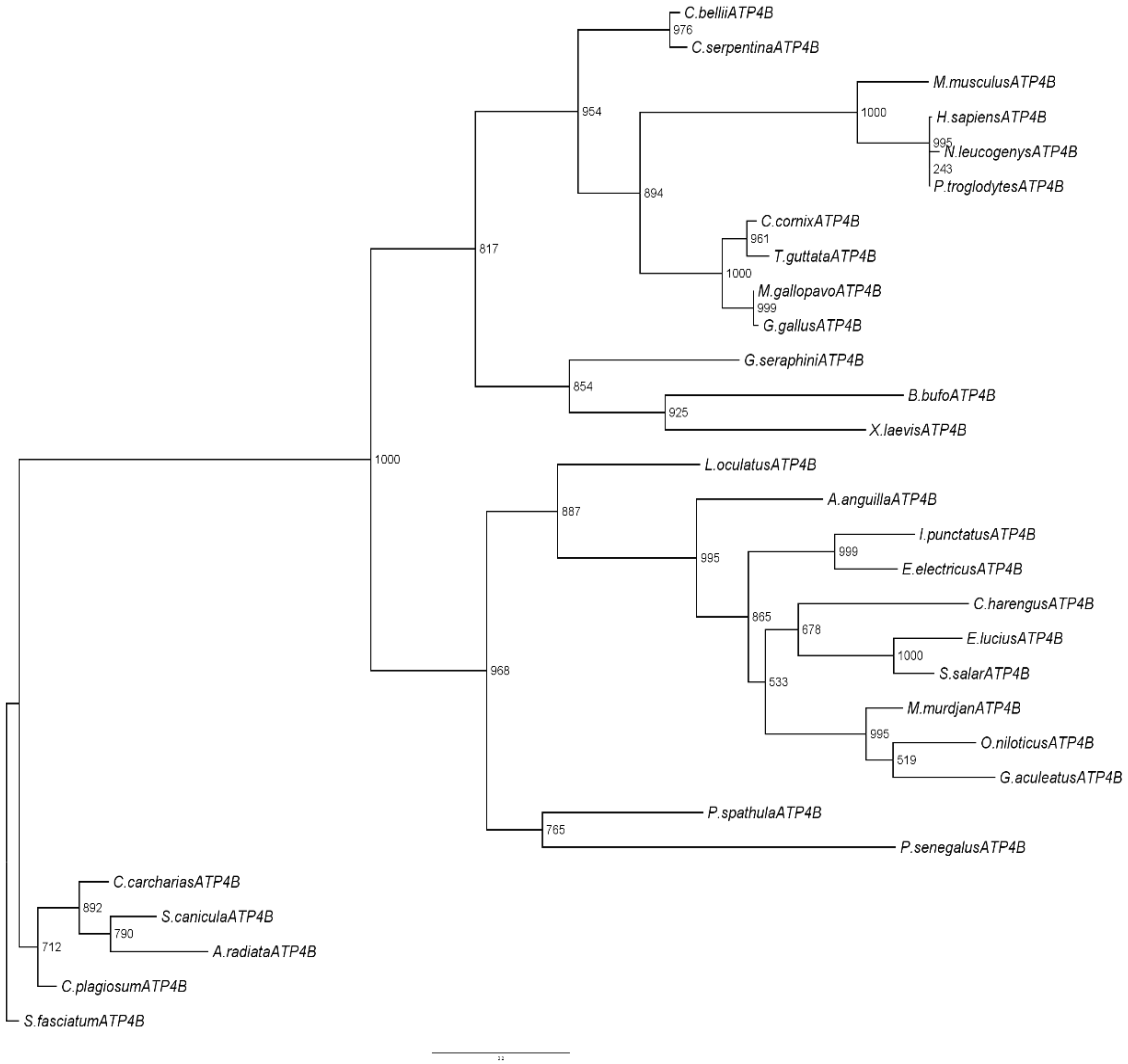

Figure 19 - ***Maximum-likelihood tree of beta H+/K+-ATPase subunit*** *gene in vertebrate species with bootstrap values (1000 replicates) shown at each node. The alignment file used contained 30 taxa of 370 characters. 100% gaps file was used. Human (Homo sapiens), Chimpanzee (Pan troglodytes), Northern white-cheeked gibbon (Nomascus leucogenys), house mouse (Mus musculus), chicken (Gallus gallus), Turkey (Meleagris gallopavo), hooded crow (Corvus cornix), Zebra finch (Taeniopygia guttata), western painted turtle (Chrysemys picta bellii), common snapping turtle (Chelydra serpentina), Common toad (Bufo bufo), African clawed frog (Xenopus laevis), Gaboon caecilian (Geotrypetes seraphini), Senegal bichir (Polypterus senegalus), Spotted gar (Lepisosteus oculatus), Mississippi paddlefish (Polyodon spathula), European eel (Anguilla anguilla), Nile tilapia (Oreochromis niloticus), Three-spined stickleback (Gasterosteus aculeatus), Atlantic herring (Clupea harengus), electric eel (Electrophorus electricus), Channel catfish (Ictalurus punctatus), Northern pike (Esox lucius), Atlantic salmon (Salmo salar), Pinecone soldierfish (Myripristis murdjan), small-spotted catshark (Scyliorhinus canicula), great white shark (Carcharodon carcharias), Thorny skate (Amblyraja radiata), zebra shark (Stegostoma fasciatum), whitespotted bamboo shark (Chiloscyllium plagiosum). Each phylogeny was rooted with Chondrichthyans sequences.*

Additionally, the genomic region of *ATP4A* and *ATP4B* genes was examined, to further access orthology and determine whether true loss or incomplete genome coverage underlined the absence of these genes in the aforementioned species.

The *ATP4A* locus displayed some degree of synteny conservation across human and cartilaginous fish (Fig. 20). *TMEM147* is a flanking gene of *ATP4A* in human, zebrafish and sharks. No evidence of *TMEM147* genes was found in the thorny skate (*A. radiata*) and ghost shark (*C. milii*) genomes. To ensure that the correct *locus* was being analyzed, the presence of another flanking gene found in Chondrichthyes and conserved across the studied species (*CD276*), was taken into consideration. Additionally, human, zebrafish and small spotted catshark (*S. canicula*) display *GAPDHS* as a flanking gene as well. Although it has been established that *ATP4A* gene is present in the chicken genome [53], it was not possible to determine its genomic location. Moreover, no *ATP4A* gene was found in *A. radiata* genome, which comes as a surprise considering the presence of a gastric phenotype in this species. This is possibly a case of incomplete genome coverage, but further studies are required to determine the cause of the *ATP4A* gene absence. As expected, the gene coding for the α subunit of the gastric proton pump gene was not found in the stomachless zebrafish and chimaera.

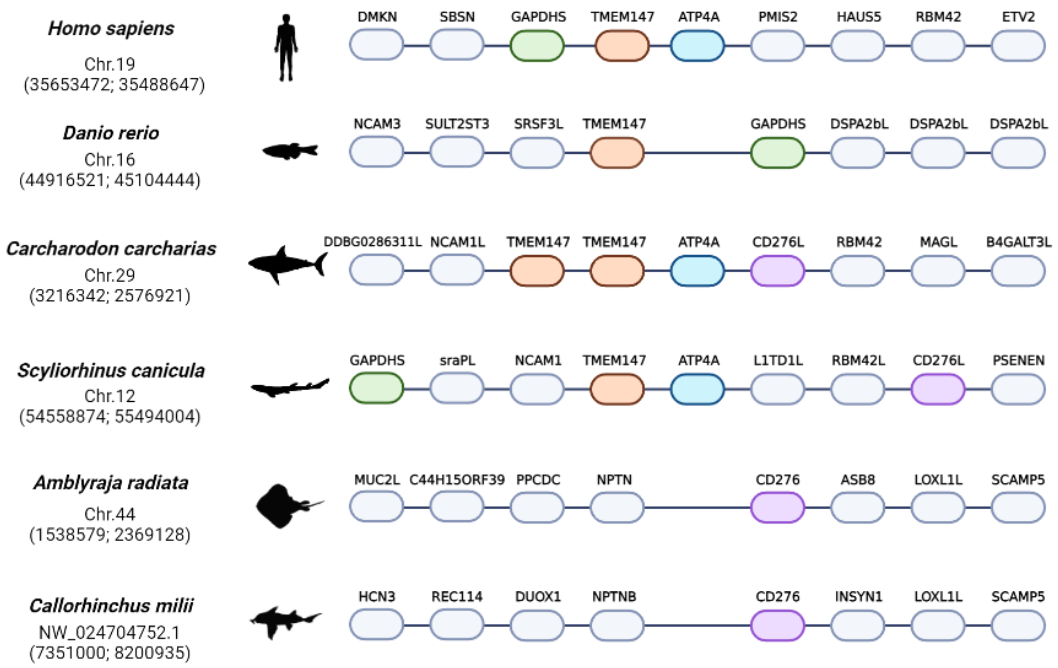

*Figure 20 - **Genomic loci of the ATP4A** gene in vertebrate species. Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); zebrafish (Danio rerio); great white shark (Carcharodon carcharias); small spotted catshark (Scyliorhinus canicula), thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

Concerning *ATP4B* genomic region, some level of synteny conservation was found in human, chicken and Chondrichthyes (Fig. 21). Across all studied species, with the exception of *G. gallus*, *GRK1* was found as one of the neighboring genes of *ATP4B*.

Similarly, *TMCO3* also flanked *ATP4B* in all analyzed species apart from sharks. As in the case of the *ATP4A* analysis, no gene was retrieved in the analyzed agastric species: teleost and cartilaginous fish. This seems to indicate a correlation between the absence of the two proton pump genes and the loss of gastric phenotype, in agreement with previous reports on agastric teleosts [53].

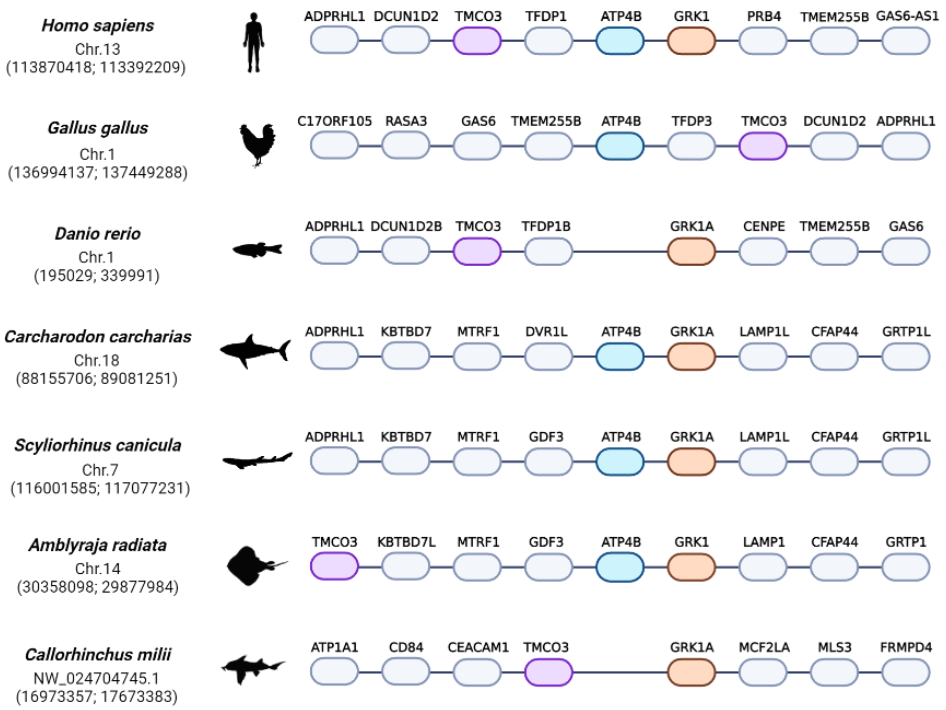

*Figure 21 - **Genomic loci of the ATP4B** gene in vertebrate species. Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); zebrafish (Danio rerio); great white shark (Carcharodon carcharias); small spotted catshark (Scyliorhinus canicula), thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

5.1.2. PGA and PGC

Pepsinogens are aspartic proteases produced in the gastric mucosa that have a critical role in protein digestion. They can be generally divided into 5 groups: *PGA*, *PGB*, *PGC*, *PGF* and *CYM*. Previous works focusing on the study of pepsinogens in tetrapods and teleost species sum-up the current knowledge regarding these genes [53, 58]. In the present study we focused our attention on pepsinogen A (*PGA*) and pepsinogen C (*PGC*).

PGA is usually found with multiple copies within the same organism (for example, the human has 3 known copies of *PGA*); while *PGC* gene family experienced processes of expansion, loss and retention, multiple copies have been described in some species (i.e.

up to five copies in marsupials and in the African clawed frog *Xenopus tropicalis*) [57, 58].

Following the previously mentioned methods, a phylogenetic analysis was carried out for both genes, yielding a well-supported tree topology, clustering Chondrichthyes as an out-group of the remaining vertebrate sequences.

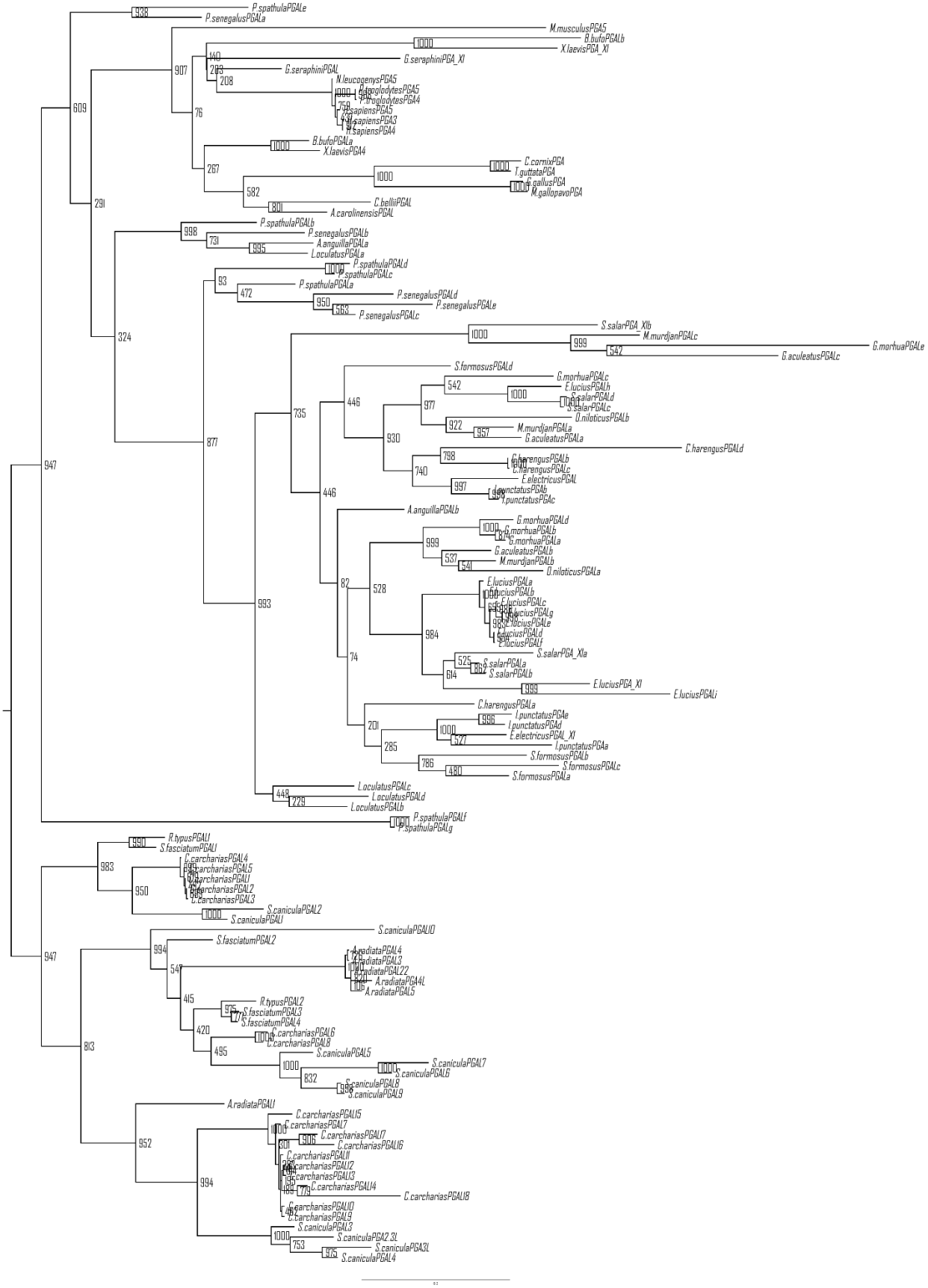

*Figure 22 - **Maximum-likelihood tree of pepsinogen A** gene in vertebrate species with bootstrap values (1000 replicates) shown at each node. The alignment file used contained 136 taxa of 644 characters. 100% gaps file was*

used. Human (Homo sapiens), Chimpanzee (Pan troglodytes), Northern white-cheeked gibbon (Nomascus leucogenys), house mouse (Mus musculus), chicken (Gallus gallus), Turkey (Meleagris gallopavo), hooded crow (Corvus cornix), Zebra finch (Taeniopygia guttata), western painted turtle (Chrysemys picta bellii), green anole (Anolis carolinensis), Common toad (Bufo bufo), African clawed frog (Xenopus laevis), Gaboon caecilian (Geotrypetes seraphini), Senegal bichir (Polypterus senegalus), Spotted gar (Lepisosteus oculatus), Mississippi paddlefish (Polyodon spathula), European eel (Anguilla anguilla), Asian arowana (Scleropages formosus), Atlantic cod (Gadus morhua), Nile tilapia (Oreochromis niloticus), Three-spined stickleback (Gasterosteus aculeatus), Atlantic herring (Clupea harengus), electric eel (Electrophorus electricus), Channel catfish (Ictalurus punctatus), Northern pike (Esox lucius), Atlantic salmon (Salmo salar), Pinecone soldierfish (Myripristis murdjan), small-spotted catshark (Scyliorhinus canicula), great white shark (Carcharodon carcharias), Thorny skate (Amblyraja radiata), whale shark (Rhincodon typus), zebra shark (Stegostoma fasciatum). Each phylogeny was rooted with Chondrichthyans sequences.

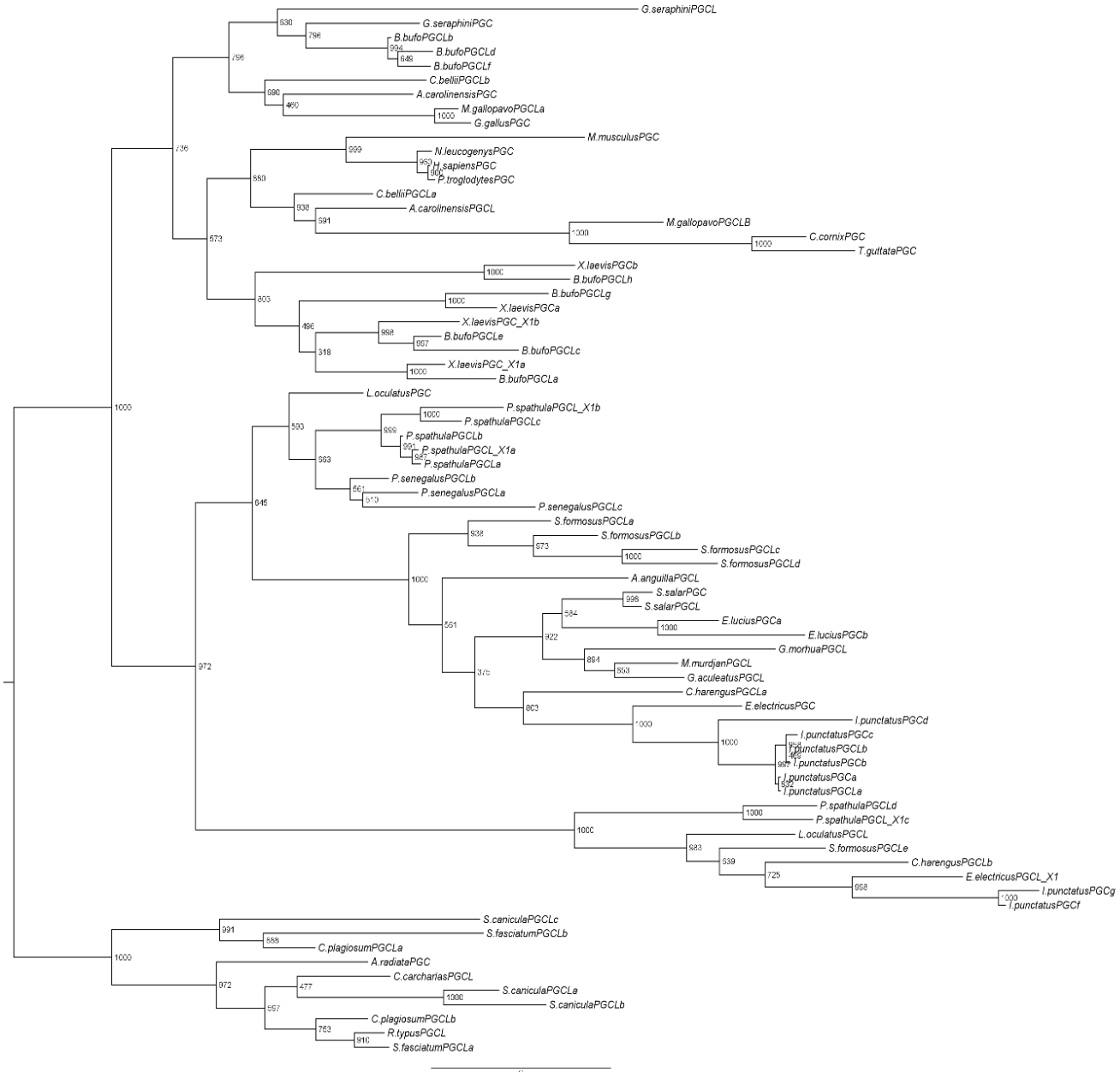

*Figure 23 - **Maximum-likelihood tree of Pepsinogen C** gene in vertebrate species with bootstrap values (1000 replicates) shown at each node. The alignment file used contained 74 taxa of 476 characters. 100% gaps file was used. Human (Homo sapiens), Chimpanzee (Pan troglodytes), Northern white-cheeked gibbon (Nomascus leucogenys), house mouse (Mus musculus), chicken (Gallus gallus), Turkey (Meleagris gallopavo), hooded crow (Corvus cornix), Zebra finch (Taeniopygia guttata), western painted turtle (Chrysemys picta bellii), green anole (Anolis carolinensis), Common toad (Bufo bufo), African clawed frog (Xenopus laevis), Gaboon caecilian (Geotrypetes seraphini), Senegal*

bichir (Polypterus senegalus), Spotted gar (Lepisosteus oculatus), Mississippi paddlefish (Polyodon spathula), European eel (Anguilla anguilla), Asian arowana (Scleropages formosus), Atlantic cod (Gadus morhua), Three-spined stickleback (Gasterosteus aculeatus), Atlantic herring (Clupea harengus), electric eel (Electrophorus electricus), Channel catfish (Ictalurus punctatus), Northern pike (Esox lucius), Atlantic salmon (Salmo salar), Pinecone soldierfish (Myripristis murdjan), small-spotted catshark (Scyliorhinus canicula), great white shark (Carcharodon carcharias), Thorny skate (Amblyraja radiata), whale shark (Rhincodon typus), zebra shark (Stegostoma fasciatum), whitespotted bamboo shark (Chiloscyllium plagiosum). Each phylogeny was rooted with Chondrichthyans sequences.

Sequences with similarity to both genes were found across all jawed vertebrate groups within selected species (Fig. 22 and Fig. 23). As expected for *PGA*, several sequences were retrieved for each species: varying from a single copy in the painted turtle (*C. bellii*) to nineteen copies in great white shark (*C. carcharias*). In fact, multiple *PGA* copies were retrieved across Chondrichthyes, with copy numbers varying from 6 in *A. radiata* to the 19 copies in *C. carcharias* and sequences clustering mostly by species. In the case of *PGC* most species had only one sequence. However, others, such as the small spotted catshark (*S. canicula*), exhibited 3 *PGC* copies. No *PGA* or *PGC* genes were found in the chimaera (*C. millii*).

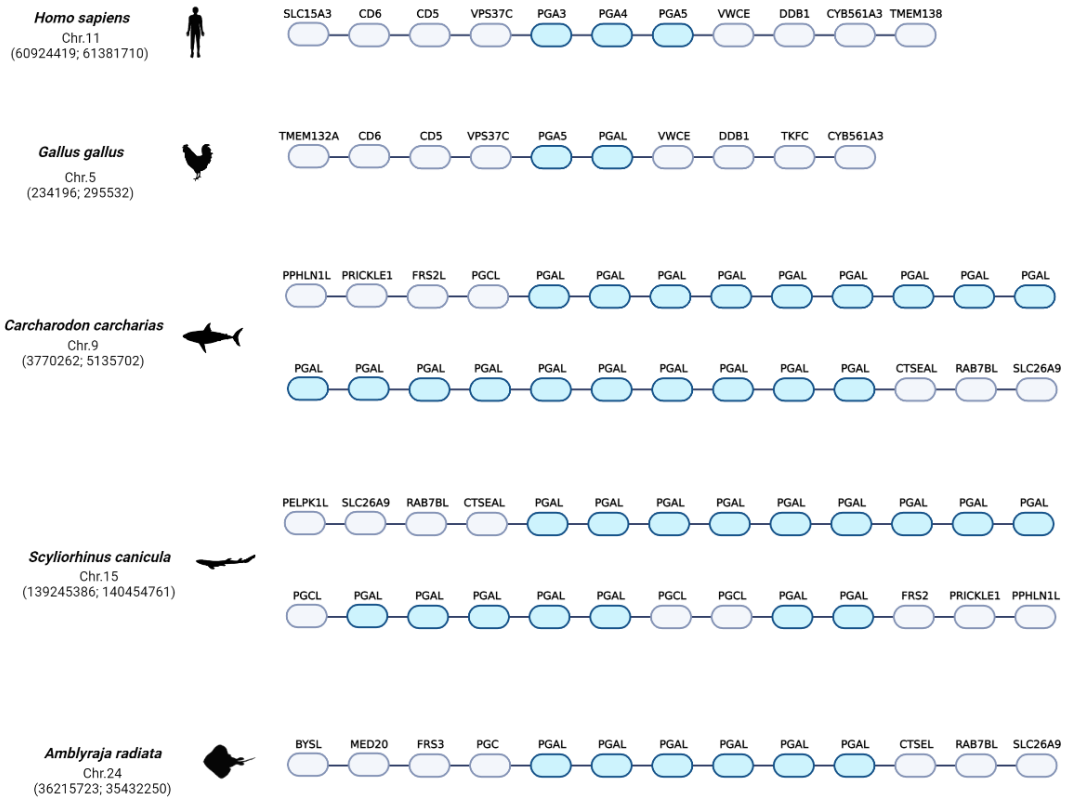

*Figure 24 - **Genomic loci of Pepsinogen A** gene in vertebrate species. Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); small spotted catshark (Scyliorhinus canicula), thorny skate (Amblyraja radiata).*

Following the phylogenetic analysis, the genome *loci* of both *PGA* and *PGC* was examined. In the case of *PGA*, a high degree of synteny conservation was found in human and chicken with *VPS37C* and *VWCE* being present as neighboring genes of our

gene of interest (Fig.24). These flanking genes were analyzed in Chondrichthyes species, despite their presence in Chondrichthyes genomes they did not flank *PGA* in these animals. In fact, Chondrichthyes *PGA* copies neighboured the *PGC* locus, suggesting a distinct rearrangement for these gene clusters in Chondrichthyes.

In the case of *PGC*, a higher level of synteny was found in all the analyzed species (Fig. 25). Although a single copy was retrieved for all analyzed species, in the Small Spotted Catshark (*S. canicula*) three copies of *PGC* were found. Considering the synteny analysis carried out, *PGC* genes are found within the same locus in both human, chicken and cartilaginous fish. Yet, in Chondrichthyes *PGA* and *PGC* are found within the same locus. This suggests that a shift in the genomic location of these genes took place in tetrapods. In teleosts *PGC* genes are found in a different locus than the one described in Human and Chicken, having *MAPK8IP2* and *ZP3* as its neighboring genes [58].

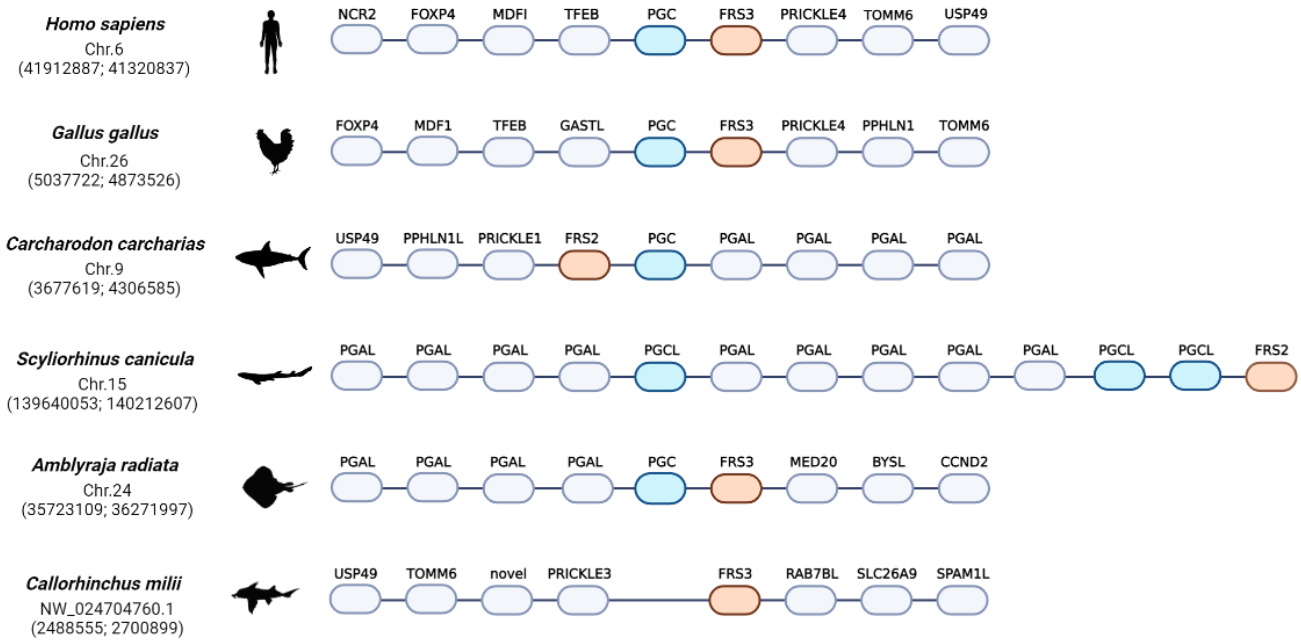

*Figure 25 - **Genomic loci of Pepsinogen C** gene in vertebrate species. Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); small spotted catshark (Scyliorhinus canicula), thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

Both pepsinogen A and pepsinogen C genes were retrieved in all gastric species but were missing in the stomachless zebrafish and chimaera. The phylogeny and synteny analysis suggest a link between the loss of the gastric phenotype and the loss of pepsinogen genes. This is in accordance with the findings reported by Castro et al., 2014 [53]. As pepsinogens give rise to pepsins once they come in contact with gastric acid, therefore the loss of the gastric proton pump and subsequent loss of HCl secretion aligns

with the loss of pepsinogens within agastric species. Moreover, cartilaginous fish possess additional copies of these genes which can represent an expansion of these gene families in Elasmobranchii species. Such massive expansion possibly paralleled specific dietary specialization within this lineage, notably their high-protein and lipid content diet and infrequent feeding, based on large meals [63]. Thus, pepsinogen expansion could underpin a requirement for increased digestibility to optimize digestion and nutrient absorption; yet further studies will be required to ascertain the impact and function of such expansion. Unlike most sharks, chimaeras are mostly bottom-dwelling opportunistic feeders preying on crabs, mollusks, and other invertebrate species.

5.2. Distinct routes to stomach loss – the case of *Cldn18*

Claudin 18 is one of the multiple genes found in the claudin gene family. In general, claudins are typically associated with tight junctions and can be found in a variety of tissues [55]. Particularly, claudin 18 is expressed in two different tissues: the lungs and the stomach. The *CLDN18* gene has two isoforms that differ only on the first exon. *CLDN18.1* is the isoform found in the lungs, while *CLDN18.2* is present in the stomach [55]. Thus, the present study aimed to evaluate whether the loss of gastric phenotype would also lead to the loss of the *CLDN18* stomach isoform.

To assess the presence of the *cldn18* gene isoforms throughout different vertebrate groups, gene sequences were retrieved for selected mammals, birds, reptiles, amphibians, teleost and cartilaginous fish, and phylogenetic analysis was carried out (Fig. 26). Apart from *C. millii*, for which no *cldn18* was found, and *P. spathula* (American paddlefish), for which two gene copies were retrieved, the remaining species held a single *cldn18* gene, yielding one or two splice variants. In human (*H. sapiens*) two isoforms of the same gene were identified which varied in the first exon usage—one for the “lung” isoform and another for the “stomach” isoform of *CLDN18*. In contrast, a single isoform was retrieved in the stomachless platypus and the zebrafish (*D. rerio*). In general mammals, birds and amphibians display two isoforms of the *cldn18* gene, while the majority of reptiles, teleosts and cartilaginous fish have a single isoform of this gene.

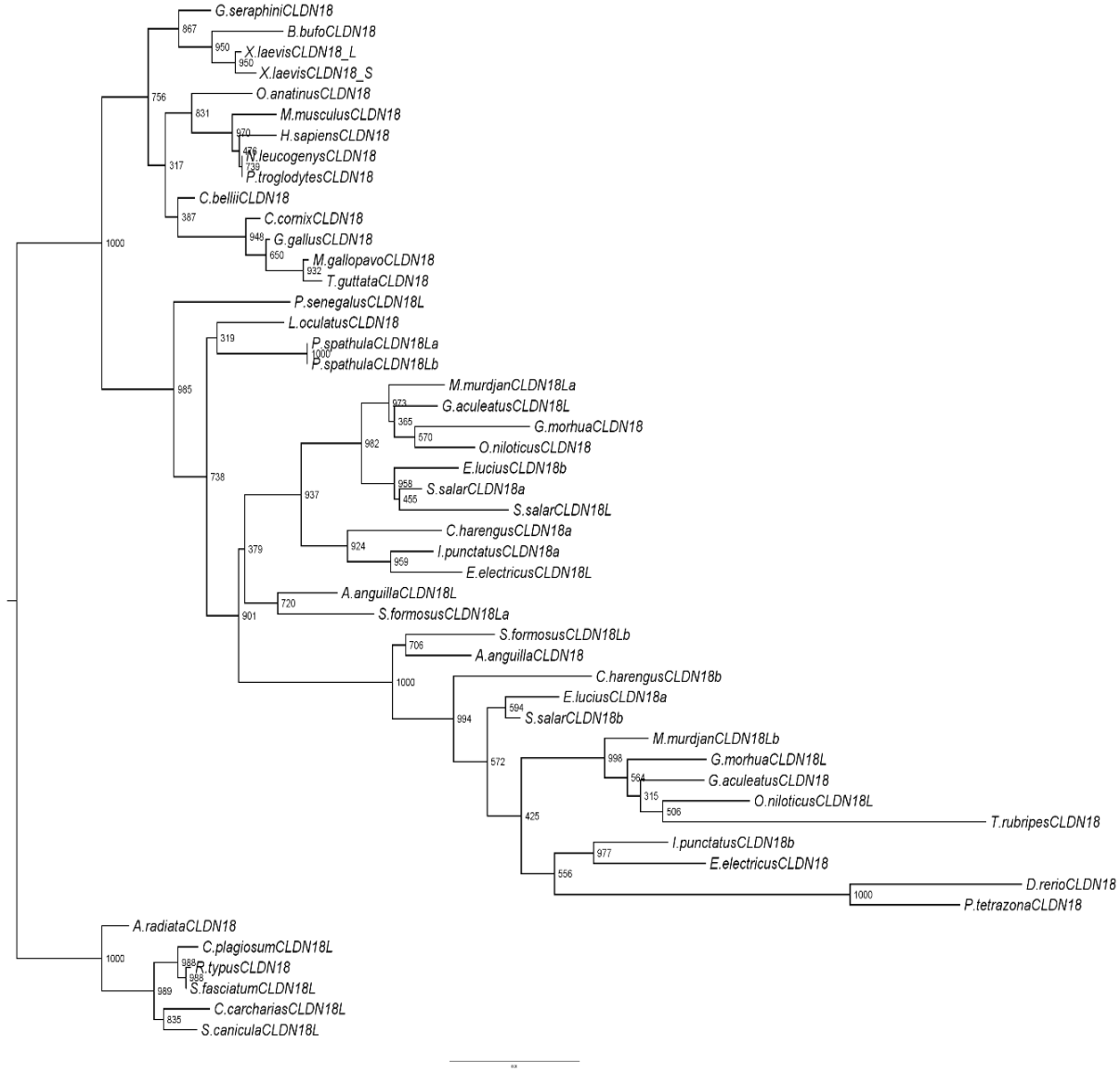

Figure 26- ***Maximum-likelihood tree of Claudin 18*** *gene in vertebrate species with bootstrap values (percentage of 1000 replicates) shown at each node. The alignment file used contained 50 taxa of 453 characters. 100% gaps file was used. Human (Homo sapiens), Chimpanzee (Pan troglodytes), Northern white-cheeked gibbon (Nomascus leucogenys), house mouse (Mus musculus), platypus (Ornithorhynchus anatinus), chicken (Gallus gallus), Turkey (Meleagris gallopavo), hooded crow (Corvus cornix), Zebra finch (Taeniopygia guttata), western painted turtle (Chrysemys picta bellii), Common toad (Bufo bufo), African clawed frog (Xenopus laevis), Gaboon caecilian (Geotrypetes seraphini), Senegal bichir (Polypterus senegalus), Spotted gar (Lepisosteus oculatus), Mississippi paddlefish (Polyodon spathula), European eel (Anguilla anguilla), Asian arowana (Scleropages formosus), Atlantic cod (Gadus morhua), Japanese puffer (Takifugu rubripes), Nile tilapia (Oreochromis niloticus), Zebrafish (Danio rerio), Three-spined stickleback (Gasterosteus aculeatus), Atlantic herring (Clupea harengus), electric eel (Electrophorus electricus), Channel catfish (Ictalurus punctatus), Tiger Barb (Puntigrus tetrazona), Northern pike (Esox lucius), Atlantic salmon (Salmo salar), Pinecone soldierfish (Myripristis murdjan), small-spotted catshark (Scyliorhinus canicula), great white shark (Carcharodon carcharias), Thorny skate (Amblyraja radiata), whale shark (Rhincodon typus), zebra shark (Stegostoma fasciatum), whitespotted bamboo shark (Chiloscyllium plagiosum). Each phylogeny was rooted with Chondrichthyans sequences.*

As a complement to the phylogenetic analysis, the genomic *loci* of *cldn18* were examined in representative species (Fig.27). In this analysis, a clear conserved pattern was found,

with *Sox14* and *Dzip1l* flanking *cldn18* in most species. A *cldn18* gene was found in human, chicken, platypus, zebrafish, great white shark, small spotted catshark, and thorny skate, making the chimaera *C. milii* the only species in which no *cldn18* gene was found. This agrees with the absence of both *cldn18* expression tissues, stomach and lung in chimaeras. Thus, this appears to be a case of true gene loss.

Another important finding in this analysis was the presence of a *cldn18* gene in zebrafish (*D. rerio*). Although this organism has no lung or stomach tissue, a *cldn18* gene, yielding a single isoform, was retrieved in the phylogeny and synteny analysis.

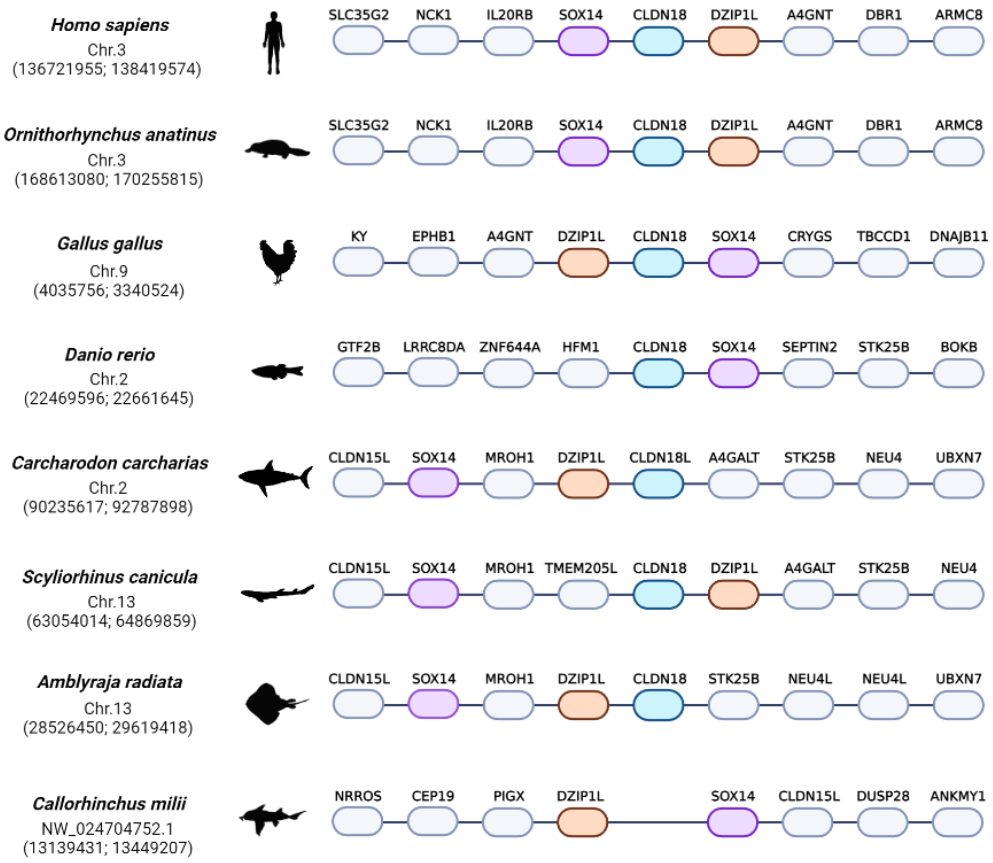

*Figure 27 - **Genomic loci of the CLDN18** gene in vertebrate species. Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); zebrafish (Danio rerio); great white shark (Carcharodon carcharias); small spotted catshark (Scyliorhinus canicula), thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

A possible explanation for such finding would be differential gene expression patterns, with the gene being expressed in other tissues that have a strong requirement for tight junctions. An initial search in VastDB highlighted the testis and pancreas as possible expression sites for the zebrafish gene, as shown in fig. 28. Yet, further analysis such as RNA sequencing, protein labelling and functional characterization will be required to clarify *cldn18* expression and function in this species. According to the Human Protein

Atlas, there are no reports of *CLDN18* gene expression in either testis or pancreas tissues in human. However, it is possible that other genes from the claudin family are expressed in these tissues. For instance, the data found in VastDB shows that genes such as *cldn2* and *cldn12* are expressed in the testis and pancreas of zebrafish.

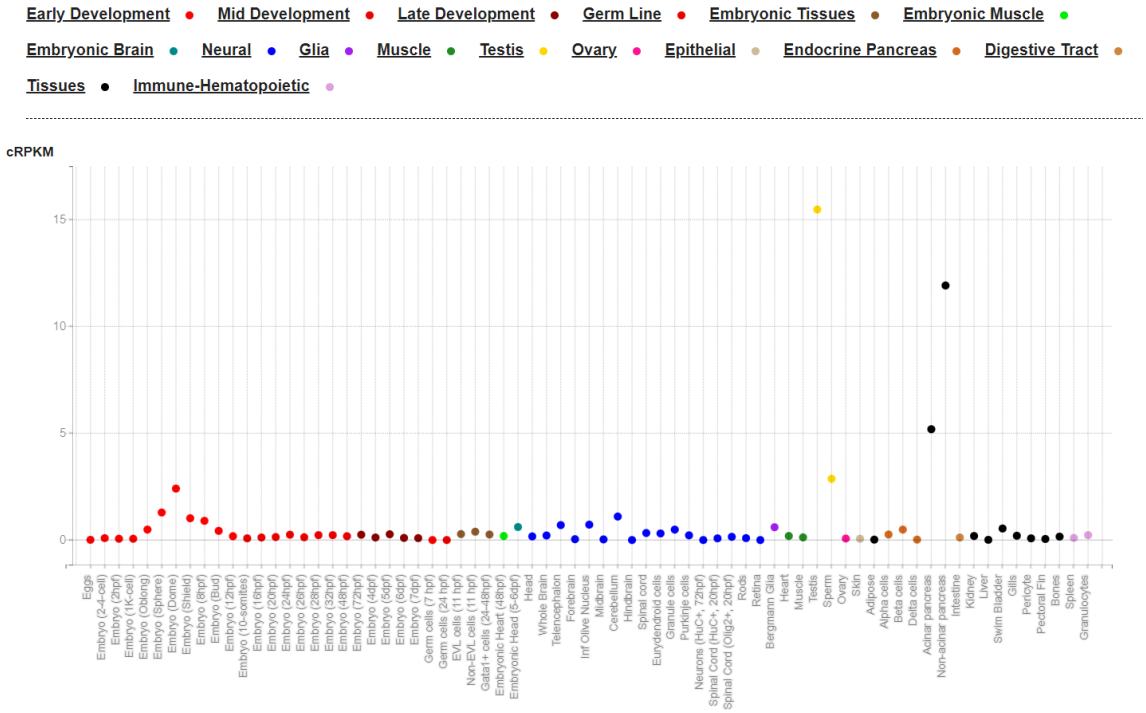

*Figure 28 - **CLDN18 gene expression in Danio rerio**. The CLDN18 gene in the stomachless zebrafish is expressed primarily in testis and pancreas. Retrieved from [ENSDARG00000103087 @ danRer10 - VastDB (crg.eu)]*

Finally, the platypus case displays an interesting case of loss as well. Although a *cldn18* gene was found in this egg-laying mammal's genome, only one isoform seems to be present. Taking into consideration this species agastric phenotype and the alignment found in Fig. 29, it is expected that the retrieved gene corresponds to the lung *CLDN18.1* isoform. Because both isoforms differ only in the first exon, the presence of a single isoform in platypus appears to be associated with an exon erosion phenomenon, in which the organism kept the first exon that gives rise to the lung isoform and lost the exon that encodes the gastric one. This type of loss-of-function, through exon loss, has not been previously reported to the best of my knowledge. In fact, alternative splicing is a powerful source of diversity in animals [64] however, the possible adaptive impacts of the secondarily loss of exons, affecting gene variants differently, is unknown. Thus, in addition to gene erosion and loss, the specific assessment of exon loss could underlie novel paths towards phenotypic variation.

Distances	Homo sapiens CLDN18.1	Homo sapiens CLDN18.2	O. anatinus CLDN18
Homo sapiens CLDN18.1		71.233%	87.671%
Homo sapiens CLDN18.2	71.233%		71.233%
O. anatinus CLDN18	87.671%	71.233%	

*Figure 29 - **Exon alignment.** Alignment and distances of the first exon of Cldn18 in human and platypus.*

5.3. Stomach-specific genes: a mosaic of retention and loss

Besides the previously mentioned genes, a group of 42 additional stomach related genes retrieved from the Human Protein Atlas (https://www.proteinatlas.org/search/tissue_category_rna%3Astomach%3BGroup+enriched+AND+sort_by%3Atissue+specific+score+AND+show_columns%3Agroupenriched) was also reviewed through synteny analysis (Supplementary materials). The goal was to determine the presence and absence of mentioned genes across Chondrichthyes species to get a better understanding of the underlying genomic differences found in elasmobranchs and holocephalans.

As shown in Fig. 30, of the analyzed genes only 4 were distinctively absent from *C. millii* and present in *C. carcharias* and *A. radiata*. This gene set included Cathepsin E (*CTSE*), Tripartite motif containing 50 (*TRIM50*), Annexin A10 (*ANXA10*) and V-Set and Immunoglobulin Domain Containing 1 (*VSIG1*). The *CTSE* gene constitutes an aspartic protease and is therefore responsible for the hydrolyzation of peptides [65]. The exact role that *TRIM50* gene plays in organisms has not been determined, yet it has been demonstrated that this gene is relevant for the formation of canaliculi in parietal cells [66]. Part of the annexin gene family, *ANXA10* functions remain unclear, although it is believed that they can be related to cell motility and differentiation and calcium signaling [67]. V-set and Ig domain-containing protein 1, *VSIG1*, is part of the superfamily of immunoglobulins (IgSF). This gene family is typically associated with cell surface proteins that take part in processes such as cell recognition and adhesion. The exact function of *VSIG1* has not been described thus far [68]. The absence of these genes in chimaeriformes could possibly relate to the fact that these have a predominant

expression in the stomach, which is lost in holocephalans. The *UCN3* gene was absent from the thorny skate genome (*A. radiata*). As previously noted, this could be due to poor coverage since the genes in question are present in both sharks and chimaeras and in some cases have also been identified in other Batoidea species. Nonetheless, further analysis (e.g. RNA sequencing) should be carried out to ascertain this issue. Most genes (26) were present in all studied species. This group includes genes such as Carbonic anhydrase 9 (*CA9*) and Fucosyltransferase 9 (*FUT9*), which are expressed in the liver and brain respectively, in addition to their presence in stomach tissues. *CA9* specifically is involved in pH regulation, cell adhesion and also intercellular communication [69], while *FUT9* catalyzes the synthesis of Lewis X antigen, carbohydrate epitopes participating in cell interaction [70, 71]. The presence of these genes across all Chondrichthyan groups is therefore explained by their expression in several tissues besides the stomach. The remaining genes were not found in any of the three Chondrichthyes species in study. This could be a case of genes that appeared later on and thus are not present in cartilaginous fish. For instance, *CYM* is a pepsinogen found in reptiles, birds and mammals [53]. In mammals, *CYM* displays important functions in the gastric environment during the early development of these animals being usually restricted to fetuses and neonates [72].

*Figure 30 - **Mosaic of Stomach-specific genes.** Synteny analysis were carried out to determine the presence of stomach specific genes across Chondrichthyes species.*

.

5.4. Final discussion

The gastric phenotype and therefore the acid digestion evolved in Gnathostomes ancestry [49]. Although the stomach can be considered a hallmark of jawed vertebrates' evolution, it has been previously shown that it is absent in some lineages. The agastric phenotype of Gnathostomes can be found in monotremes, lungfish, some teleost fish such as cyprinids, and in chimaeriformes [50]. Taking into consideration the distribution

of gastric and agastric lineages from a phylogenetic standpoint, it seems that the loss of stomach is related to several secondary loss events.

Previous studies [53, 73] have shown a connection between the loss of gastric genes and the loss of the stomach, here the repertoire of gastric function genes was taken into consideration and examined on a variety of vertebrate species, with both gastric and agastric representatives.

To validate if the lack of stomach aligned with the deletion of the gastric function genes (*ATP4A*, *ATP4B*, *PGA*, *PGC*, *CLDN18*) in the examined species, a combination of phylogenetic analyses and comparative genomics was used.

Although it is not possible to determine precisely what lead to the secondary loss events of the gastric phenotype in some jawed vertebrates' lineages, some hypothesis can be stated. The stomach is a remarkable feature of vertebrate evolution, and its acidic environment represents an advantage by preventing infections, improving protein digestion as well as increasing the incorporation of important vitamins and minerals into the organism. However, it could be that maintaining such a complex organ is not viable for certain species, possibly due to a high energetic cost [74]. Another possible explanation for the absence of a gastric phenotype in some lineages could be associated with shifts in the organisms' diets. Some cases of pseudogenization have been described for pepsinogen genes in some species, including the human [57]. These processes can be related to the selective pressures the genes face as a response to the different protein components in the animals' diets. Thus, this could mean that the function carried out by specific pepsinogens is no longer required in some organisms. This in turn cancels the need for an acidic environment within the stomach, leading to the disappearance of the gastric function and consequently the proton pump genes (*ATP4A* and *ATP4B*). On the other hand, the opposite is also a possibility. A pseudogenization phenomenon could be taking place within the proton pump genes, leading to the loss of the energy-demanding stomach acidity which in turn would cause secondary loss of pepsinogen genes, since an acid environment is required for their activation. Altogether, it seems reasonable to point diet or environmental shifts as a trigger for the loss of gastric acidification.

Additionally, some of the studied genes have been retained despite the loss of a gastric phenotype, as in the case of *cldn18* in the stomachless *D. rerio*. In these cases, the genes have possibly taken over non-gastric functions and are therefore still found in the organism's genome even though the stomach is absent.

Overall, the results presented in this study show a clear link between the loss of stomach-specific genes and the absence of a gastric phenotype within some Gnathostome lineages.

6. Conclusion

The rise of genomics and the increase in available genome data and improvement in technologies has allowed researchers to take the next step in evolutionary biology and related fields.

Taking advantage of the latest comparative genomics tools, this project aimed to provide some insight into the genetic differences found within Chondrichthyes subclasses and bring to light the mechanisms behind the development of novel phenotypes.

Considering that the stomach embodies one of the key adaptations found in Gnathostomata, it made sense that the present work focused exclusively on this organ, highlighting the mechanisms that lead to the presence and absence of the gastric phenotype across different lineages, with special attention to the cartilaginous fish representatives.

The use of phylogeny and synteny analysis within this project showed that in fact both processes of gene duplication and gene loss are significant drivers of evolution, leading to the appearance of novel traits.

Furthermore, specific cases previously documented in this work emphasize the clear correlation between the loss of the gastric phenotype in some vertebrate lineages and the loss of stomach-specific genes or exons, such as *CLDN18*, the gastric proton pump genes and pepsinogens.

7. Bibliography

1. Dittmar, K. and D. Liberles, *Evolution after gene duplication*. 2011: John Wiley & Sons.
2. Lund-Ricard, Y. and A. Boutet, *Current Trends in Chondrichthyes*. 2021.
3. Stephens, Z., et al., *Big Data: Astronomical or Genomical?* PLoS biology, 2015. **13**: p. e1002195.
4. Albalat, R. and C. Cañestro, *Evolution by gene loss.* Nature Reviews Genetics, 2016. **17**(7): p. 379-391.
5. Hirasawa, T. and S. Kuratani, *Evolution of the vertebrate skeleton: morphology, embryology, and development.* Zoological Lett, 2015. **1**: p. 2.
6. Rittmeyer, E.N., et al., *Ecological Guild Evolution and the Discovery of the World's Smallest Vertebrate.* PLOS ONE, 2012. **7**(1): p. e29797.
7. Bagnat, M. and R.S. Gray, *Development of a straight vertebrate body axis.* Development, 2020. **147**(21).
8. Irie, N., N. Satoh, and S. Kuratani, *The phylum Vertebrata: a case for zoological recognition.* Zoological Lett, 2018. **4**: p. 32.
9. Satoh, N., K. Tagawa, and H. Takahashi, *How was the notochord born?* Evolution & Development, 2012. **14**(1): p. 56-75.
10. Chen, J.-Y., *The origins and key innovations of vertebrates and arthropods.* Palaeoworld, 2011. **20**(4): p. 257-278.
11. Shimeld, S.M. and P.W.H. Holland, *Vertebrate innovations.* Proceedings of the National Academy of Sciences, 2000. **97**(9): p. 4449-4452.
12. Delsuc, F., et al., *Tunicates and not cephalochordates are the closest living relatives of vertebrates.* Nature, 2006. **439**(7079): p. 965-8.
13. Edvardsen, R.B., et al., *Remodelling of the homeobox gene complement in the tunicate Oikopleura dioica.* Curr Biol, 2005. **15**(1): p. R12-3.
14. Denoeud, F., et al., *Plasticity of animal genome architecture unmasked by rapid evolution of a pelagic tunicate.* Science, 2010. **330**(6009): p. 1381-5.
15. Cañestro, C. and J.H. Postlethwait, *Development of a chordate anterior-posterior axis without classical retinoic acid signaling.* Dev Biol, 2007. **305**(2): p. 522-38.
16. Donoghue, P.C.J. and I.J. Sansom, *Origin and early evolution of vertebrate skeletonization.* Microscopy Research and Technique, 2002. **59**(5): p. 352-372.
17. Mayor, R. and E. Theveneau, *The neural crest.* Development, 2013. **140**(11): p. 2247-2251.

18. Baker, C.V.H. and M. Bronner-Fraser, *Vertebrate Cranial Placodes I. Embryonic Induction.* Developmental Biology, 2001. **232**(1): p. 1-61.
19. Schlosser, G., *Evolutionary origins of vertebrate placodes: insights from developmental studies and from comparisons with other deuterostomes.* Journal of Experimental Zoology Part B: Molecular and Developmental Evolution, 2005. **304B**(4): p. 347-399.
20. Schlosser, G., *Chapter Four - Making Senses: Development of Vertebrate Cranial Placodes*, in *International Review of Cell and Molecular Biology*, K. Jeon, Editor. 2010, Academic Press. p. 129-234.
21. Martik, M.L. and M.E. Bronner, *Riding the crest to get a head: neural crest evolution in vertebrates.* Nature Reviews Neuroscience, 2021. **22**(10): p. 616-626.
22. Martín-Durán, J.M. and A. Hejnol, *A developmental perspective on the evolution of the nervous system.* Developmental Biology, 2021. **475**: p. 181-192.
23. Brazeau, M.D. and M. Friedman, *The origin and early phylogenetic history of jawed vertebrates.* Nature, 2015. **520**(7548): p. 490-497.
24. Dulvy, N.K. and J.D. Reynolds, *Evolutionary transitions among egg–laying, live–bearing and maternal inputs in sharks and rays.* Proceedings of the Royal Society of London. Series B: Biological Sciences, 1997. **264**(1386): p. 1309-1315.
25. Chapman, D.D., et al., *Virgin birth in a hammerhead shark.* Biology Letters, 2007. **3**(4): p. 425-427.
26. Capapé, C., A. Seck, and Y. Diatta, *Reproductive biology of the common torpedo, Torpedo torpedo (Linnaeus, 1758)(Pisces, Torpedinidae) from the coast of Senegal (Eastern Tropical Atlantic).* Miscel· lània Zoològica, 2000: p. 9-21.
27. Joung, S.-J., et al., *The whale shark, Rhincodon typus, is a livebearer: 300 embryos found in one 'megamamma' supreme.* Environmental Biology of Fishes, 1996. **46**(3): p. 219-223.
28. Cailliet, G.M., et al., *Age determination and validation studies of marine fishes: do deep-dwellers live longer?* Experimental gerontology, 2001. **36**(4-6): p. 739-764.
29. Nielsen, J., et al., *Eye lens radiocarbon reveals centuries of longevity in the Greenland shark (Somniosus microcephalus).* Science, 2016. **353**(6300): p. 702-704.
30. Inoue, J.G., et al., *Evolutionary Origin and Phylogeny of the Modern Holocephalans (Chondrichthyes: Chimaeriformes): A Mitogenomic Perspective.* Molecular Biology and Evolution, 2010. **27**(11): p. 2576-2586.
31. Didier, D.A., *Phylogeny and classification of extant Holocephali.* Biology of sharks and their relatives. CRC Press, Boca Raton, 2004. **4**: p. 115-138.

32. Wegner, N.C., *3 - Elasmobranch Gill Structure*, in *Fish Physiology*, R.E. Shadwick, A.P. Farrell, and C.J. Brauner, Editors. 2015, Academic Press. p. 101-151.
33. Tozer, H. and D. Didier, *Husbandry of spotted ratfish, Hydrolagus colliei.* The elasmobranch husbandry manual: captive care of sharks, rays and their relatives, 2004: p. 487-491.
34. Szaniawski, H., *The earliest known venomous animals recognized among conodonts.* Acta Palaeontologica Polonica, 2009. **54**(4): p. 669-676.
35. Priede, I.G., et al., *The absence of sharks from abyssal regions of the world's oceans.* Proceedings of the Royal Society B: Biological Sciences, 2006. **273**(1592): p. 1435-1441.
36. De Iuliis, G. and D. Pulerà, *The dissection of vertebrates*. 2019: Academic press.
37. Meekan, M.G., et al., *The world's largest omnivore is a fish.* Ecology, 2022. **103**(12): p. e3818.
38. Lallemand, T., et al., *An Overview of Duplicated Gene Detection Methods: Why the Duplication Mechanism Has to Be Accounted for in Their Choice.* Genes (Basel), 2020. **11**(9).
39. Fan, C., Y. Chen, and M. Long, *Recurrent tandem gene duplication gave rise to functionally divergent genes in Drosophila.* Mol Biol Evol, 2008. **25**(7): p. 1451-8.
40. Peatman, E. and Z. Liu, *Evolution of CC chemokines in teleost fish: a case study in gene duplication and implications for immune diversity.* Immunogenetics, 2007. **59**(8): p. 613-623.
41. Hofberger, J.A., et al., *Whole Genome and Tandem Duplicate Retention Facilitated Glucosinolate Pathway Diversification in the Mustard Family.* Genome Biology and Evolution, 2013. **5**(11): p. 2155-2173.
42. Zhang, Z., et al., *Millions of years of evolution preserved: a comprehensive catalog of the processed pseudogenes in the human genome.* Genome research, 2003. **13**(12): p. 2541-2558.
43. Kubiak, M.R., M.W. Szcześniak, and I. Makałowska, *Complex Analysis of Retroposed Genes' Contribution to Human Genome, Proteome and Transcriptome.* Genes (Basel), 2020. **11**(5).
44. Samonte, R.V. and E.E. Eichler, *Segmental duplications and the evolution of the primate genome.* Nature Reviews Genetics, 2002. **3**(1): p. 65-72.
45. Ohno, S., *Evolution by gene duplication New York: springer*. 1970.
46. Meyer, A. and M. Schartl, *Gene and genome duplications in vertebrates: the one-to-four (-to-eight in fish) rule and the evolution of novel gene functions.* Current opinion in cell biology, 1999. **11**(6): p. 699-704.

47. Lien, S., et al., *The Atlantic salmon genome provides insights into rediploidization.* Nature, 2016. **533**(7602): p. 200-205.
48. Jeffery, W.R., *Regressive evolution in Astyanax cavefish.* Annual review of genetics, 2009. **43**: p. 25.
49. Koelz, H., *Gastric acid in vertebrates.* Scandinavian Journal of Gastroenterology, 1992. **27**(sup193): p. 2-6.
50. Stevens, C.E. and I.D. Hume, *Comparative physiology of the vertebrate digestive system*. 2004: Cambridge University Press.
51. Holmgren, S. and C. Olsson, *Autonomic control of glands and secretion: A comparative view.* Autonomic Neuroscience, 2011. **165**(1): p. 102-112.
52. Li, H., et al., *H+/K+-ATPase inhibitors: a patent review.* Expert Opinion on Therapeutic Patents, 2013. **23**(1): p. 99-111.
53. Castro, L.F.C., et al., *Recurrent gene loss correlates with the evolution of stomach phenotypes in gnathostome history.* Proceedings of the Royal Society B: Biological Sciences, 2014. **281**(1775): p. 20132669.
54. Zhou, B., et al., *Claudin-18–mediated YAP activity regulates lung stem and progenitor cell homeostasis and tumorigenesis.* The Journal of Clinical Investigation, 2018. **128**(3): p. 970-984.
55. Türeci, Ö., et al., *Claudin-18 gene structure, regulation, and expression is evolutionary conserved in mammals.* Gene, 2011. **481**(2): p. 83-92.
56. RICHTER, C., T. TANAKA, and R.Y. YADA, *Mechanism of activation of the gastric aspartic proteinases: pepsinogen, progastricsin and prochymosin.* Biochemical Journal, 1998. **335**(3): p. 481-490.
57. Kageyama, T., *Pepsinogens, progastricsins, and prochymosins: structure, function, evolution, and development.* Cellular and Molecular Life Sciences CMLS, 2002. **59**(2): p. 288-306.
58. Castro, L.F., et al., *The evolution of pepsinogen C genes in vertebrates: duplication, loss and functional diversification.* PLoS One, 2012. **7**(3): p. e32852.
59. Lopes-Marques, M., et al., *Unusual loss of chymosin in mammalian lineages parallels neo-natal immune transfer strategies.* Mol Phylogenet Evol, 2017. **116**: p. 78-86.
60. Madeira, F., et al., *Search and sequence analysis tools services from EMBL-EBI in 2022.* Nucleic acids research, 2022: p. gkac240.
61. Guindon, S., et al., *New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0.* Systematic Biology, 2010. **59**(3): p. 307-321.

62. Schubert, M.L., *Gastric acid secretion.* Current Opinion in Gastroenterology, 2016. **32**(6): p. 452-460.
63. Leigh, S.C., Y. Papastamatiou, and D.P. German, *The nutritional physiology of sharks.* Reviews in Fish Biology and Fisheries, 2017. **27**(3): p. 561-585.
64. Wright, C.J., C.W.J. Smith, and C.D. Jiggins, *Alternative splicing as a source of phenotypic diversity.* Nature Reviews Genetics, 2022. **23**(11): p. 697-710.
65. Page, J.L., S.C. Strom, and C.J. Omiecinski, *Regulation of the human cathepsin E gene by the constitutive androstane receptor.* Archives of biochemistry and biophysics, 2007. **467**(1): p. 132-138.
66. Nishi, M., et al., *TRIM50 protein regulates vesicular trafficking for acid secretion in gastric parietal cells.* Journal of Biological Chemistry, 2012. **287**(40): p. 33523-33532.
67. Ishikawa, A., et al., *Annexin A10 is involved in the induction of pancreatic duodenal homeobox-1 in gastric cancer tissue, cells and organoids.* Oncology Reports, 2020. **43**(2): p. 581-590.
68. Jung, Y., et al., *V-Set and Immunoglobulin Domain-Containing 1 (VSIG1), Predominantly Expressed in Testicular Germ Cells, Is Dispensable for Spermatogenesis and Male Fertility in Mice.* Animals, 2021. **11**(4): p. 1037.
69. Shin, H.-J., et al., *Carbonic anhydrase IX (CA9) modulates tumor-associated cell migration and invasion.* Journal of cell science, 2011. **124**(7): p. 1077-1087.
70. Auslander, N., et al., *An integrated computational and experimental study uncovers FUT 9 as a metabolic driver of colorectal cancer.* Molecular systems biology, 2017. **13**(12): p. 956.
71. Nishihara, S., et al., *α1, 3-Fucosyltransferase IX (Fut9) determines Lewis X expression in brain.* Glycobiology, 2003. **13**(6): p. 445-455.
72. Janiak, M.C., *Digestive enzymes of human and nonhuman primates.* Evolutionary Anthropology: Issues, News, and Reviews, 2016. **25**(5): p. 253-266.
73. Ordoñez, G.R., et al., *Loss of genes implicated in gastric function during platypus evolution.* Genome biology, 2008. **9**(5): p. 1-11.
74. Moffatt, K., et al., *Inhibition of gastric acid secretion with omeprazole affects fish specific dynamic action and growth rate: Implications for the development of phenotypic stomach loss.* Frontiers in Physiology, 2022. **13**.

8. Supplementary Material

8.1 Accession Numbers of Retrieved Gene Sequences

Table 2 - Accession numbers of ATP4A sequences.

Species	***Accession Number***
Homo sapiens ATP4A	NP_000695.2
Pan troglodytes ATP4A	XP_016791181.1
Nomascus leucogenys ATP4A	XP_030675993.1
Mus musculus ATP4A	NP_001277556.1
Xenopus laevis ATP4A	NP_001084343.1
Bufo bufo ATP4A	XP_040271459.1
Geotrypetes seraphini ATP4A	XP_033770265.1
Chrysemys picta bellii ATP4A	XP_042701799.1
Chelydra serpentina ATP4A	KAG6924301.1
Gallus gallus ATP4A	AHE78442.1
Taeniopygia guttata ATP4A	XP_041568322.1
Polypterus senegalus ATP4A	XP_039629560.1
Lepisosteus oculatus ATP4A	XP_015191976.1
Polyodon spathula ATP4A	XP_041082596.1
Anguilla anguilla ATP4A	XP_035244218.1
Scleropages formosus ATP4A	XP_018582433.1
Gadus morhua ATP4A	XP_030226282.1
Oreochromis niloticus ATP4A	XP_003450965.1
Gaterosteus aculeatus ATP4A	XP_040020180.1
Clupea harengus ATP4A	XP_012673538.1
Philodryas nattereri ATP4A	XP_017564192.1
Electrophorus electricus ATP4A	XP_026856953.2
Ictalurus punctatus ATP4A	XP_017308505.1
Esox lucius ATP4A	XP_010883375.1
Salmo salar ATP4A	XP_014056965.1
Myripristis murdjan ATP4A	XP_029927877.1
Scyliorhinus canicula ATP4A	XP_038669150.1
Carcharodon carcharias ATP4A	XP_041037095.1
Stegostoma fasciatum ATP4A	XP_048379720.1

Table 3 - Accession numbers of ATP4B sequences.

Species	***Accession Number***
Homo sapiens ATP4B	NP_000696.1
Pan troglodytes ATP4B	XP_001146058.3
Nomascus leucogenys ATP4B	XP_030668397.1
Mus musculus ATP4B	AAI25293.1
Gallus gallus ATP4B	NP_989749.1
Meleagris gallopavo ATP4B	XP_003203249.2
Taeniopygia guttata ATP4B	XP_004174770.3
Corvus cornix ATP4B	XP_010392606.1
Chrysemys picta bellii ATP4B	XP_005309941.1
Chelydra serpentina ATP4B	KAG6938908.1
Bufo bufo ATP4B	XP_040277914.1
Xenopus laevis ATP4B	XP_041438910.1
Geotrypetes seraphini ATP4B	XP_033805743.1
Polypterus senegalus ATP4B	XP_039602205.1
Lepisosteus oculatus ATP4B	XP_015219566.1
Polyodon spathula ATP4B	XP_041128104.1
Anguilla anguilla ATP4B	XP_035265917.1
Oreochromis niloticus ATP4B	XP_003445766.1
Gasterosteus aculeatus ATP4B	XP_040019560.1
Clupea harengus ATP4B	XP_031414030.1
Electrophorus electricus ATP4B	XP_026866015.1
Ictalurus punctatus ATP4B	XP_017313691.1
Esox lucius ATP4B	XP_010885041.1
Salmo salar ATP4B	XP_045555533.1
Myripristis murdjan ATP4B	XP_029936975.1
Amblyraja radiata ATP4B	XP_032889408.1
Scyliorhinus canicula ATP4B	XP_038658680.1
Carcharodon carcharias ATP4B	XP_041067797.1

Stegostoma fasciatum ATP4B	XP_048396998.1
Chiloscyllium plagiosum ATP4B	XP_043557482.1

Table 4 - Accession numbers of CLDN18 sequences.

Species	***Accession Number***
Homo sapiens CLDN18	NP_001002026.1
Pan troglodytes CLDN18	XP_526318.3
Nomascus leucogenys CLDN18	XP_003265323.1
Mus musculus CLDN18	NP_062789.1
Ornithorhynchus anatinus CLDN18	XP_028925772.1
Gallus gallus CLDN18	XP_040535069.1
Meleagris gallopavo CLDN18	XP_003209013.1
Corvus cornix CLDN18	XP_039413104.1
Taeniopygia guttata CLDN18	XP_002196241.1
Chrysemys picta bellii CLDN18	XP_005303714.1
Bufo bufo CLDN18	XP_040286930.1
Xenopus laevis CLDN18_L	NP_001083443.1
Xenopus laevis CLDN18_S	XP_018120300.1
Geotrypetes seraphini CLDN18	XP_033815148.1
Polypterus senegalus CLDN18L	XP_039617881.1
Lepisosteus oculatus CLDN18	XP_006637728.1
Polyodon spathula CLDN18La	XP_041125902.1
Polyodon spathula CLDN18Lb	XP_041125902.1
Anguilla anguilla CLDN18L	XP_035272052.1
Anguilla anguilla CLDN18	XP_035278220.1
Danio rerio CLDN18	XP_002660862.3
Puntigrus tetrazona CLDN18	XP_043080518.1
Scleropages formosus CLDN18La	XP_018593928.1
Scleropages formosus CLDN18Lb	XP_018586618.1
Gadus morhua CLDN18	XP_030228284.1
Gadus morhua CLDN18L	XP_030220814.1
Oreochromis niloticus CLDN18	XP_003444253.1
Oreochromis niloticus CLDN18L	XP_003438075.1
Gasterosteus aculeatus CLDN18L	XP_040039231.1
Gasterosteus aculeatus CLDN18	XP_040027686.1
Takifugu rubripes CLDN18	XP_011614961.1
Clupea harengus CLDN18	XP_031430652.1
Clupea harengus CLDN18	XP_012681094.1

Electrophorus electricus CLDN18L	XP_026863571.1
Electrophorus electricus CLDN18	XP_026882749.2
Ictalurus punctatus CLDN18a	XP_017334649.1
Ictalurus punctatus CLDN18b	NP_001316217.1
Esox lucius CLDN18a	XP_010891673.1
Esox lucius CLDN18b	XP_019900146.1
Salmo salar CLDN18a	XP_013979138.1
Salmo salar CLDN18L	XP_014024098.1
Salmo salar CLDN18b	XP_014045392.1
Myripristis murdjan CLDN18La	XP_029905880.1
Myripristis murdjan CLDN18Lb	XP_029930832.1
Stegostoma fasciatum CLDN18L	XP_048398771.1
Rhincodon typus CLDN18	XP_048460550.1
Chiloscyllium plagiosum CLDN18L	XP_043557721.1
Carcharodon carcharias CLDN18L	XP_041029914.1
Amblyraja radiata CLDN18	XP_032887427.1
Scyliorhinus canicula CLDN18L	XP_038671664.1

Table 5 - Accession numbers of PGA sequences

Species	***Accession Number***
Homo sapiens PGA3	AAI71815.1
Homo sapiens PGA4	AAI71814.1
Homo sapiens PGA5	AAI71897.1
Pan troglodytes PGA4	XP_024202947.1
Pan troglodytes PGA5	XP_024202950.1
Nomascus leucogenys PGA5	XP_003282561.2
Mus musculus PGA5	NP_067428.2
Gallus gallus PGA	NP_990209.1
Meleagris gallopavo PGA	XP_010709067.1
Taeniopygia guttata PGA	XP_030129447.3
Corvus cornix PGA	XP_019136587.1
Chrysemys bellii PGAL	XP_005309696.1
Anolis carolinensis PGAL	XP_003224161.1
Bufo bufo PGALa	XP_040291904.1
Bufo bufo PGALb	XP_040276550.1
Geotrypetes seraphini PGAL	XP_033776360.1
Geotrypetes seraphini PGA_X1	XP_033776305.1
Xenopus laevis PGA4	NP_001079037.1
Xenopus laevis PGA_X1	XP_041442323.1
Polypterus senegalus PGALa	XP_039604463.1

Polypterus senegalus PGALb	XP_039605925.1
Polypterus senegalus PGALc	XP_039605923.1
Polypterus senegalus PGALd	XP_039604464.1
Polypterus senegalus PGALe	XP_039604466.1
Lepisosteus oculatus PGALa	XP_006628507.1
Lepisosteus oculatus PGALb	XP_006628508.1
Lepisosteus oculatus PGALc	XP_015197905.1
Lepisosteus oculatus PGALd	XP_015197904.1
Polyodon spathula PGALa	XP_041083519.1
Polyodon spathula PGALb	XP_041091851.1
Polyodon spathula PGALc	XP_041083520.1
Polyodon spathula PGALd	XP_041091848.1
Polyodon spathula PGALe	XP_041092101.1
Polyodon spathula PGALf	XP_041084632.1
Polyodon spathula PGALg	XP_041092046.1
Anguilla anguilla PGALa	XP_035238658.1
Anguilla anguilla PGALb	XP_035237009.1
Scleropages formosus PGALa	XP_029114792.1
Scleropages formosus PGALb	XP_018606330.1
Scleropages formosus PGALc	XP_018606240.1
Scleropages formosus PGALd	XP_018617085.1
Gadus morhua PGALa	XP_030237681.1
Gadus morhua PGALb	XP_030192894.1
Gadus morhua PGALc	XP_030195301.1
Gadus morhua PGALd	XP_030237581.1
Gadus morhua PGALe	XP_030230695.1
Oreochromis niloticus PGALa	XP_003444873.1
Oreochromis niloticus PGALb	XP_025765624.1
Gasterosteus aculeatus PGALa	XP_040032590.1
Gasterosteus aculeatus PGALb	XP_040050298.1
Gasterosteus aculeatus PGALc	XP_040059424.1
Clupea harengus PGALa	XP_031421353.2
Clupea harengus PGALb	XP_031414371.1
Clupea harengus PGALc	XP_012675128.2
Clupea harengus PGALd	XP_042565836.1
Electrophorus electricus PGAL_X1	XP_026851318.1
Electrophorus electricus PGAL	XP_035389448.1
Ictalurus punctatus PGAa	XP_017323545.1
Ictalurus punctatus PGAb	XP_017338171.1
Ictalurus punctatus PGAc	XP_017338196.1
Ictalurus punctatus PGAd	NP_001187944.1
Ictalurus punctatus PGAe	XP_017323541.1
Esox lucius PGALa	XP_010900709.4
Esox lucius PGALb	XP_028980024.2
Esox lucius PGALc	XP_034151761.1
Esox lucius PGALd	XP_028980026.2

Esox lucius PGALe	XP_034151845.1
Esox lucius PGALf	XP_034151760.1
Esox lucius PGALg	XP_034151759.1
Esox lucius PGALh	XP_010890229.2
Esox lucius PGALi	XP_010891481.3
Esox lucius PGA_X1	XP_034151843.1
Salmo salar PGALa	XP_045577449.1
Salmo salar PGALb	XP_014002860.1
Salmo salar PGA_X1a	XP_045548866.1
Salmo salar PGALc	XP_014013617.2
Salmo salar PGALd	XP_014013619.1
Salmo salar PGA_X1b	XP_014022342.1
Myripristis murdjan PGALa	XP_029933656.1
Myripristis murdjan PGALb	XP_029911515.1
Myripristis murdjan PGALc	XP_029907485.1
Carcharodon carcharias PGAL1	XP_041051482.1
Carcharodon carcharias PGAL2	XP_041050881.1
Carcharodon carcharias PGAL3	XP_041050880.1
Carcharodon carcharias PGAL4	XP_041050883.1
Carcharodon carcharias PGAL5	XP_041050882.1
Carcharodon carcharias PGAL6	XP_041051017.1
Carcharodon carcharias PGAL7	XP_041051243.1
Carcharodon carcharias PGAL8	XP_041051574.1
Carcharodon carcharias PGAL9	XP_041051678.1
Carcharodon carcharias PGAL10	XP_041050879.1
Carcharodon carcharias PGAL11	XP_041051246.1
Carcharodon carcharias PGAL12	XP_041050651.1
Carcharodon carcharias PGAL13	XP_041051116.1
Carcharodon carcharias PGAL14	XP_041051244.1
Carcharodon carcharias PGAL15	XP_041050884.1
Carcharodon carcharias PGAL16	XP_041051249.1
Carcharodon carcharias PGAL17	XP_041051248.1
Carcharodon carcharias PGAL18	XP_041051247.1
Amblyraja radiata PGAL1	XP_032898812.1
Amblyraja radiata PGAL2	XP_032872962.1
Amblyraja radiata PGA4L	XP_032898814.1
Amblyraja radiata PGAL3	XP_032898086.1
Amblyraja radiata PGAL4	XP_032898811.1
Amblyraja radiata PGAL5	XP_032898813.1
Scyliorhinus canicula PGAL1	XP_038675276.1
Scyliorhinus canicula PGAL2	XP_038675278.1
Scyliorhinus canicula PGAL3	XP_038675946.1
Scyliorhinus canicula PGA2.3L	XP_038675945.1
Scyliorhinus canicula PGAL4	XP_038675279.1
Scyliorhinus canicula PGA3L	XP_038675275.1
Scyliorhinus canicula PGAL5	XP_038675205.1

Scyliorhinus canicula PGAL6	XP_038675282.1
Scyliorhinus canicula PGAL7	XP_038675281.1
Scyliorhinus canicula PGAL8	XP_038675283.1
Scyliorhinus canicula PGAL9	XP_038675206.1
Scyliorhinus canicula PGAL10	XP_038675280.1
Rhincodon typus PGAL1	XP_020375383.1
Rhincodon typus PGAL2	XP_048468044.1
Stegostoma fasciatum PGAL1	XP_048409337.1
Stegostoma fasciatum PGAL2	XP_048409058.1
Stegostoma fasciatum PGAL3	XP_048409293.1
Stegostoma fasciatum PGAL4	XP_048409059.1

Table 6 - Accession numbers of PGC sequences.

Species	***Accession Number***
Homo sapiens PGC	AAH73740.1
Pan troglodytes PGC	XP_016810948.1
Nomascus leucogenys PGC	XP_030652580.1
Mus musculus PGC	NP_080249.2
Gallus gallus PGC	NP_990208.1
Meleagris gallopavo PGCLa	XP_003213026.1
Meleagris gallopavo PGCLb	XP_010722700.1
Taeniopygia guttata PGC	XP_030147988.3
Corvus cornix PGC	XP_039420997.1
Chrysemys bellii PGCLa	XP_005308354.1
Chrysemys bellii PGCLb	XP_042705002.1
Anolis carolinensis PGCL	XP_003220379.1
Anolis carolinensis PGC	XP_003220377.1
Bufo bufo PGCLa	XP_040278134.1
Bufo bufo PGCLb	XP_040277812.1
Bufo bufo PGCLc	XP_040277810.1
Bufo bufo PGCLd	XP_040277813.1
Bufo bufo PGCLe	XP_040277808.1
Bufo bufo PGCLf	XP_040277811.1
Bufo bufo PGCLg	XP_040277809.1
Bufo bufo PGCLh	XP_040278135.1
Xenopus laevis PGC_X1a	XP_041437935.1
Xenopus laevis PGC_X1b	XP_018096610.1
Xenopus laevis PGCa	XP_018104214.1
Xenopus laevis PGCb	XP_018104213.1
Geotrypetes seraphini PGC	XP_033774487.1
Geotrypetes seraphini PGCL	XP_033774785.1

Polypterus senegalus PGCLa	XP_039620720.1
Polypterus senegalus PGCLb	XP_039620428.1
Polypterus senegalus PGCLc	XP_039620429.1
Lepisosteus oculatus PGC	XP_015198218.1
Lepisosteus oculatus PGCL	XP_006636729.1
Polyodon spathula PGCL_X1a	XP_041083283.1
Polyodon spathula PGCLa	XP_041084521.1
Polyodon spathula PGCLb	XP_041083922.1
Polyodon spathula PGCLc	XP_041083920.1
Polyodon spathula PGCL_X1b	XP_041083914.1
Polyodon spathula PGCLd	XP_041120540.1
Polyodon spathula PGCL_X1c	XP_041117467.1
Anguilla anguilla PGCL	XP_035257999.1
Scleropages formosus PGCLa	XP_018611523.1
Scleropages formosus PGCLb	XP_018611520.1
Scleropages formosus PGCLc	XP_029104614.1
Scleropages formosus PGCLd	XP_018611521.1
Scleropages formosus PGCLe	XP_018591035.2
Gadus morhua PGCL	XP_030222924.1
Gasterosteus aculeatus PGCL	XP_040018008.1
Clupea harengus PGCLa	XP_031420519.1
Clupea harengus PGCLb	XP_031434048.1
Electrophorus electricus PGC	XP_026870930.2
Electrophorus electricus PGCL_X1	XP_035390644.1
Ictalurus punctatus PGCa	XP_017321906.2
Ictalurus punctatus PGCLa	XP_047011144.1
Ictalurus punctatus PGCLb	XP_047011145.1
Ictalurus punctatus PGCb	XP_017321907.2
Ictalurus punctatus PGCc	XP_017321909.1
Ictalurus punctatus PGCd	XP_017321905.1
Ictalurus punctatus PGCf	XP_017325116.1
Ictalurus punctatus PGCg	XP_017325115.1
Esox lucius PGCa	XP_010888694.1
Esox lucius PGCb	XP_010888692.2
Salmo salar PGCL	XP_013980433.2
Salmo salar PGC	XP_013980432.2
Myripristis murdjan PGCL	XP_029909168.1
Scyliorhinus canicula PGCLa	XP_038676616.1
Scyliorhinus canicula PGCLb	XP_038675277.1
Scyliorhinus canicula PGCLc	XP_038675944.1
Carcharodon carcharias PGCL	XP_041052341.1
Amblyraja radiata PGC	XP_032898815.1
Rhincodon typus PGCL	XP_020368736.1
Stegostoma fasciatum PGCLa	XP_048409491.1
Stegostoma fasciatum PGCLb	XP_048408842.1
Chiloscyllium plagiosum PGCLa	XP_043572465.1

Chiloscyllium plagiosum PGCLb	XP_043572467.1

8.2 Synteny analysis of stomach-related genes

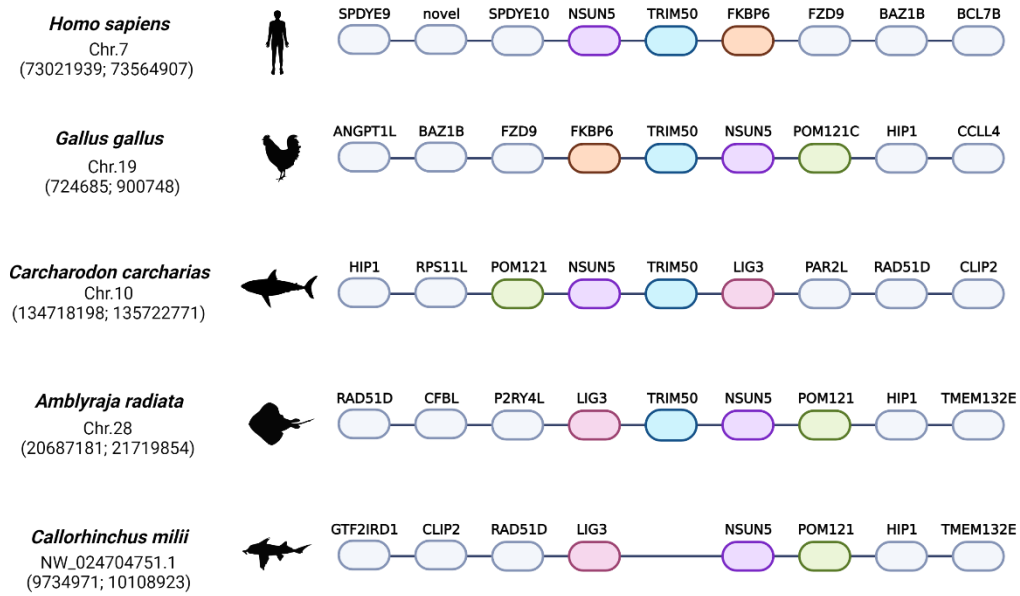

*Figure 31 - **Genomic loci of the TRIM50 gene in vertebrate species.** Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

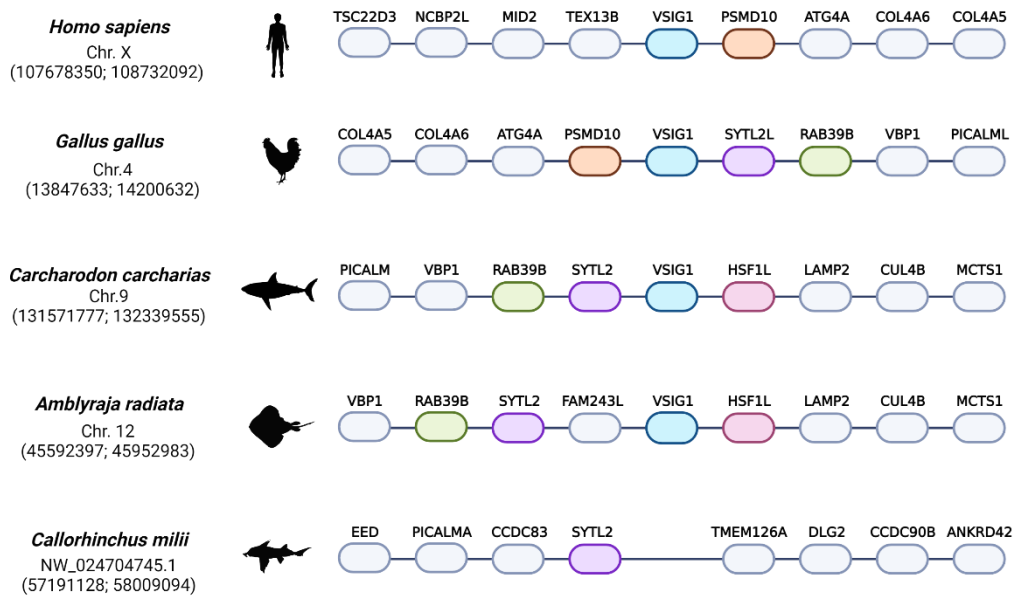

*Figure 32 - **Genomic loci of the VSIG1 gene in vertebrate species.** Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

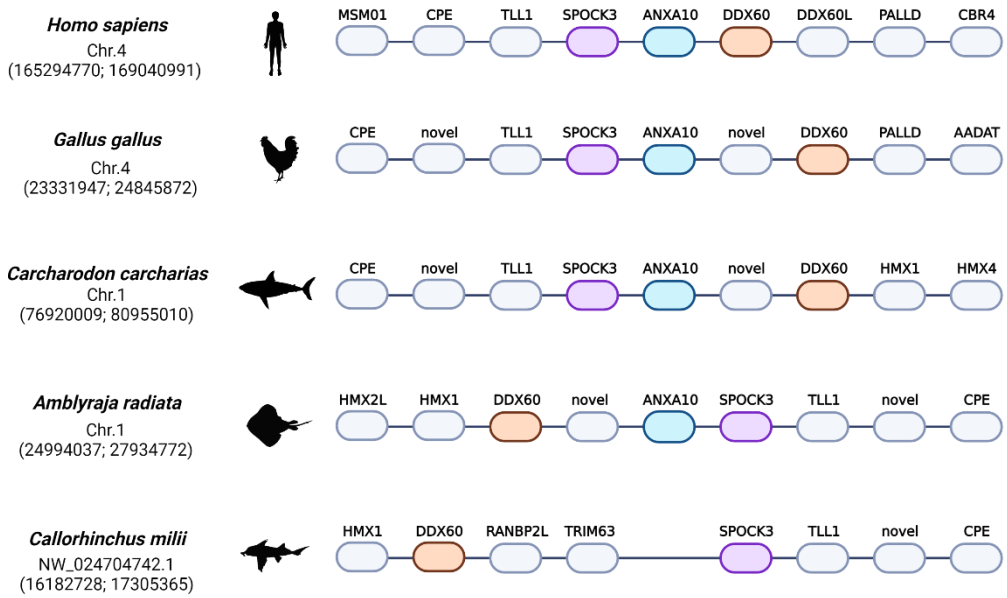

*Figure 33 - **Genomic loci of the ANXA10 gene in vertebrate species.** Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

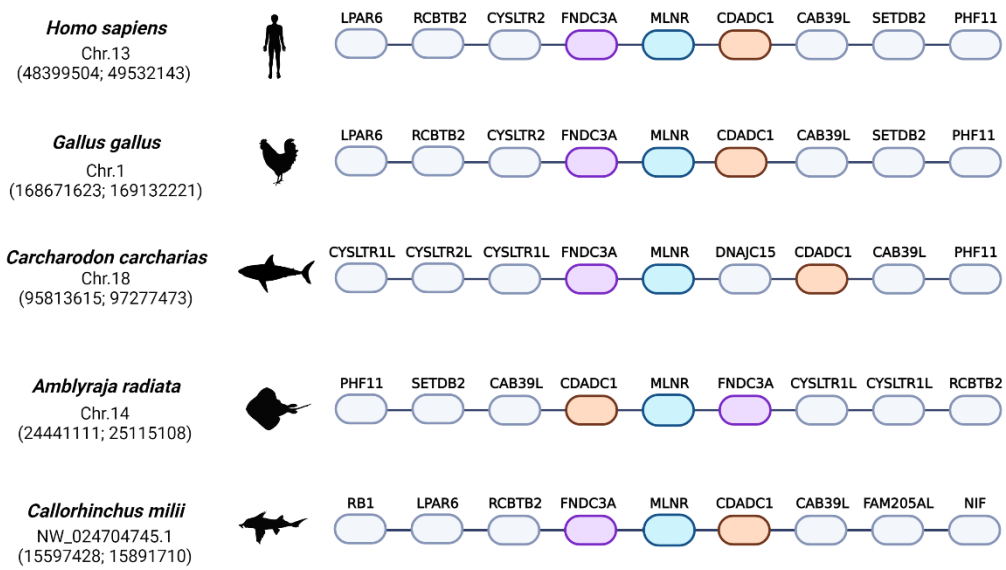

*Figure 34 - **Genomic loci of the MLNR gene in vertebrate species.** Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias);, thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

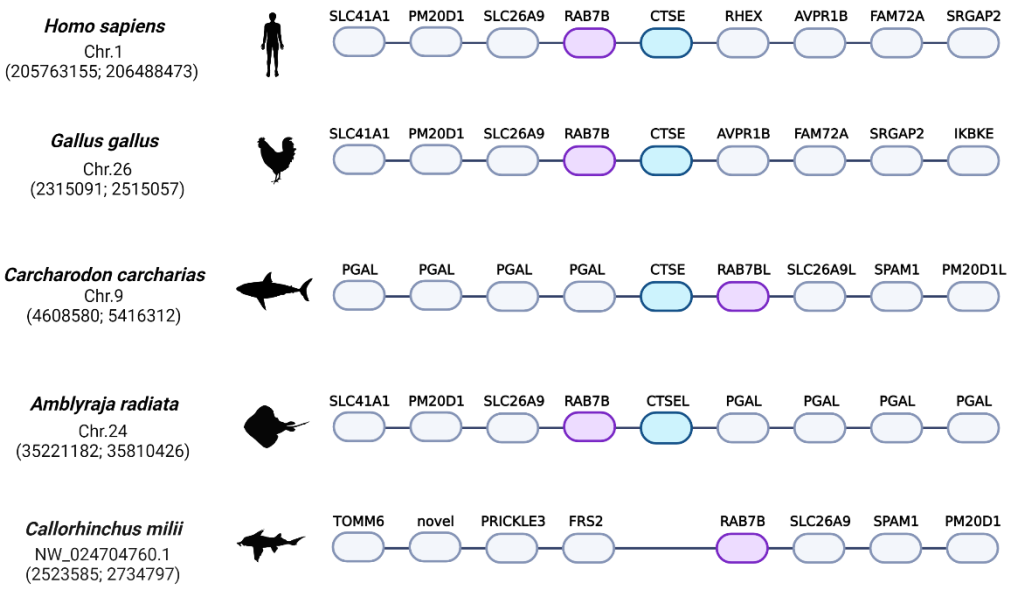

*Figure 35 - **Genomic loci of the CTSE gene in vertebrate species.** Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

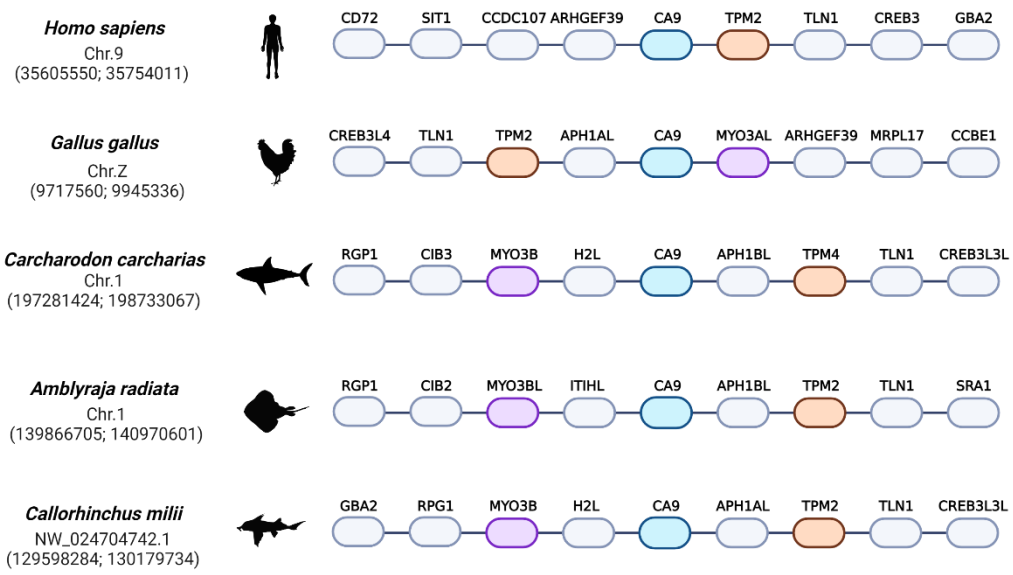

*Figure 36 - **Genomic loci of the CA9 gene in vertebrate species.** Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

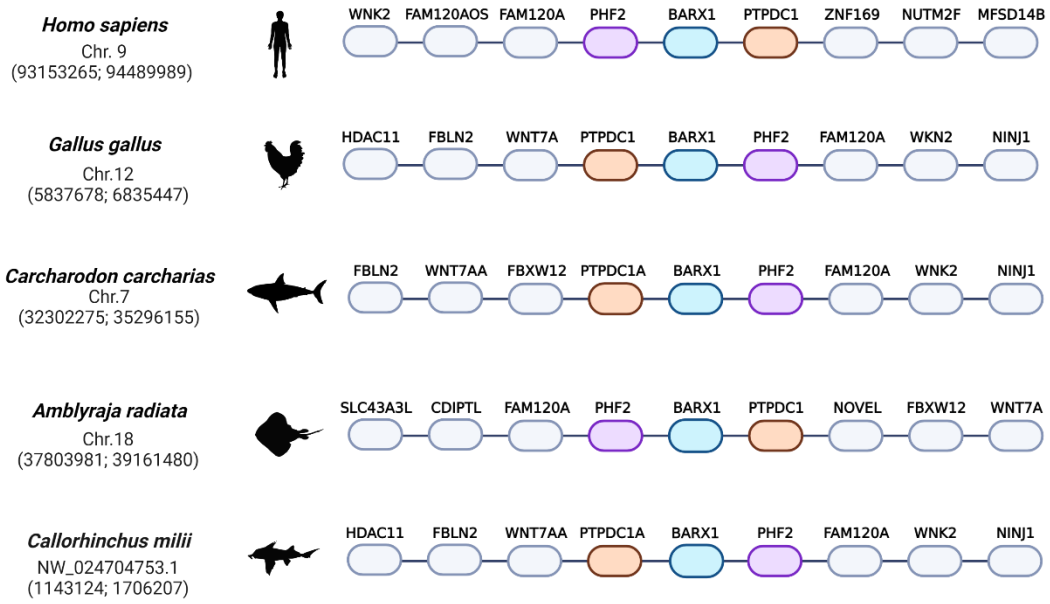

*Figure 38 - **Genomic loci of the BARX1 gene in vertebrate species.** Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

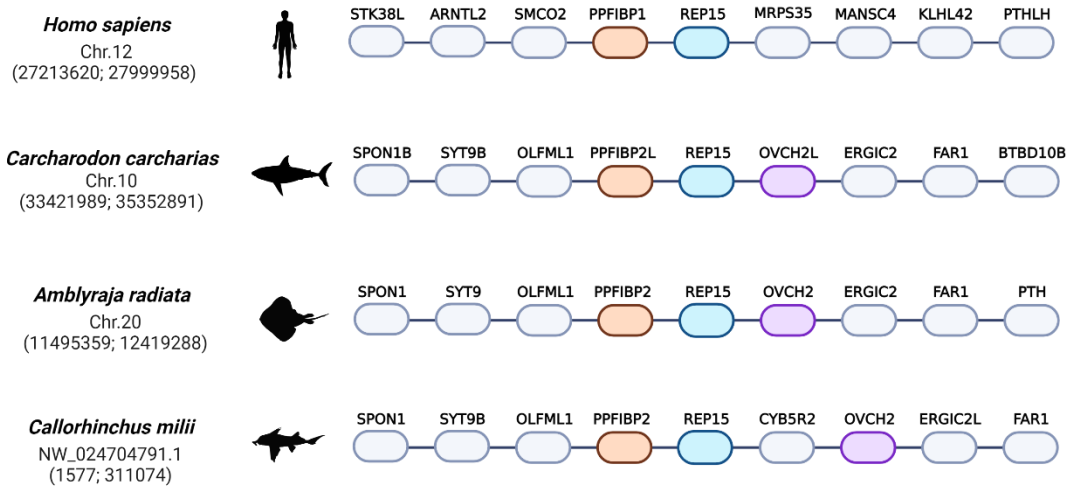

*Figure 37 - **Genomic loci of the REP15 gene in vertebrate species.** Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

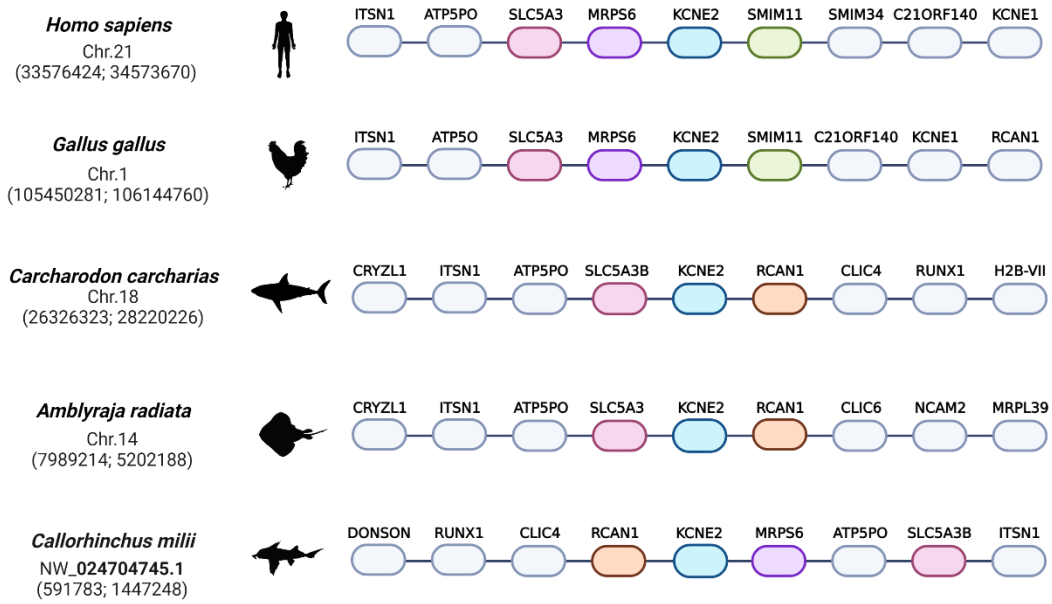

*Figure 39 - **Genomic loci of the KCNE2 gene in vertebrate species.** Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

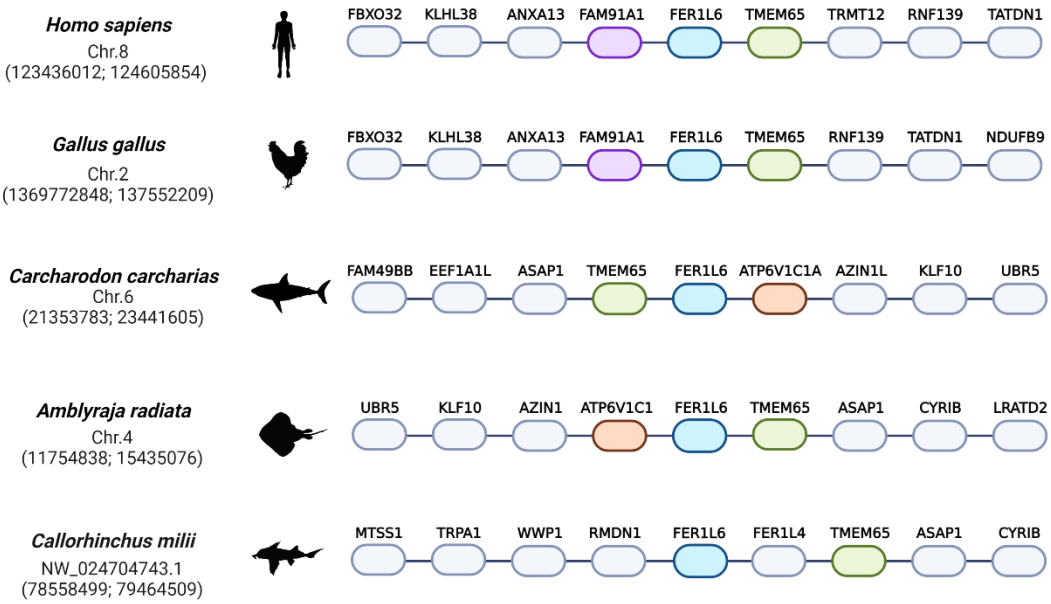

*Figure 40 - **Genomic loci of the FER1L6 gene in vertebrate species.** Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

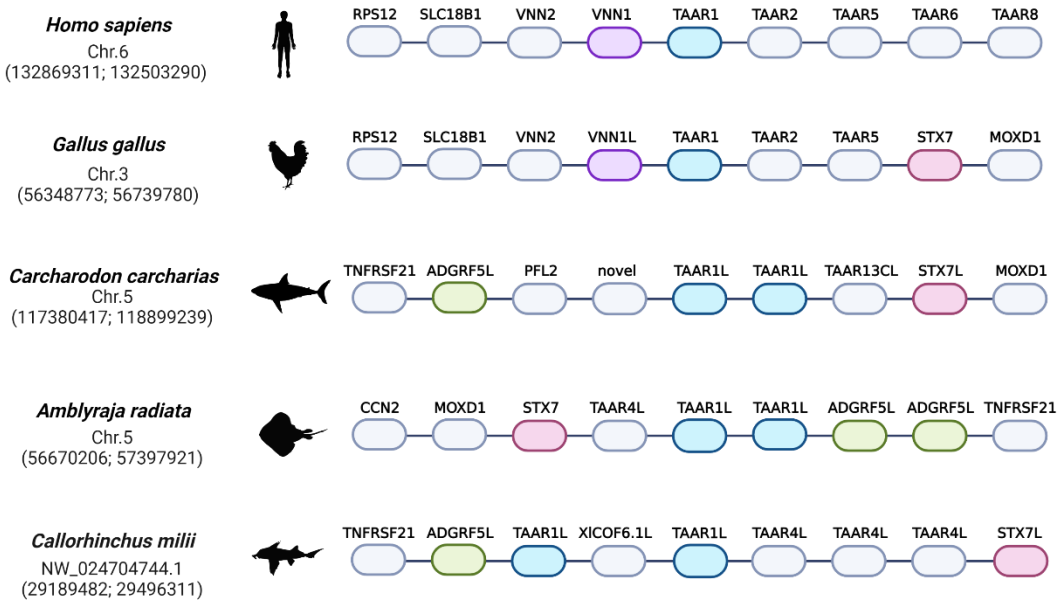

Figure 41 - ***Genomic loci of the TAAR1 gene in vertebrate species.*** *Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

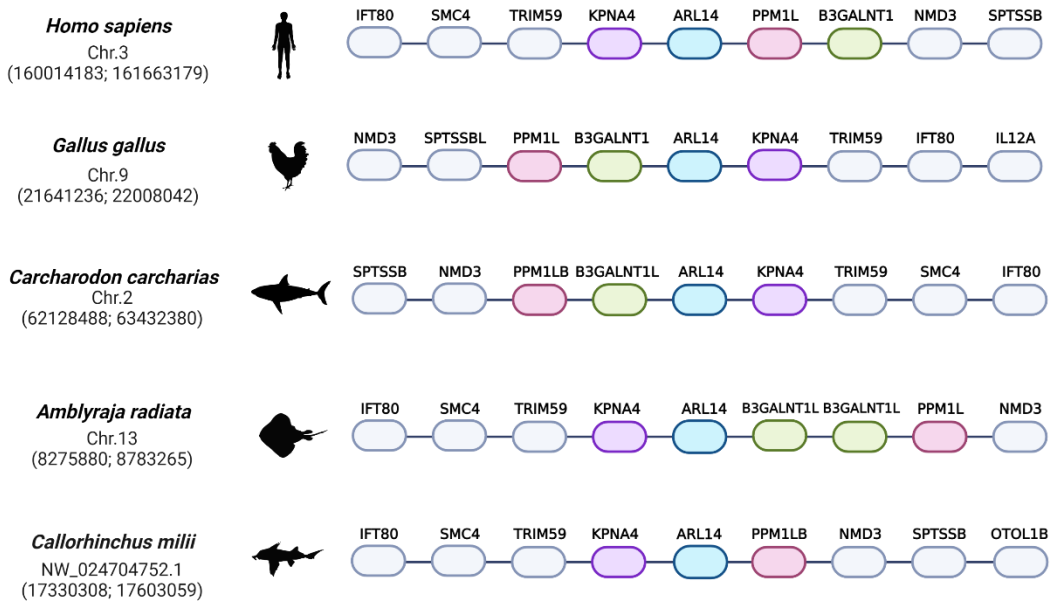

Figure 42 - ***Genomic loci of the ARL14 gene in vertebrate species.*** *Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

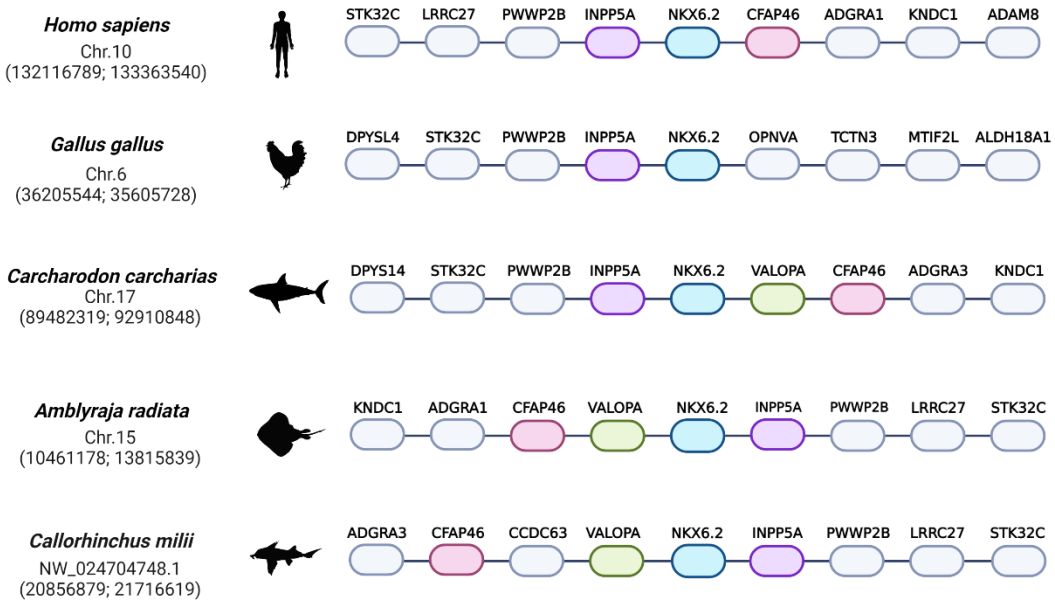

Figure 43 - ***Genomic loci of the NKX6.2 gene in vertebrate species.*** *Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

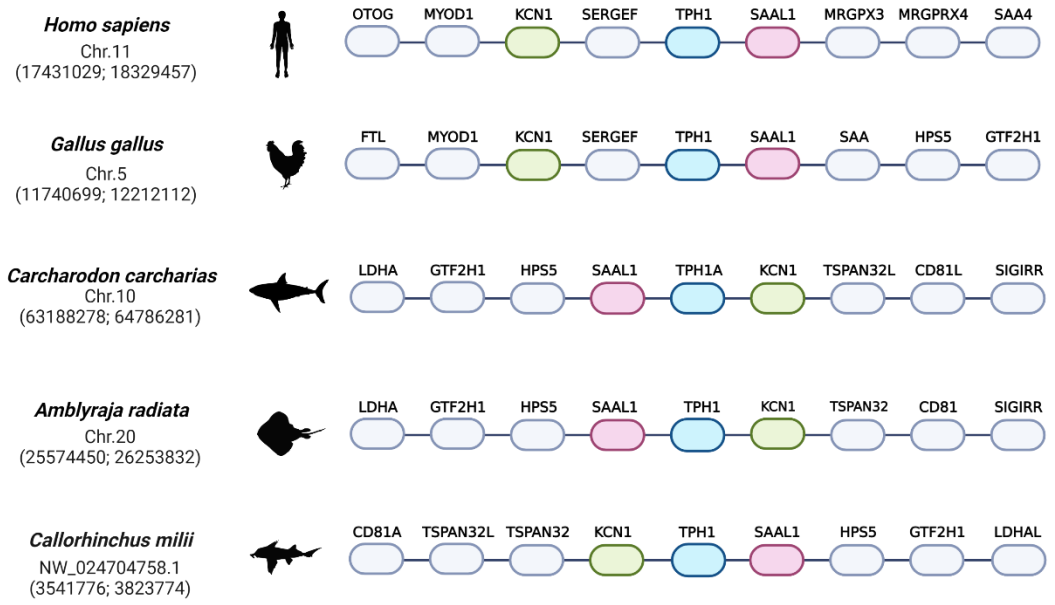

Figure 44 – ***Genomic loci of the TPH1 gene in vertebrate species.*** *Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

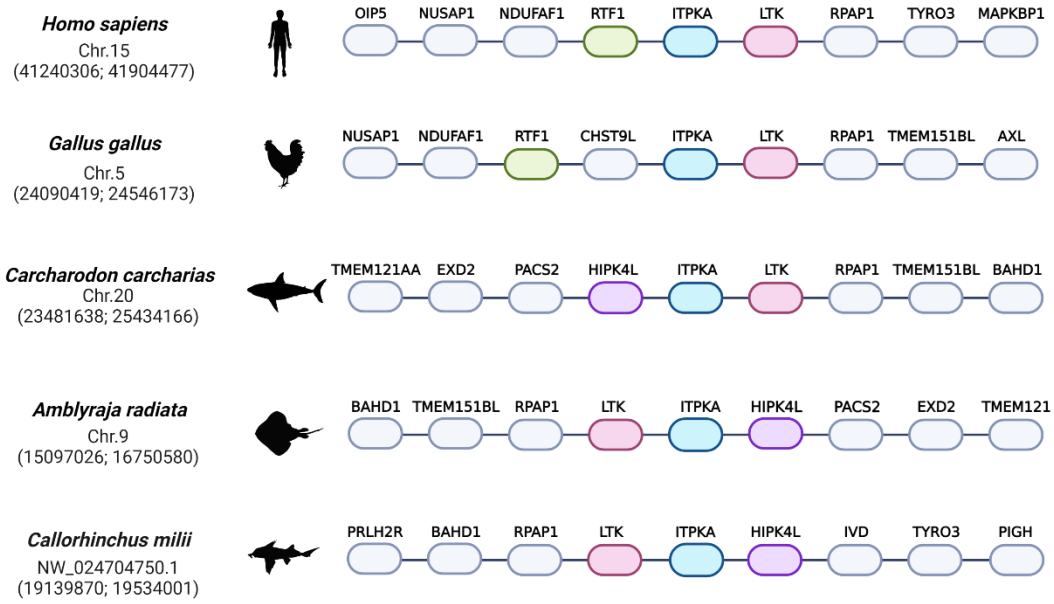

*Figure 46 - **Genomic loci of the ITPKA gene in vertebrate species.** Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

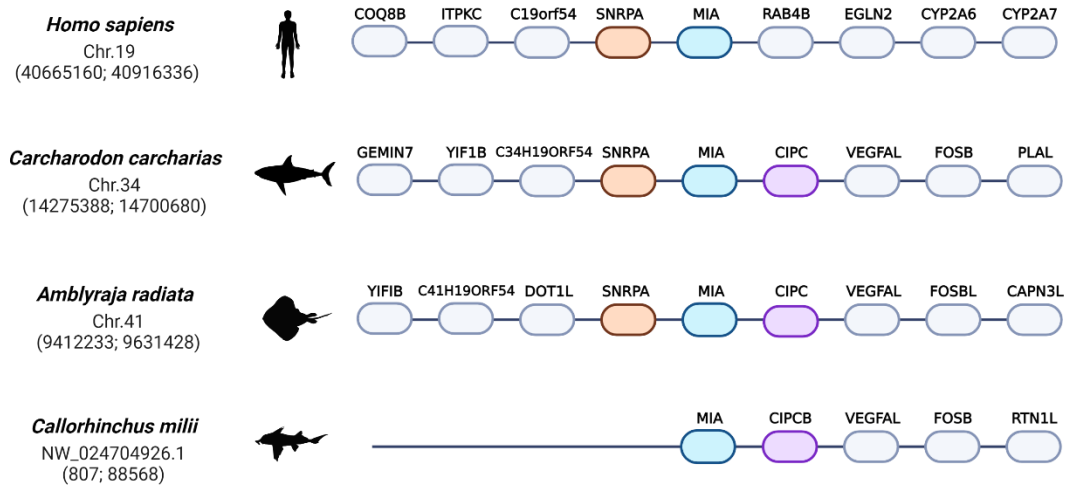

*Figure 45 - **Genomic loci of the MIA gene in vertebrate species.** Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

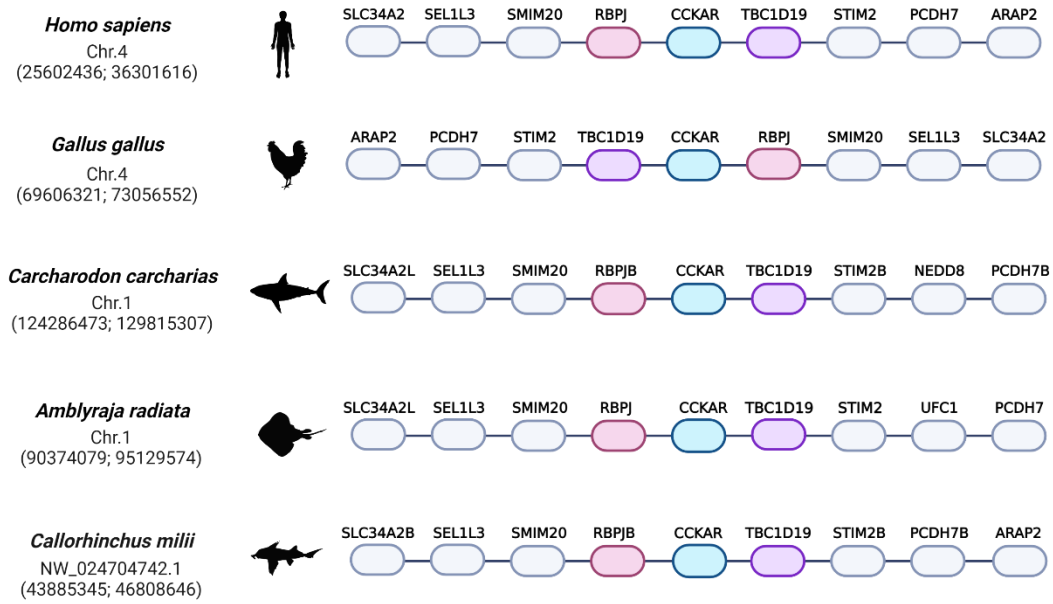

Figure 47 - ***Genomic loci of the CCKAR gene in vertebrate species.*** *Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

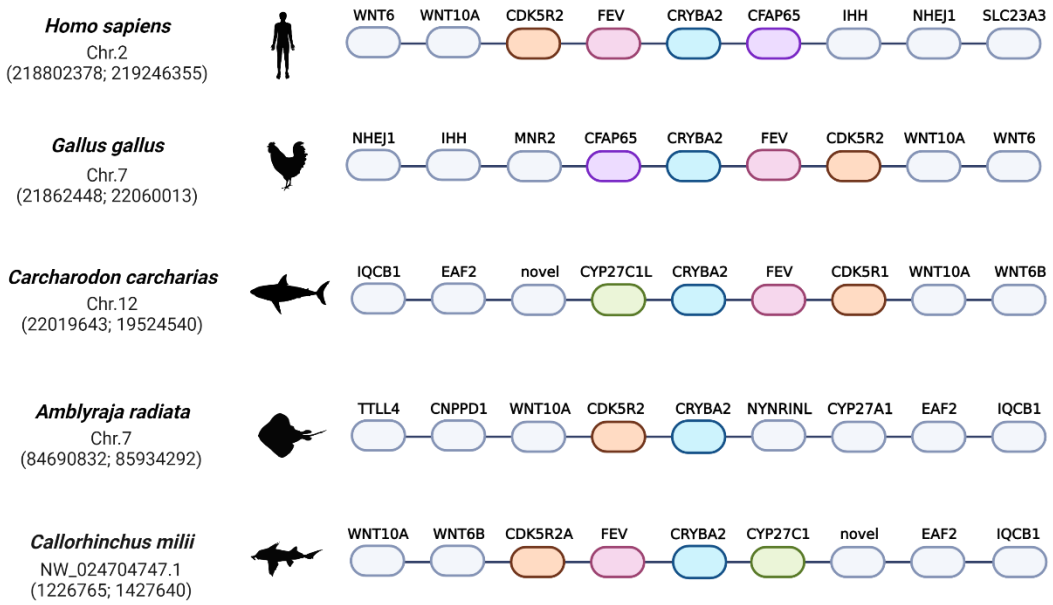

Figure 48 - ***Genomic loci of the CRYBA2 gene in vertebrate species.*** *Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

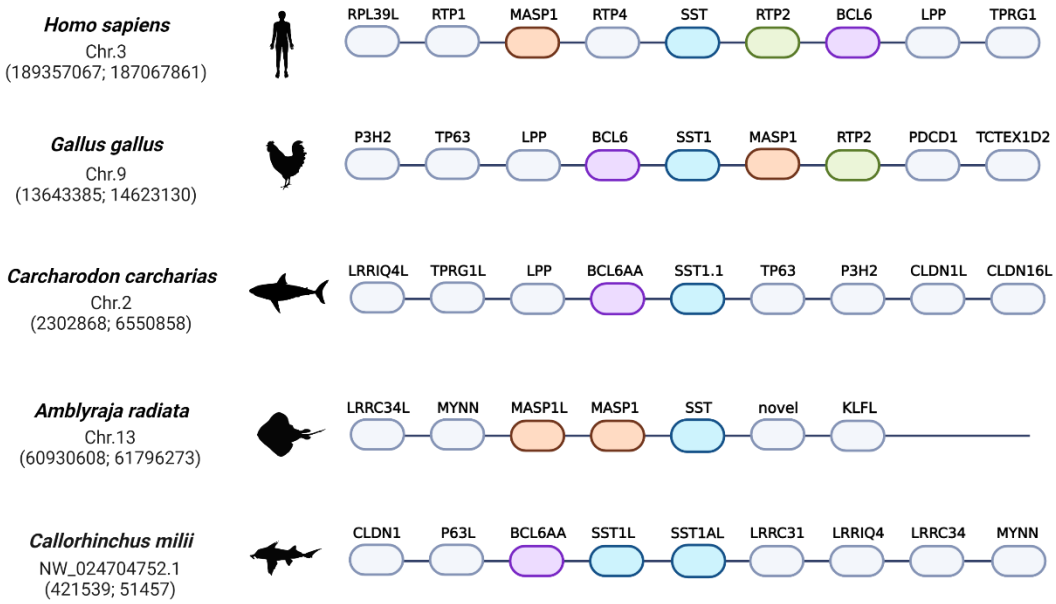

*Figure 49 - **Genomic loci of the SST gene in vertebrate species.** Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

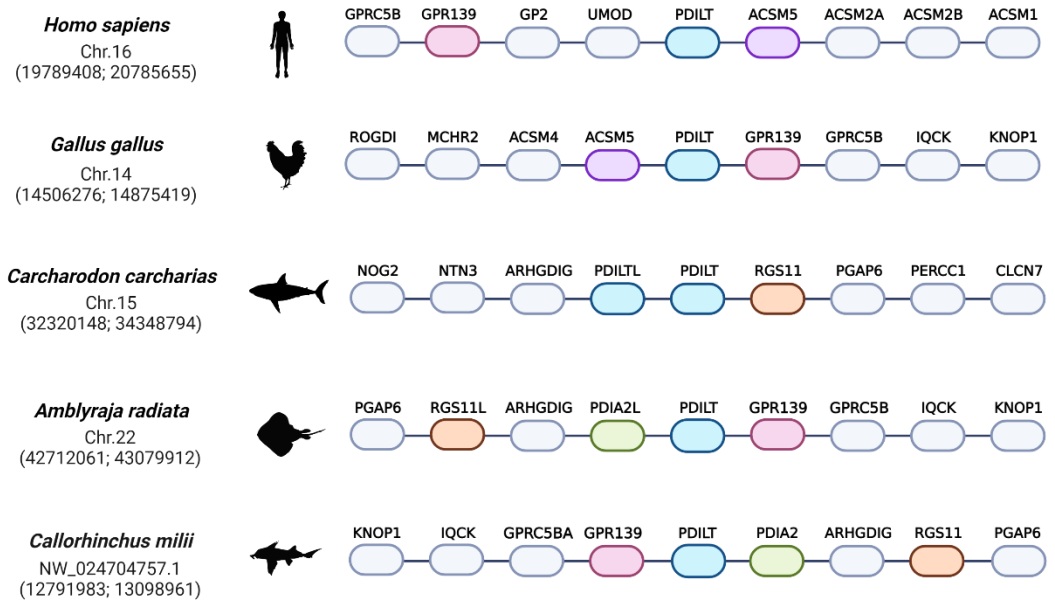

*Figure 50 - **Genomic loci of the PDILT gene in vertebrate species.** Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

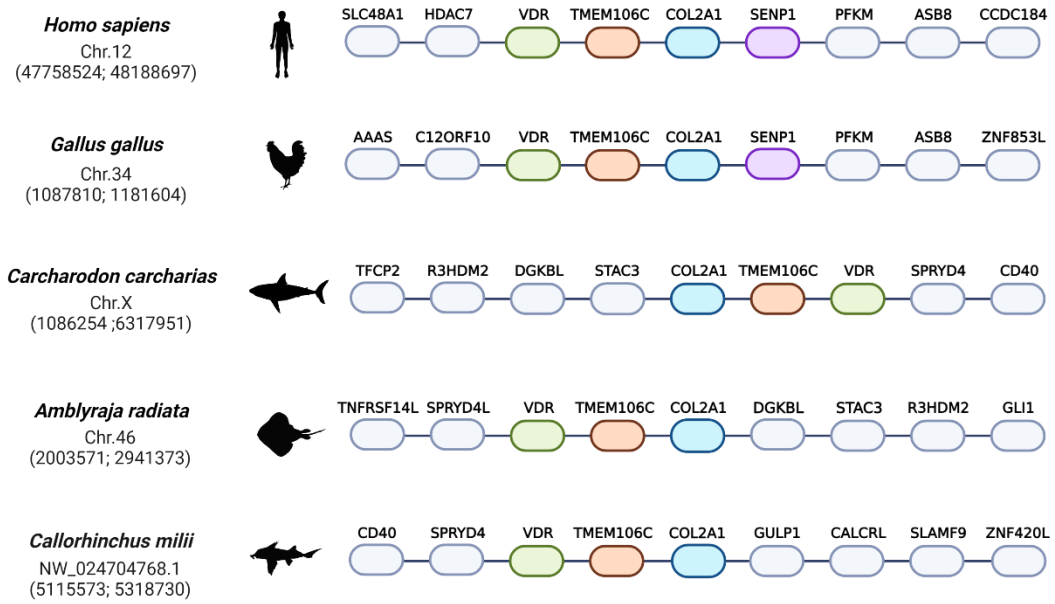

*Figure 51 - **Genomic loci of the COL2A1 gene in vertebrate species.** Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

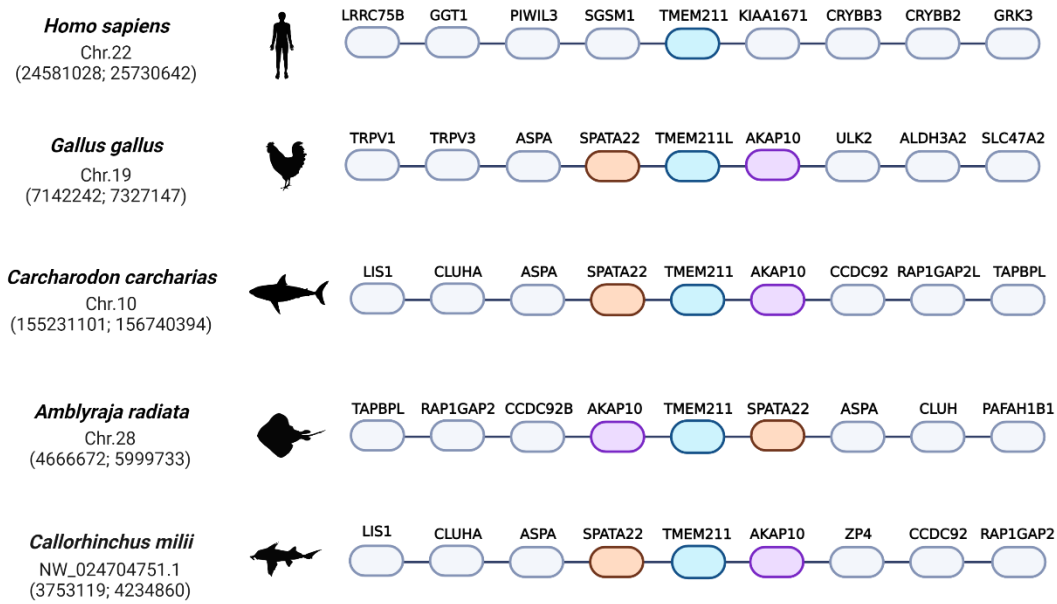

*Figure 52 - **Genomic loci of the TMEM211 gene in vertebrate species.** Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

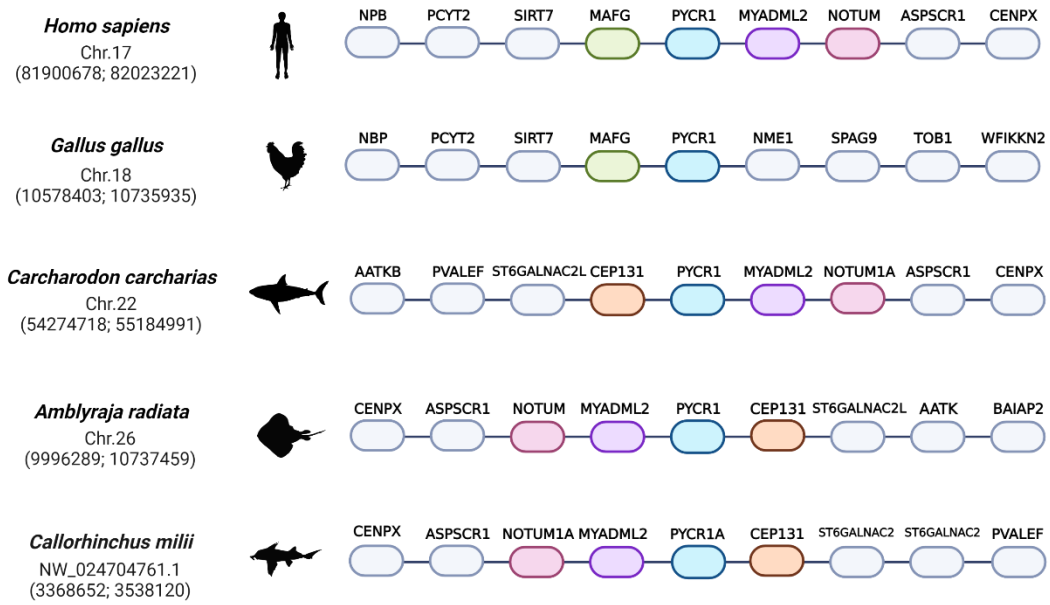

*Figure 53 - **Genomic loci of the PYCR1 gene in vertebrate species.** Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

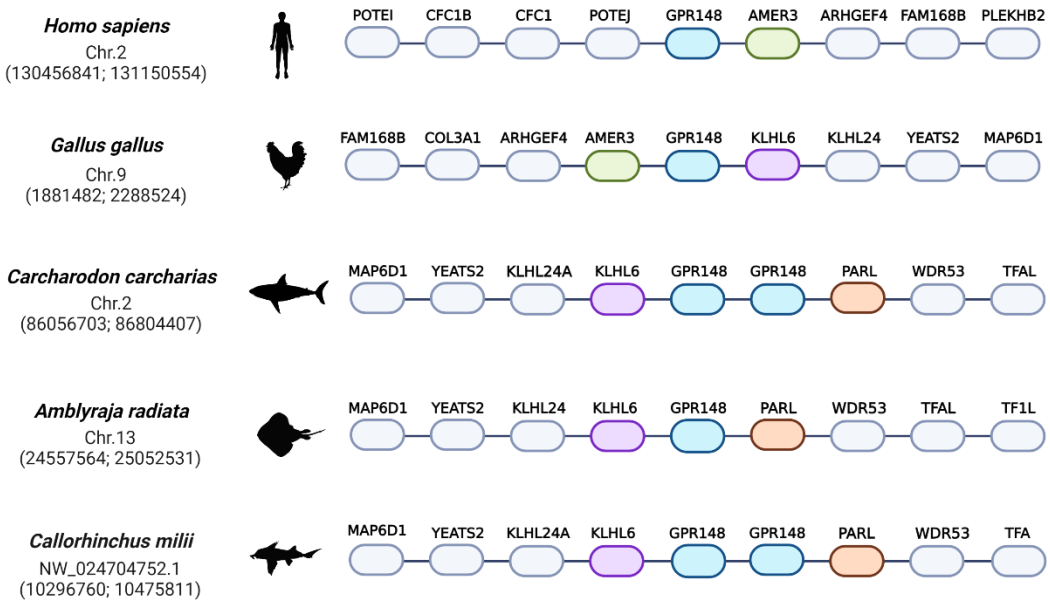

*Figure 54 - **Genomic loci of the GPR148 gene in vertebrate species.** Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

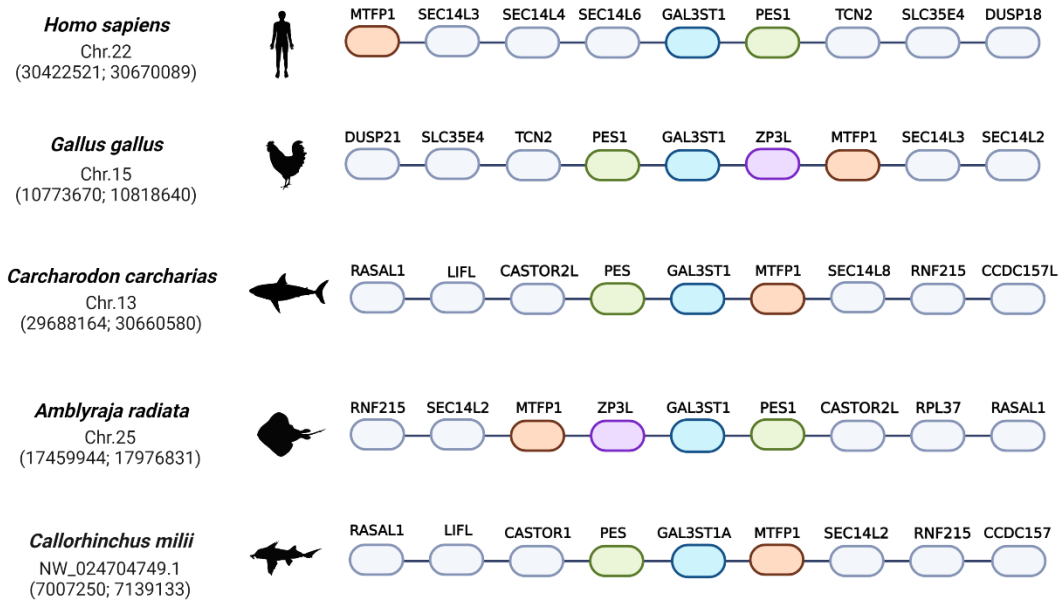

*Figure 55 - **Genomic loci of the GAL3ST1 gene in vertebrate species.** Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

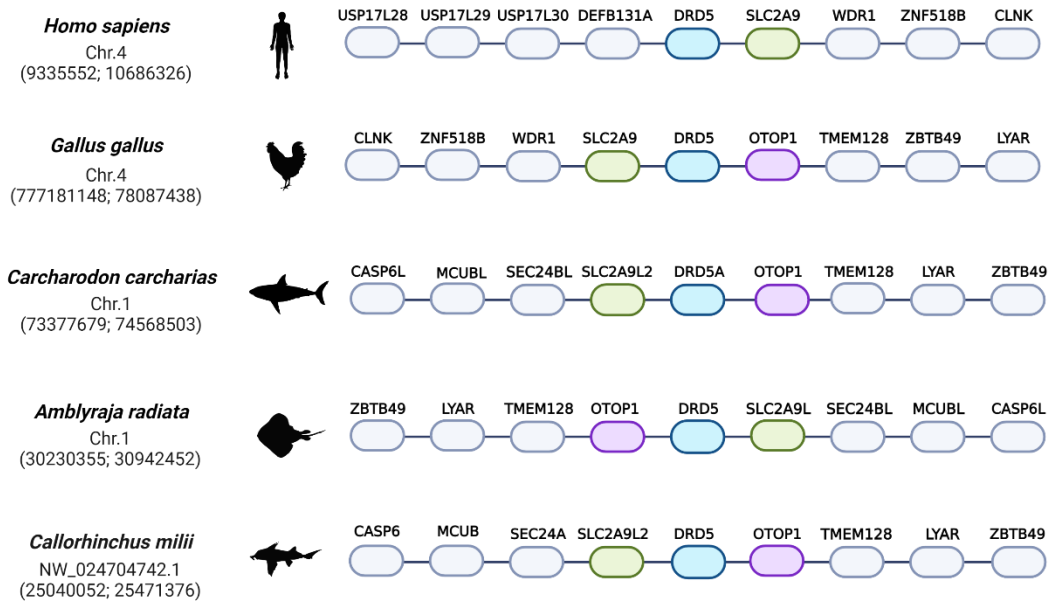

*Figure 56 - **Genomic loci of the DRD5 gene in vertebrate species.** Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

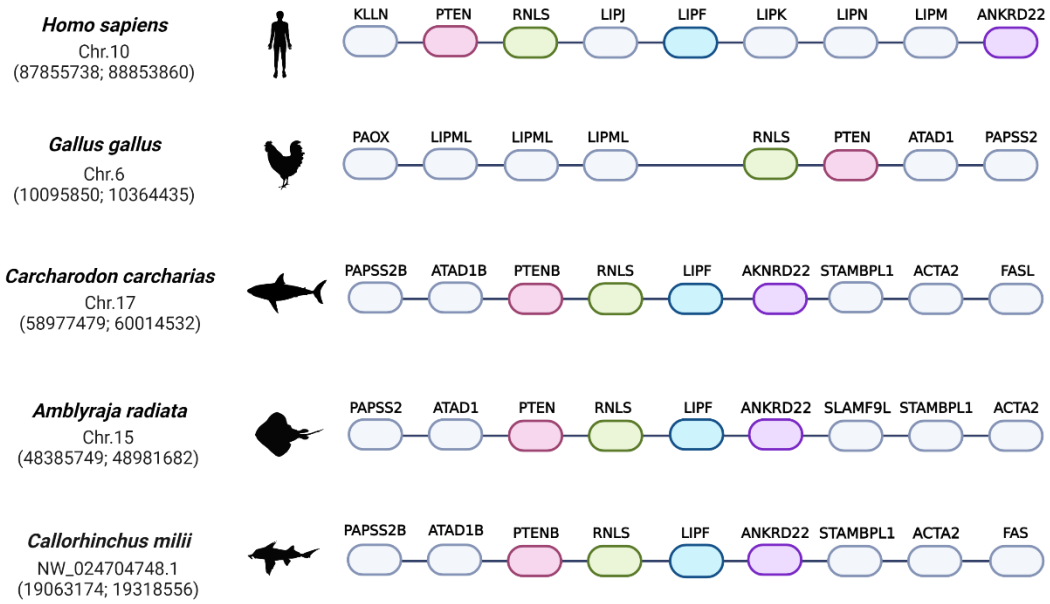

*Figure 57 - **Genomic loci of the LIPF gene in vertebrate species.** Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

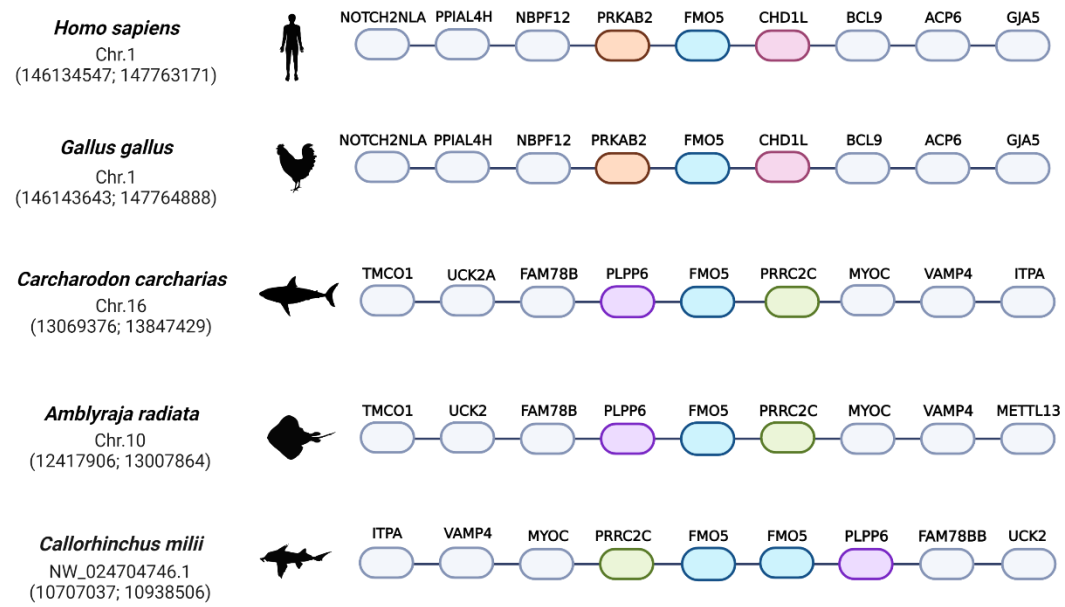

*Figure 58 - **Genomic loci of the FMO5 gene in vertebrate species.** Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

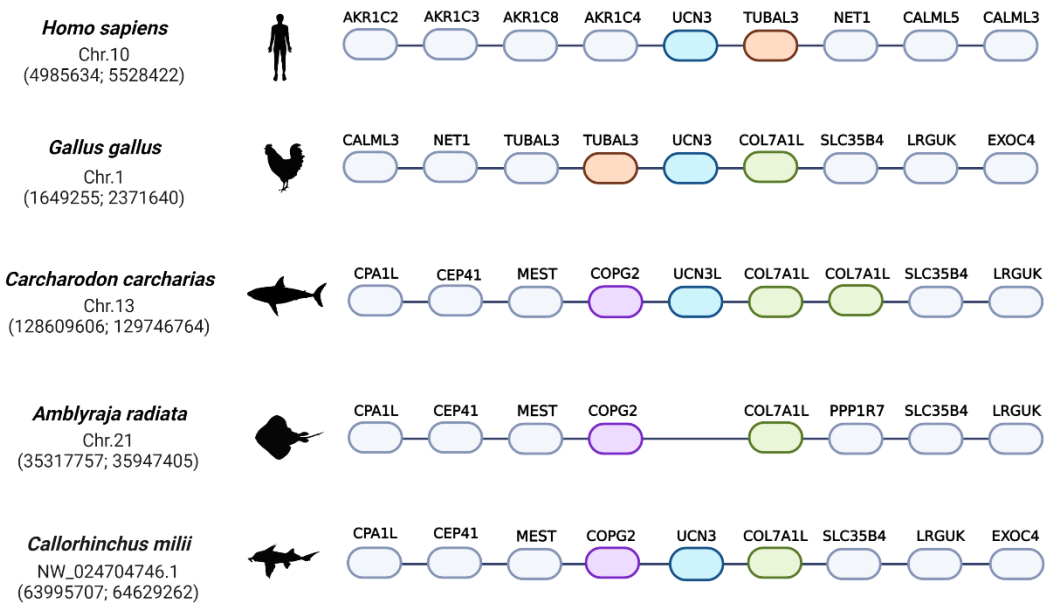

*Figure 59 - **Genomic loci of the UCN3 gene in vertebrate species.** Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

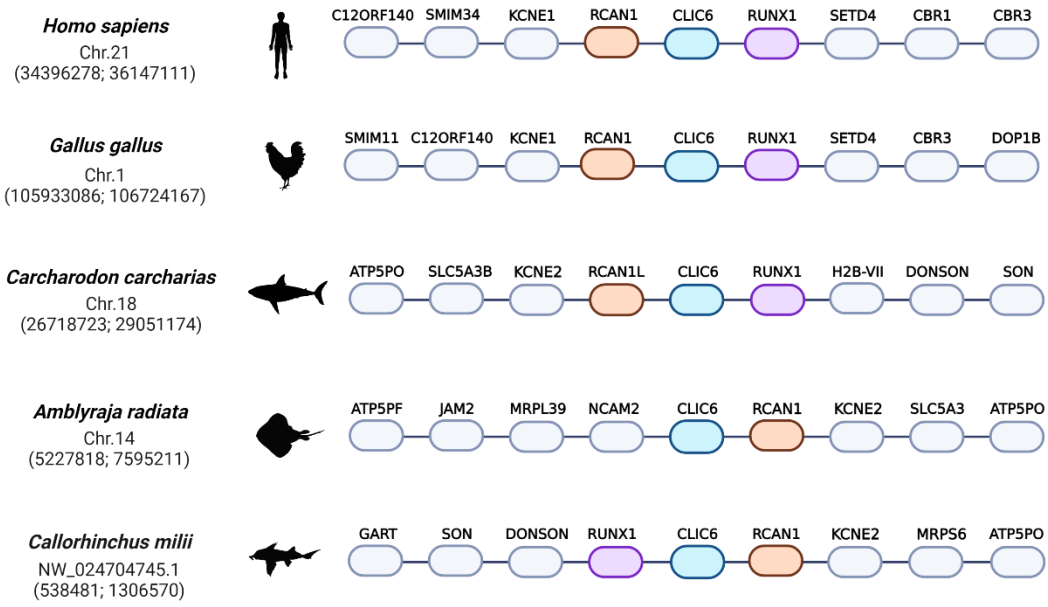

*Figure 60 - **Genomic loci of the CLIC6 gene in vertebrate species.** Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

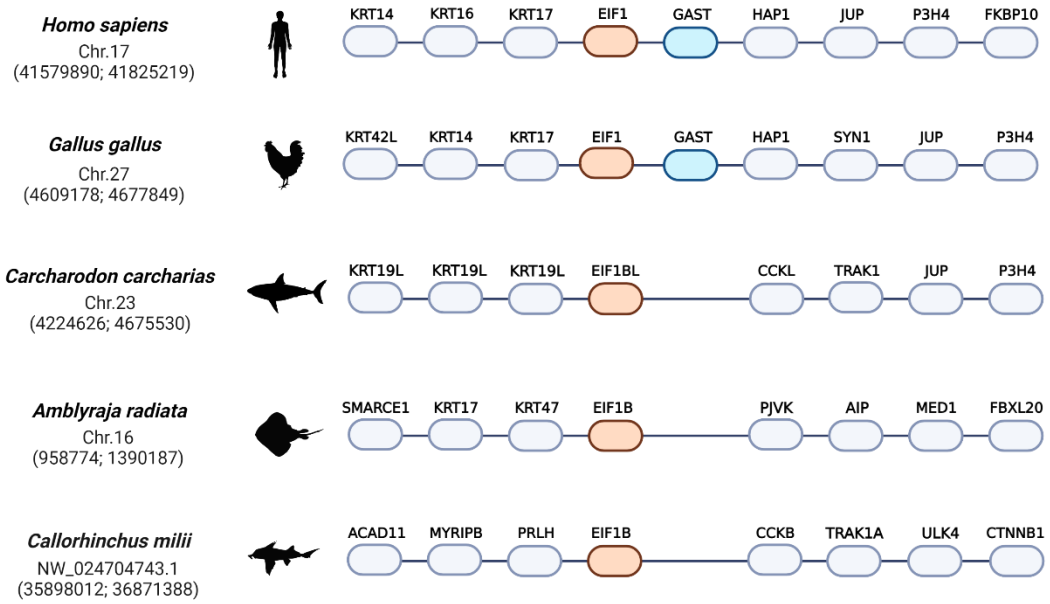

Figure 61 ***Genomic loci of the GAST gene in vertebrate species.*** *Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

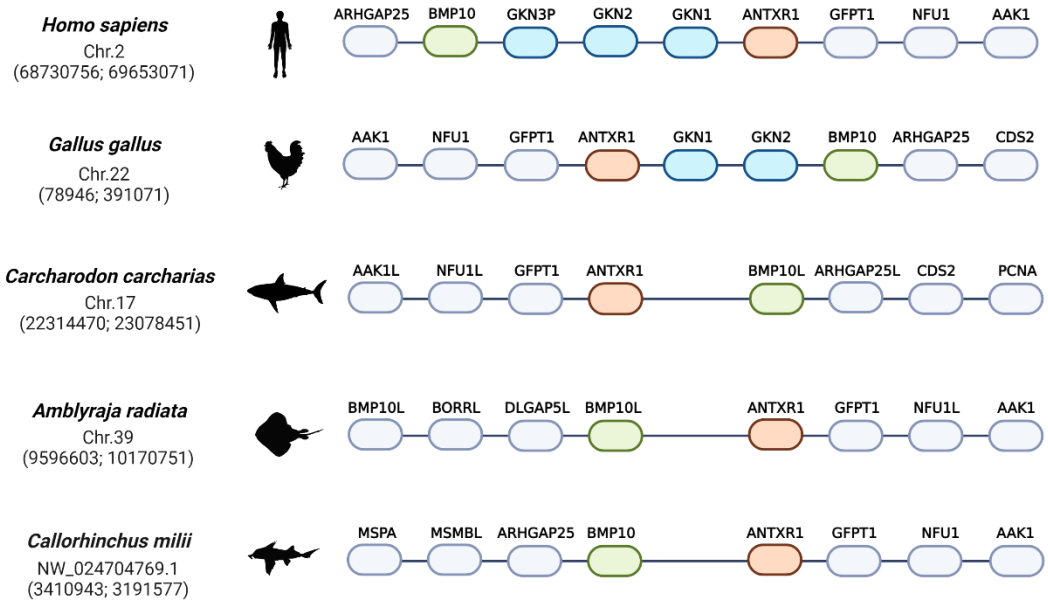

Figure 62 - ***Genomic loci of the GRK1/GRK2/GRK3 gene in vertebrate species.*** *Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

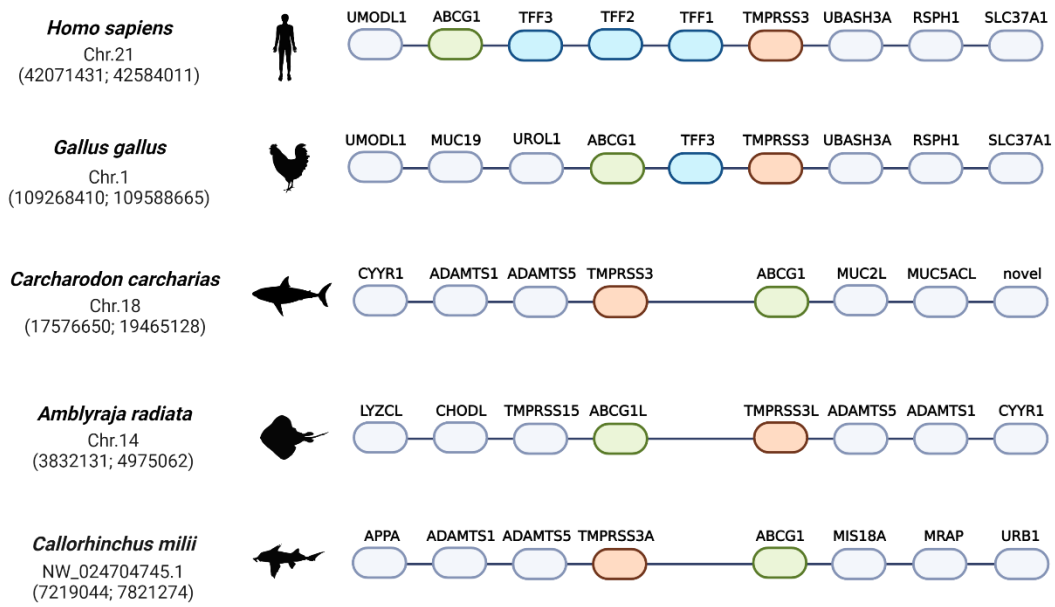

*Figure 63 - **Genomic loci of the TFF1/TFF2 gene in vertebrate species.** Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

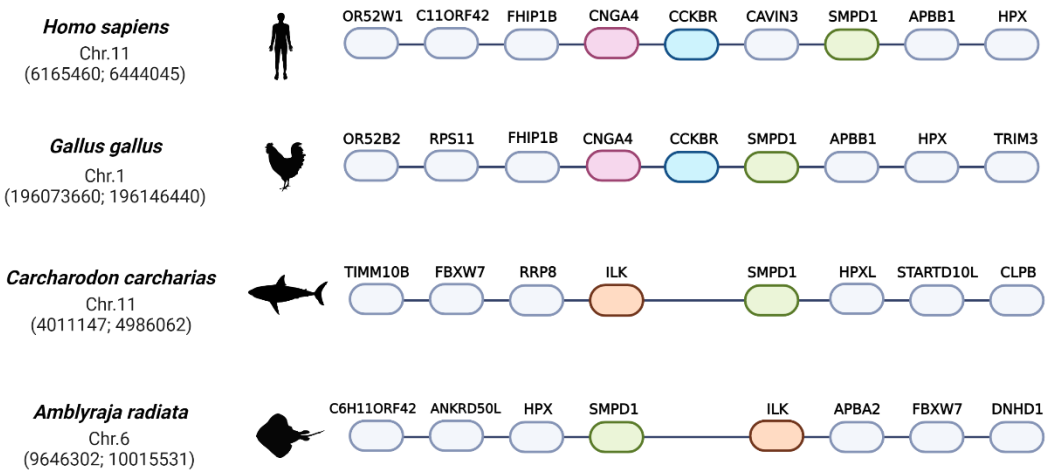

*Figure 64 - **Genomic loci of the CCKBR gene in vertebrate species.** Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata)*

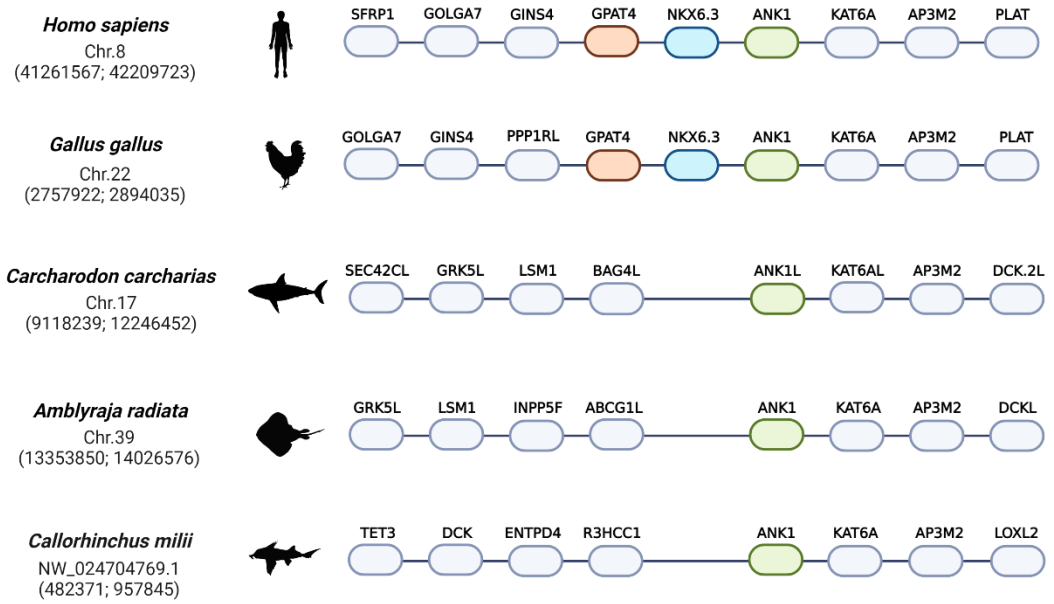

*Figure 65 - **Genomic loci of the NKX6.3 gene in vertebrate species.** Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

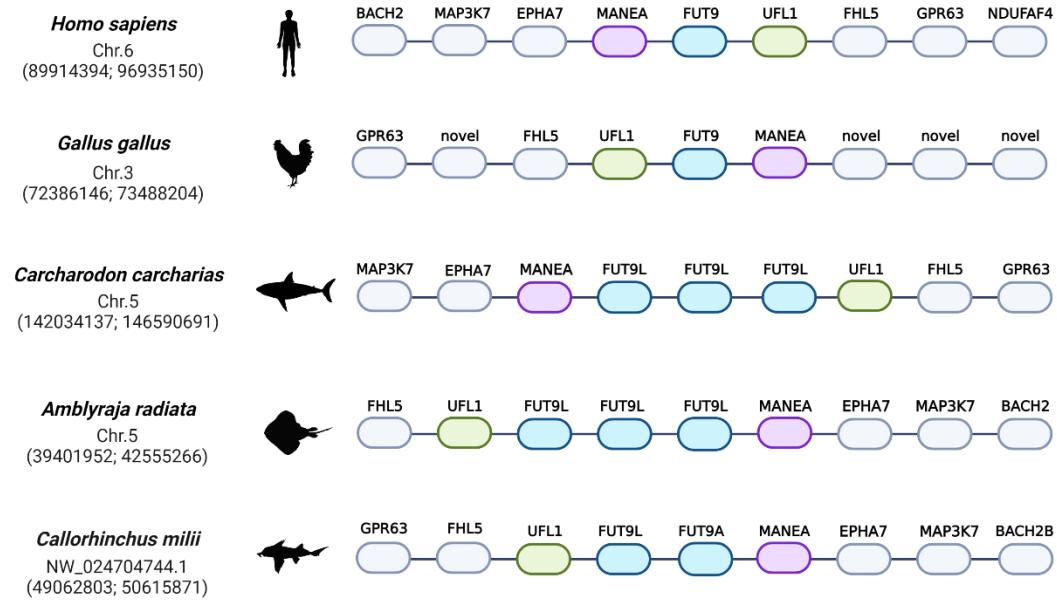

*Figure 66 - **Genomic loci of the FUT9 gene in vertebrate species.** Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

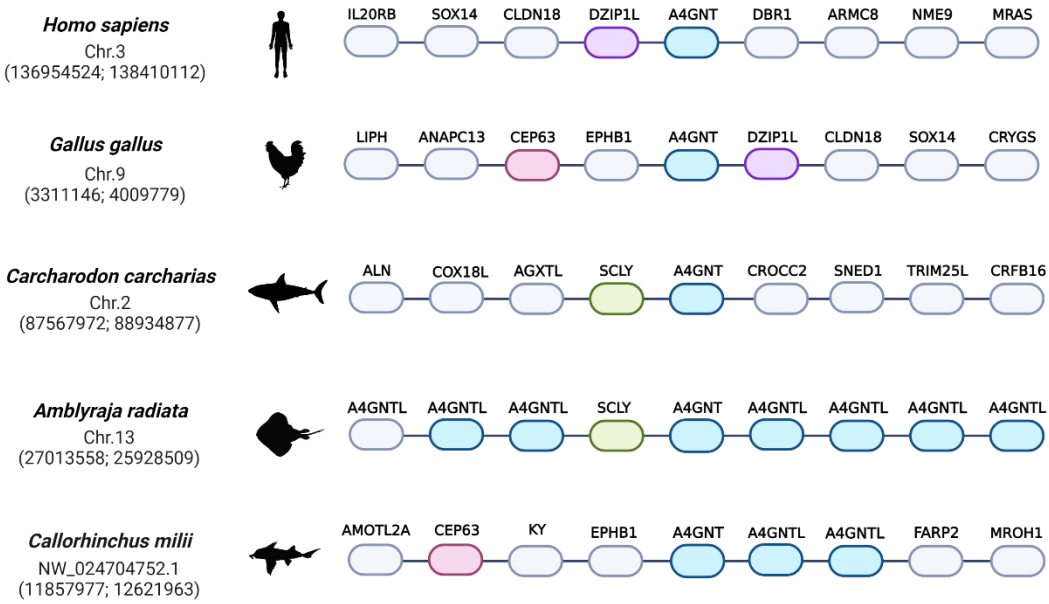

*Figure 68 - **Genomic loci of the A4GNT gene in vertebrate species.** Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*

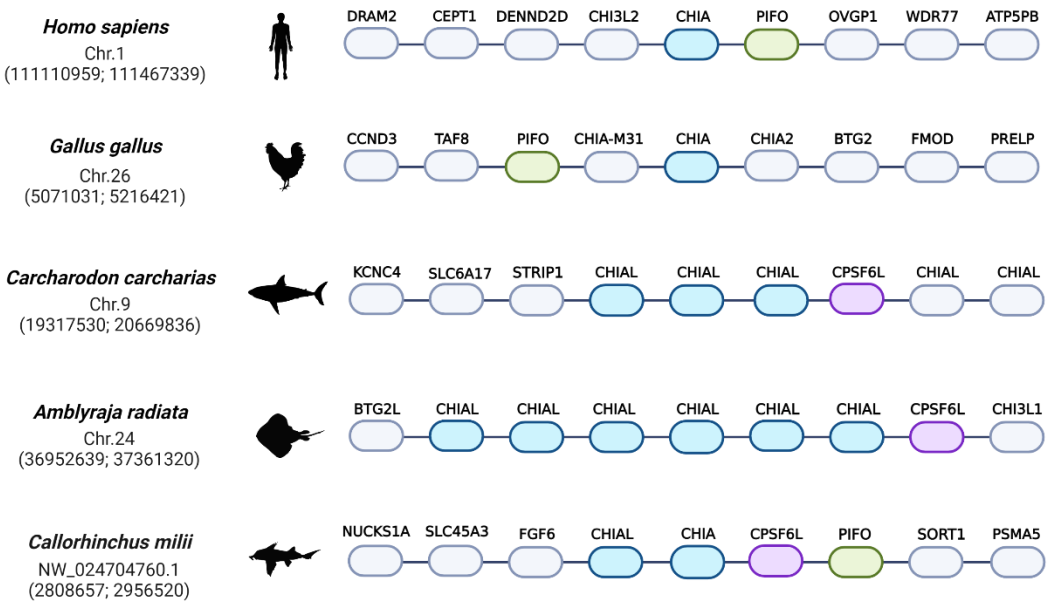

*Figure 67 - **Genomic loci of the CHIA gene in vertebrate species.** Chromosome or scaffold number is indicated. Studied species include Human (Homo sapiens); chicken (Gallus gallus); great white shark (Carcharodon carcharias); thorny skate (Amblyraja radiata) and ghost shark (Callorhinchus milii).*