A peer-reviewed version of this preprint was published in PeerJ on 13 April 2018.

View the peer-reviewed version (peerj.com/articles/4593), which is the preferred citable publication unless you specifically need to cite this preprint.

https://doi.org/10.7717/peerj.4593
Phylogenetic evidence suggests a later origin of the DRD$_{2l}$ and DRD$_{4rs}$ dopamine receptor gene lineages

Juan C Opazo Corresp., 1, Kattina Zavala 1, Soledad Miranda-Rottmann 2, Roberto Araya 2

1 Instituto de Ciencias Ambientales y Evolutivas, Universidad Austral de Chile, Valdivia, Chile
2 Faculty of Medicine, University of Montreal, Montreal, Canada

Corresponding Author: Juan C Opazo
Email address: jopazo@gmail.com

Dopamine receptors are integral membrane proteins whose endogenous ligand is dopamine. They play a fundamental role in the central nervous system and dysfunction of dopaminergic neurotransmission is responsible for the generation of a variety of neuropsychiatric disorders. From an evolutionary standpoint, phylogenetic relationships among the DRD$_1$ class of dopamine receptors are still a matter of debate as in the literature different tree topologies have been proposed. In contrast, phylogenetic relationships among the DRD$_2$ group of receptors are well understood. Understanding the time of origin of the different dopamine receptors is also an issue that needs further study, especially for the genes that have restricted phyletic distributions (e.g. DRD$_{2l}$ and DRD$_{4rs}$).

Thus, the goal of this study was to investigate the evolution of dopamine receptors, with emphasis on shedding light on the phylogenetic relationships among the D$_1$ class of dopamine receptors and the time of origin of the DRD$_{2l}$ and DRD$_{4rs}$ gene lineages. Our results recovered the monophyly of the two groups of dopamine receptors. Within the DRD$_1$ group the monophyly of each paralog was recovered with strong support, and phylogenetic relationships among them were well resolved. Within the DRD$_1$ class of dopamine receptors we recovered the sister group relationship between the DRD$_{1C}$ and DRD$_{1E}$, and this clade was recovered sister to a cyclostome sequence. The DRD$_1$ clade was recovered sister to the aforementioned clade, and the group containing DRD$_5$ receptors was sister to all other DRD$_1$ paralogs. In agreement with the literature, among the DRD$_2$ class of receptors, DRD$_2$ was recovered sister to DRD$_3$, whereas DRD$_4$ was sister to the DRD$_2$/DRD$_3$ clade. According to our phylogenetic tree, the DRD$_{2l}$ and DRD$_{4rs}$ gene lineages would have originated in the ancestor of gnathostomes between 615 and 473 mya. Conservation of sequences required for dopaminergic neurotransmission and small changes in regulatory regions suggest a functional refinement of the dopaminergic pathways along evolution.
Phylogenetic evidence suggests a later origin of the DRD$_2$ and DRD$_4$ dopamine receptor gene lineages

Juan C. Opazo$^1$, Kattina Zavala$^1$, Soledad Miranda-Rottmann$^2$, Roberto Araya$^2$

$^1$Instituto de Ciencias Ambientales y Evolutivas, Facultad de Ciencias, Universidad Austral de Chile, Valdivia, Chile. $^2$Department of Neurosciences, Faculty of Medicine, University of Montreal, Montreal, QC, Canada

Corresponding author
Juan C. Opazo
Instituto de Ciencias Ambientales y Evolutivas
Facultad de Ciencias
Universidad Austral de Chile
Valdivia, Chile
Phone: +56632221674
Email: jopazo@gmail.com

Keywords: Dopamine Receptors, Gene Family Evolution, whole genome duplications.
Abstract

Dopamine receptors are integral membrane proteins whose endogenous ligand is dopamine. They play a fundamental role in the central nervous system and dysfunction of dopaminergic neurotransmission is responsible for the generation of a variety of neuropsychiatric disorders.

From an evolutionary standpoint, phylogenetic relationships among the DRD₁ class of dopamine receptors are still a matter of debate as in the literature different tree topologies have been proposed. In contrast, phylogenetic relationships among the DRD₂ group of receptors are well understood. Understanding the time of origin of the different dopamine receptors is also an issue that needs further study, especially for the genes that have restricted phyletic distributions (e.g. DRD₂₁ and DRD₄ₑ). Thus, the goal of this study was to investigate the evolution of dopamine receptors, with emphasis on shedding light on the phylogenetic relationships among the D₁ class of dopamine receptors and the time of origin of the DRD₂₁ and DRD₄ₑ gene lineages. Our results recovered the monophyly of the two groups of dopamine receptors. Within the DRD₁ group the monophyly of each paralog was recovered with strong support, and phylogenetic relationships among them were well resolved. Within the DRD₁ class of dopamine receptors we recovered the sister group relationship between the DRD₁ₑ and DRD₁ₑ, and this clade was recovered sister to a cyclostome sequence. The DRD₁ clade was recovered sister to the aforementioned clade, and the group containing DRD₃ receptors was sister to all other DRD₁ paralogs. In agreement with the literature, among the DRD₂ class of receptors, DRD₂ was recovered sister to DRD₃, whereas DRD₄ was sister to the DRD₂/DRD₃ clade. According to our phylogenetic tree, the DRD₂₁ and DRD₄ₑ gene lineages would have originated in the ancestor of gnathostomes between 615 and 473 mya. Conservation of sequences required for dopaminergic neurotransmission and small changes in regulatory regions suggest a functional refinement of the dopaminergic pathways along evolution.
Introduction

The availability of whole genome sequences offers a great opportunity to study the evolution of genes involved in physiological processes in a variety of living organisms. The diversity of gene content and its evolutionary history are fundamental pieces of information that should be taken into account when comparing the physiology of different species. To understand the evolution of genes it is necessary to reconcile their evolutionary history by comparing relationships among genes –i.e. gene trees– and among species involved in the study –i.e. species trees. Thus, comparing both trees represents a powerful approach to infer homology, time of origin, birth-and-death processes, gene conversion events among others.

Dopamine receptors are integral membrane proteins that mediate the action of dopamine (Beaulieu & Gainetdinov, 2011). They play fundamental roles in functions associated with the central nervous system including learning, cognition, memory, feeding, sleep, and motor control among others (Beaulieu & Gainetdinov, 2011). Peripherally, these receptors are also involved in hormonal regulation, cardiovascular function, renal function, and olfaction among others (Beaulieu & Gainetdinov, 2011). Several human disorders are associated with dopamine receptors including parkinson’s disease, schizophrenia, Tourette’s syndrome, huntington’s disease, drug abuse and addiction, bipolar disorder, depression, and hypertension among others (Hussain & Lokhandwala, 1998; Hisahara & Shimohama, 2011; Chu et al., 2012; Chen et al., 2013; Denys et al., 2013; Brisch, 2014; Ashok et al., 2017). Based on their pharmacological properties, dopamine receptors are classified into two major groups: the DRD₁ group, which includes DRD₁, DRD₃, DRD₁C, and DRD₁E; and the DRD₂ group that includes DRD₂, DRD₂l, DRD₃, DRD₄, and DRD₄rs (Yamamoto et al., 2015). Today it is well known that these groups originated independently such that the ability to bind dopamine was acquired twice during the
evolution of biogenic amine receptors (Callier et al., 2003; Yamamoto et al., 2013, 2015; Spielman, Kumar & Wilke, 2015). Although both groups share the ability to bind dopamine, they also show the signature of their independent histories as they differ in several other characteristics (Sibley, 1999; Beaulieu & Gainetdinov, 2011).

From an evolutionary standpoint, evolutionary relationships among the members of the DRD\(_1\) class of dopamine receptors are still a matter of debate; different phylogenetic hypotheses have been proposed in the literature. For example, DRD\(_1\) has been recovered sister to DRD\(_5\), a clade that in turn is recovered sister to DRD\(_{1C}\); in these studies DRD\(_{1E}\) is recovered sister to all other DRD\(_1\) members (Callier et al., 2003; Yamamoto et al., 2013). In other cases, the clade containing DRD\(_1\) sequences has been recovered sister to DRD\(_{1C}\), and this group is sister to DRD\(_5\) (Le Crom et al., 2004). A case in which the monophyly of DRD\(_{1E}\) has not been recovered has also been reported (Haug-Baltzell et al., 2015). There is also a case in which the members of the DRD\(_1\) class of dopamine receptors have been recovered as two distinct clades, one that includes DRD\(_1\) and DRD\(_3\) and another grouping DRD\(_{1C}\) and DRD\(_{1E}\) (Yamamoto et al., 2015). In contrast to the lack of phylogenetic agreement among the DRD\(_1\) class of dopamine receptors, phylogenetic relationships among the members of the DRD\(_2\) class of dopamine receptors are well resolved as in most studies DRD\(_2\) is recovered sister to DRD\(_3\), whereas DRD\(_4\) is recovered sister to the DRD\(_2/DRD_3\) clade (Callier et al., 2003; Haug-Baltzell et al., 2015; Spielman, Kumar & Wilke, 2015; Yamamoto et al., 2015). Understanding the time of origin of the different dopamine receptors is also an issue that needs further study, especially for the genes that possess restricted phyletic distributions (e.g. DRD\(_{2L}\) and DRD\(_{4N}\)). Regarding the time of origin, different hypotheses are associated with different phylogenetic predictions. Therefore, a phylogenetic tree that is built on adequate taxonomic sampling and an adequate number of genes should provide valuable information to understand the time of origin of dopamine receptors and also about their sister group relationships.
The goal of this study was to investigate the evolution of dopamine receptors, with emphasis on shedding light on the phylogenetic relationships among the DRD₁ class of dopamine receptors and the time of origin of the DRD₂₁ and DRD₄ᵣᵣ gene lineages. Our results recovered the monophyly of the two groups of dopamine receptors. Within the DRD₁ class of receptors, the monophyly of each paralog was recovered with strong support, and phylogenetic relationships among them were well resolved. We recovered the sister group relationship between the DRD₁C and DRD₁E receptors, and this clade was recovered sister to a cyclostome sequence. The DRD₁ clade was recovered sister to the aforementioned clade, and the group containing the DRD₅ receptors was sister to all other DRD₁ paralogs. This topology represents a new phylogenetic hypothesis for the evolution of this group of dopamine receptors. In agreement with the literature, among the D₂ class of dopamine receptors, DRD₂ was recovered sister to DRD₃ whereas DRD₄ was sister to the DRD₂/DRD₃ clade. Finally, our phylogenetic evidence suggests a later origin of the DRD₂₁ and DRD₄ᵣᵣ gene lineages.

Materials and methods

DNA data and phylogenetic analyses

We used bioinformatic procedures to annotate dopamine receptors in species of all major groups of vertebrates. Our sampling included mammals, birds, reptiles, amphibians, coelacanths, teleost fish, holostean fish, cartilaginous fish and cyclostomes (Supplementary Table S1). We also included sequences of the α₂-adrenoreceptors (ADRA2A, ADRA2B, ADRA2C, ADRA2D), and β-adrenoreceptors (ADRB1, ADRB2 and ADRB3)(Supplementary Table S1). Our final dataset contained 396 dopamine receptor sequences. Amino acid sequences were aligned using the FFT-NS-i strategy from MAFFT v.7 (Katoh & Standley, 2013). We used the proposed model tool of IQ-Tree(Trifinopoulos et al., 2016) to select the best-fitting model of amino acid substitution (JTT + R9). We performed a maximum likelihood analysis to obtain the best tree using the
program IQ-Tree (Trifinopoulos et al., 2016); support for the nodes was assessed with 1,000 bootstrap pseudoreplicates using the ultrafast routine. Human ADRA1A, ADRA1B, and ADRA1D sequences were used as outgroups.

Assessments of Conserved Synteny

We examined genes found upstream and downstream of the dopamine receptor genes of representative vertebrate species. We used the estimates of orthology and paralogy derived from the EnsemblCompara database (Herrero et al., 2016); these estimates are obtained from an automated pipeline that considers both synteny and phylogeny to generate orthology mappings. These predictions were visualized using the program Genomicus v90.01 (Louis et al., 2015). Our analyses were performed in humans (Homo sapiens), chicken (Gallus gallus), spotted gar (Lepisosteus oculatus) and elephant shark (Callorhinchus milii). In the case of the elephant shark (http://esharkgenome.imcb.a-star.edu.sg/), the genomic pieces containing the dopamine receptor genes were annotated, and predicted genes were then compared with the non-redundant protein database using Basic Local Alignment Search Tool (BLAST)(Altschul et al., 1990).

Molecular structure and graphics

Molecular graphics and analyses of the human DRD4 protein structure were performed with the UCSF Chimera package (Pettersen et al., 2004) using the 1.96Å resolution structural file PDB ID: 5WIV (Wang et al., 2017). Sequences were aligned using Vector NTI Express (Thermo Fisher). Human protein sequences DRD2: NP_000786.1 and DRD4: NP_000788 were use as reference for the numbering and alignment.

Results and Discussion

Overview of the evolution of dopamine receptors
In this work we performed an evolutionary study of dopamine receptors in representative species of all major groups of vertebrates. We combined gene phylogenies and synteny analyses with the main goal of understanding the duplicative history of the DRD₁ class of dopamine receptors and the time of origin of the DRD₂₁ and DRD₂₄₃ gene lineages.

Our phylogenetic tree recovered the monophyly of the two groups of dopamine receptors (Fig. 1). In the first clade we recovered the sister group relationship between the DRD₁ class of receptors and a clade containing β-adrenoceptors (Fig. 1); in the second clade, the DRD₂ receptors were recovered sister to the α₂-adrenoceptors (Fig. 1). This phylogenetic arrangement is in agreement with previous results (Yamamoto et al., 2013; Spielman, Kumar & Wilke, 2015; Céspedes et al., 2017; Zavala et al., 2017) and reflects the fact that the ability to bind dopamine was acquired twice during the evolutionary history of biogenic amine receptors (Callier et al., 2003; Yamamoto et al., 2015). Although the DRD₁ and DRD₂ receptor families share the ability to bind dopamine, they also show the signature of their independent histories as they differ in several other characteristics. From a structural standpoint, the DRD₁ class of receptors is characterized by the lack of introns, a short third cytoplasmatic loop and a long C-terminal tail. Conversely, DRD₂ possess up to six introns, encoding a long third cytoplasmatic loop and a short C-terminal tail (Gingrich & Caron, 1993). From a biochemical perspective, the DRD₁ group of receptors activates the Gαsolf family of G proteins stimulating adenilate cyclase activity and production of cAMP, whereas DRD₂ group of receptors activate the Gαi/o family of G proteins inhibiting adenilate cyclase activity and reducing levels of cAMP (Sibley, 1999; Beaulieu & Gainetdinov, 2011). From a synaptic anatomy standpoint, the DRD₁ class of receptors is located exclusively at the postsynaptic site whereas the DRD₂ class is found both in pre- and postsynaptic terminals (Sibley, 1999; Beaulieu & Gainetdinov, 2011).
According to our phylogenetic analyses, the monophyly of the DRD₁ class of dopamine receptors, as well as the monophyly of each paralog (DRD₁, DRD₅, DRD₁C, and DRD₁E), were recovered with strong support (Fig. 1). In all cases synteny analyses provided further support for the identity of the four DRD₁ clades recovered in our phylogenetic tree (Fig. 2). Phylogenetic relationships among the different DRD₁ lineages were well resolved (Fig. 1). We recovered the sister group relationship between the DRD₁C and DRD₁E dopamine receptors (Fig. 1), and this clade was recovered sister to a cyclostome sequence (Fig. 1). The DRD₁ clade was recovered sister to the aforementioned clade, and the group containing DRD₅ sequences was recovered sister to all other DRD₁ paralogs (Fig. 1). Although in the literature there are studies reporting dopamine receptor phylogenies (Callier et al., 2003; Le Crom et al., 2004; Yamamoto et al., 2013, 2015; Haug-Baltzell et al., 2015), they are not directly comparable as the taxonomic and/or family membership sampling differ. Beyond this point, phylogenetic relationships among the DRD₁ class of dopamine receptors seem to still be a matter of debate. In some cases DRD₁ has been recovered sister to DRD₅, a clade that in turn is recovered sister to DRD₁C; in these studies DRD₁E is recovered sister to all other DRD₁ (Callier et al., 2003; Le Crom et al., 2003; Yamamoto et al., 2013). In other studies the clade containing DRD₁ sequences has been recovered sister to DRD₁C, and this group is sister to DRD₅ (Le Crom et al., 2004). A case in which the monophyly of DRD₁E is not recovered has also been reported (Haug-Baltzell et al., 2015). Finally, there is also a case in which the DRD₁ class of receptors has been recovered as two different clades, one that includes DRD₁ and DRD₅ and another grouping DRD₁C and DRD₁E (Yamamoto et al., 2015).

Thus, our results propose a new phylogenetic hypothesis regarding the evolution of the DRD₁ class of dopamine receptors (Fig. 1). Overall, we believe that our hypothesis is well supported based on a taxonomic sampling that covered all main groups of vertebrates, as well as, the phylogenetic context of the monoamine receptors (Spielman, Kumar & Wilke, 2015).
Phylogenetic relationships among the D₂ class of dopamine receptors

We recovered the monophyly of the DRD₂ class of dopamine receptors with strong support (Fig. 1). The monophyly of all paralogs of this class of receptors are also well supported, defining clear orthology and paralogy (Fig. 1). Synteny analyses provide further support for the evolutionary identity of all D₂ dopamine receptors (Fig. 3). In our phylogenetic tree DRD₂ was recovered sister to DRD₃ with strong support (Fig. 1), whereas DRD₄ was sister to the DRD₂/DRD₃ clade (Fig. 1).

In contrast to the lack of phylogenetic resolution among the DRD₁ class of dopamine receptors, phylogenetic relationships among the DRD₂ class of receptors seem to be well resolved as all studies, including ours, show the same topology ((DRD₂,DRD₃),DRD₄)(Callier et al., 2003; Le Crom et al., 2003; Haug-Baltzell et al., 2015; Spielman, Kumar & Wilke, 2015; Yamamoto et al., 2015).

Phylogenetic evidence for the origin of the DRD₂₁ gene lineage in the ancestor of gnathostomes

In agreement with Yamamoto et al. (2015) (Yamamoto et al., 2015), our phylogenetic analyses also suggest the presence of an extra dopamine receptor gene lineage that is related to DRD₂ gene (Boehmier et al., 2004; Boehmler et al., 2007) (Fig. 4). Although our results agree with Yamamoto et al. (2015)(Yamamoto et al., 2015) regarding the presence of a new dopamine receptor gene lineage, our results suggest a different time of origin.

According to our results, we recovered a strongly supported clade containing the DRD₂₁ sequences of teleost fish, holostean fish, and coelacanths (Fig. 4) sister to the clade containing DRD₂ sequences of gnathostomes (Fig. 4). This tree topology suggests that in the ancestor of gnathostomes, between 615 and 473 mya, the DRD₂ gene underwent a duplication event that gave rise to an extra DRD₂ gene copy – the DRD₂₁ – that was independently lost in the ancestor of tetrapods and cartilaginous fish (Fig. 5). In support of this scenario, our phylogenetic tree recovered a cyclostome sequence sister to the DRD₂/DRD₂₁ clade (Fig. 4). The pattern of gene
conservation found up, and downstream of DRD\textsubscript{2} and DRD\textsubscript{4} genes, provides further support for the presence of two DRD\textsubscript{2} dopamine receptor gene lineages (Fig. 3). For example, in the spotted gar (*Lepisosteus oculatus*), a species that possesses both DRD\textsubscript{2} gene copies, DRD\textsubscript{2} and DRD\textsubscript{2l} are found in different chromosomal locations. The identity of their genomic locations is defined by the presence of upstream and downstream flanking genes all across gnathostome vertebrates.

Thus, the upstream genes ANKK1 and TTC12 and the downstream genes TMPRSS, ZW10, USP28 and HTR3B define the genomic location of the DRD\textsubscript{2} gene lineage, whereas the upstream gene XRCC1 and downstream genes ETHE1, PHLDB3 and IRQQ1 define the genomic location of the DRD\textsubscript{2l} gene lineage (Fig. 3). Importantly, this pattern of conservation is also found in species that lost the DRD\textsubscript{2l} gene from their genomes (Fig. 3).

The evolutionary hypothesis proposed here is different from that proposed by Yamamoto et al. (2015) in which the clade containing DRD\textsubscript{2l} sequences was recovered sister to a clade containing DRD\textsubscript{2} sequences of vertebrates. Thus, according to their phylogeny the duplication event that gave rise to the DRD\textsubscript{2l} gene would have occurred in the ancestor of vertebrates, between 676 and 615 mya, even though they claim that the origin of this gene occurred after the Osteichthyes-Chondrichthyes divergence, between 473 and 435 mya (Yamamoto et al., 2015). Beyond this discrepancy, both evolutionary scenarios proposed in the study of Yamamoto et al. (2015) are different from ours.

An amino acid alignment of both DRD\textsubscript{2} gene lineages revealed that in the case of the spotted gar (*Lepisosteus oculatus*) and the coelacanth (*Latimeria chalumnae*) the divergence between DRD\textsubscript{2} and DRD\textsubscript{2l} receptors is approximately 30% whereas it is approximately 45% in zebrafish (*Danio rerio*). These estimates are in agreement with previous reports (Boehmier et al., 2004). Additionally, the human DRD\textsubscript{2} amino acid sequence was aligned to the zebrafish, coelacanth and spotted gar DRD\textsubscript{2l} sequence to infer functionally significant changes (Fig. 6). The binding sites for dopamine and DRD\textsubscript{2} agonists and antagonists are conserved among these
species. However, the adjacent hydrophobic pocket, which confers ligand specificity to DRD$_2$ is not conserved (Fig. 6). While in humans, coelacanths and spotted gar the second amino acid of the third transmembrane domain (TM3) is phenylalanine (F), it is leucine (L) in zebrafish. This change from an aromatic to an aliphatic amino acid could change the zebrafish DRD$_2$ ligand specificity and therefore its function. The site that confers specificity to the human G protein subunit G$_{ai}$ (Senogles et al., 2004) (Fig. 6; orange asterisks) is not conserved among species. The side chain size, shape and polarity changes observed could potentially influence the receptor/G protein coupling specificity, suggesting important evolutionary differences.

Phylogenetic evidence for the origin of DRD$_{4s}$ gene lineage in the ancestor of gnathostomes

Also in agreement with Yamamoto et al. (2015) (Yamamoto et al., 2015) our phylogenetic reconstruction identified an extra dopamine receptor gene lineage that is related to the DRD$_4$ gene (Fig. 1 and 7). According to our phylogenetic tree, a strongly supported clade that contains dopamine receptors of bony fish and coelacanths was recovered sister to the DRD$_4$ clade of gnathostomes (Fig. 7). Similarly to the DRD$_{2l}$ gene lineage, this topology suggests that the DRD$_4$ gene underwent a duplication event in the ancestor of gnathostomes, between 615 and 473 mya, giving rise to an extra copy of the DRD$_4$ gene. During the radiation of the group, one of the copies (DRD$_4$) was retained in all main groups of gnathostomes, whereas the other was only retained in bony fish and coelacanths (Fig. 5). In agreement with this hypothesis, our phylogenetic reconstruction recovered a lamprey sequence sister to the DRD$_{4u}$/DRD$_{4s}$ clade (Fig. 7). Synteny analyses provide further support to our phylogenetic tree, as the genomic locations that harbor both DRD$_4$ gene lineages are different (Fig. 3). Thus, there are four upstream genes (SCT, CDHR5, IRF7 and PHRF1) and four genes downstream (DEAF1, TMEM80, EPS8L2 and TALDO1) that define the identity of the DRD$_4$ genomic location (Fig. 3). Similarly, there are upstream genes (KCP, CDHR5, IRF5 and TNOP3) and downstream genes (ATP6V1F) of the
The DRD₄ gene that define the identity of its genomic location (Fig. 3). Similar to that found for the DRD₂ genes, our evolutionary hypothesis regarding the origin of the DRD₄ gene lineage is different from the scenario proposed by Yamamoto et al. (2015) (Yamamoto et al., 2015). According to their results, the clade containing DRD₄ sequences was recovered sister to a clade containing DRD₄ sequences of vertebrates. Therefore, their phylogenetic tree suggests that the evolutionary origin of the DRD₄ gene lineage would be in the ancestor of vertebrates, between 676 and 615 mya, as a product of two rounds of whole genome duplication (Yamamoto et al., 2015). Thus, both studies suggest different evolutionary scenarios regarding the time of origin of the DRD₄ gene lineage.

The divergence between the DRD₄ and DRD₄rs gene lineages was found to be higher compared to that estimated for the DRD₂ gene lineages. In the case of the spotted gar (Lepisosteus oculatus) and the coelacanth (Latimeria chalumnae) divergence was approximately 45% whereas in zebrafish (Danio rerio) it was approximately 49%. The human DRD₄ amino acid sequence was aligned to the zebrafish, coelacanth and spotted gar DRD₄rs sequence (Fig. 8). The binding sites for dopamine and DRD₂ agonists and antagonists are conserved among species.

Interestingly, two sites in the hydrophobic pocket of the dopamine receptor differ. The first site is located in the selectivity region of DRD₄, where a change from tyrosine (Y) to phenylalanine (F) occurs at position 91 (F91) of the human receptor sequence (Fig. 8; green asterisk). At the second site (Fig. 8; green asterisk) in position 193 of the human DRD₄, the isoleucine (I) in the corresponding spotted gar sequence is changed to valine (V) in the other species (V193). To understand the potential effects that these changes might have on DRD₄ function we used the recently uncovered crystal structure of the human DRD₄ sequence coupled to the antipsychotic drug nemonapride (Wang et al., 2017). All amino acids within 4Å of the active site are conserved (Fig. 9A and 9A’, red amino acids; 9B, red dots) except two. First, F91, which is, located in the recently discovered extended binding pocket, a region poorly conserved among dopamine
receptors that is key for receptor class specificity (Wang et al., 2017). Second, V193 that is located in the orthosteric-binding pocket that modulates agonist response (Lane et al., 2013) (Fig. 9A green amino acids, 9B, green dots). Simulated mutagenesis to the amino acids present in the spotted gar sequence (Fig. 9B’, green amino acids) shows how F91 changes the shape of the extended binding pocket compared to the human sequence, suggesting an important evolutionary change in ligand specificity and receptor function.

Duplicative history and ancestral gene repertoires

To understand the duplicative history of dopamine receptors, including the definition of ancestral repertoires, it is necessary to reconcile the evolutionary history of the gene lineages with the sister group relationships among the species involved. According to our results, the presence of differentiated dopamine receptors in vertebrates (Fig. 5) allowed us to infer that at some point of time the vertebrate ancestor possessed two dopamine receptors, one of each class (Fig. 10). After the two rounds of whole genome duplications (WGD) that occurred in the ancestor of the group (Garcia-Fernàndez & Holland, 1994; Dehal & Boore, 2005) each ancestral gene (DRD\textsubscript{1anc} and DRD\textsubscript{2anc}) gave rise to four genes in each class of dopamine receptors (Fig. 10). In support of this hypothesis, the DRD\textsubscript{1} and DRD\textsubscript{2} classes of dopamine receptors appear in the repository of genes that originated and were retained after the WGDs occurred in the ancestor of vertebrates (Singh, Arora & Isambert, 2015). The fact that non-vertebrate chordates possess just one DRD\textsubscript{1} (Kamesh, Aradhyam & Manoj, 2008; Burman et al., 2009) and that the four chromosomal locations where the DRD\textsubscript{1} class of receptors are located in humans derive from a single linkage group in the chordate ancestor(Putnam et al., 2008) provide support to our hypothesis. Overall, three out of the four DRD\textsubscript{1} originated as a product of the WGDs were retained in the genome of the vertebrate ancestor (DRD\textsubscript{1}, DRD\textsubscript{5} and DRD\textsubscript{1C/E}; Fig. 10). After that, in the gnathostome ancestor the DRD\textsubscript{1C/E} gene underwent a duplication event that gave rise to the actual DRD\textsubscript{1C} and DRD\textsubscript{1E} genes (Fig.
In support of this, we recovered a cyclostome sequence sister to the clade containing the DRD$_{1C}$ and DRD$_{1E}$ genes. Thus, the gnathostome ancestor that existed between 615 and 473 mya had a repertoire of four DRD$_1$ genes: DRD$_1$, DRD$_5$, DRD$_{1C}$ and DRD$_{1E}$ (Fig. 10). In teleost fish, a group that experienced an extra round of whole genome duplication (Meyer & Van de Peer, 2005; Kasahara et al., 2007; Sato & Nishida, 2010; Glasauer & Neuhauss, 2014), all DRD$_1$ doubled in number, however, three out of the four gene lineages retained duplicated copies (Fig. 5) (Yamamoto et al., 2013, 2015).

Similarly to the DRD$_1$ class of receptor, the vertebrate specific WGDs originated four DRD$_2$ genes, three of which were maintained in the genome of extant species (DRD$_2$, DRD$_3$ and DRD$_4$; Fig. 5 and 10). In the ancestor of gnathostomes the DRD$_2$ gene underwent a duplication event that gave rise to an extra copy of the DRD$_2$ gene (DRD$_{2l}$; Fig. 4 and 10). In this case both genes followed different evolutionary trajectories. On one hand DRD$_2$ was retained in the genome of all of the main groups of vertebrates (Fig. 5) whereas DRD$_{2l}$ was only retained in coelacanths and bony fish (Fig. 5) (Yamamoto et al., 2015). Similarly, the DRD$_4$ gene also underwent a duplication event that gave rise to an extra copy of the gene (DRD$_{4r}$; Fig. 7 and 10). This case is similar to that found for the DRD$_2$ gene, as one of the copies (DRD$_4$) was retained in the genome of all of the main groups of vertebrates, while the other was independently lost in tetrapods and cartilaginous fish (Fig. 6). Consequently, the ancestor of gnathostome vertebrates possessed a repertoire of five DRD$_2$ class of dopamine receptors: DRD$_2$, DRD$_{2l}$, DRD$_3$, DRD$_4$ and DRD$_{4r}$ (Fig. 10). As a consequence of the teleost-specific genome duplication (Meyer & Van de Peer, 2005; Kasahara et al., 2007; Sato & Nishida, 2010; Glasauer & Neuhauss, 2014), teleost fish doubled their number of DRD$_2$ receptors, however extant species retained duplicated copies in just two gene lineages (Fig. 5) (Yamamoto et al., 2015).

Concluding remarks
We present an evolutionary study of the dopamine receptors with special emphasis on unraveling the phylogenetic relationships of the D₁ class of receptors and the time of origin of the DRD₂ and DRD₄ gene lineages. Our study comprised taxonomic sampling that included representative species of all main groups of vertebrates in addition to other vertebrate biogenic amine receptors. Thus, we were able to reconstruct in a single phylogenetic tree the evolutionary history of both classes of dopamine receptors. In the case of the DRD₁ class, our results propose a new phylogenetic hypothesis in which DRD₁C was recovered sister to DRD₁E and this clade was recovered sister to a cyclostome sequence. DRD₁ was recovered sister to the aforementioned clade, and the group containing the DRD₃ sequences was sister to all other DRD₁ paralogs. According to our phylogenetic tree, the evolutionary origin of the DRD₂ and DRD₄ gene lineages would have happened in the ancestor of gnathostomes between 615 and 473 mya, which differs from current proposed scenarios. Of special interest is the analysis of sequences required for dopaminergic neurotransmission. We found high conservation of agonist and antagonist sites suggesting evolutionary conserved dopaminergic pathways. We also found small variation in the dopamine-binding regulatory regions showing a refinement of ligand specificity and big variations in G protein-coupling sequences suggesting differences in downstream signaling cascades through evolution. These new data on evolutionary divergence may help with the rational design of new agonist and antagonist to modulate the dopaminergic pathway.

**Funding**

This work was supported by the Fondo Nacional de Desarrollo Científico y Tecnológico (FONDECYT 1160627) to JCO, from the Canadian Institutes of Health Research (CIHR) grant RNI00109 and Canadian Foundation for Innovation (CFI) equipment grant *Fonds des leaders* 29970 to RA.
**Figure legends**

**Figure 1.** Maximum likelihood tree depicting evolutionary relationships among dopamine receptors in vertebrates. Numbers on the nodes correspond to maximum likelihood ultrafast bootstrap support values. Human ADRA\textsubscript{1A}, ADRA\textsubscript{1B}, and ADRA\textsubscript{1D} sequences were used as outgroups.

**Figure 2.** Patterns of conserved synteny in the chromosomal regions that harbor the DRD\textsubscript{1} class of dopamine receptors.

**Figure 3.** Patterns of conserved synteny in the chromosomal regions that harbor the DRD\textsubscript{2} class of dopamine receptors.

**Figure 4.** Maximum likelihood trees depicting evolutionary relationships among DRD\textsubscript{2} and DRD\textsubscript{2i} dopamine receptors in vertebrates. Numbers on the nodes correspond to maximum likelihood ultrafast bootstrap support values. This tree topology does not represent novel phylogenetic analyses; they are the DRD\textsubscript{2}/DRD\textsubscript{2i} clade that was recovered from Fig. 1.

**Figure 5.** Maximum likelihood trees depicting evolutionary relationships among DRD\textsubscript{4} and DRD\textsubscript{4rs} dopamine receptors in vertebrates. Numbers on the nodes correspond to maximum likelihood ultrafast bootstrap support values. This tree topology does not represent novel phylogenetic analyses; they are the DRD\textsubscript{4}/DRD\textsubscript{4rs} clade that was recovered from Fig. 1.

**Figure 6.** Phyletic distribution of dopamine receptor genes in vertebrates.
Figure 7. Alignment of the human dopamine receptor 2 (DRD$_2$) with zebrafish (Danio rerio), coelacanth (Latimeria chalumnae) and spotted gar (Lepisosteus oculatus) dopamine receptor 21 (DRD$_{21}$). Shaded regions denote transmembrane domains according to UniProt. Dopamine binding sites, agonist and antagonist binding sites were predicted with theoretical and computational techniques (Yashar et al., 2004) and experimental evidence (Shi & Javitch, 2002). Amino acids in the third intracellular loop conferring G protein subunit Gai specificity (Senogles et al., 2004) are indicated by orange asterisks.

Figure 8. Alignment of the human dopamine receptor 4 (DRD$_4$) with zebrafish (Danio rerio), coelacanth (Latimeria chalumnae) and spotted gar (Lepisosteus oculatus) dopamine receptor 4rs (DRD$_{4rs}$). Shaded regions denote transmembrane domains according to UniProt. Dopamine binding sites (red dots) were determined by site directed mutagenesis (Cummings et al., 2010) and homology to DRD$_2$. Antagonist binding sites and hydrophobic pocket-including selectivity region were obtained from mutagenesis studies (Cummings et al., 2010) and from the crystal structure of the receptor coupled to the antagonist nemonapride (Wang et al., 2017). Non-conserved amino acids in the nemonapride binding pocket are labeled with green asterisks. Binding sites for the selective agonist UCSF924 are also shown (light blue dot).

Figure 9. Structural details of the human DRD$_4$ binding site to the antipsychotic drug nemonapride (in blue). (A) Conserved amino acids within 4Å of the drug molecule are shown with functional groups (in red). Non-conserved amino acids (in green) were changed (inset A’) to the residue present in the fish species: F91Y and V193I. Mutagenesis was simulated choosing the rotamer with the highest probability (B). Partial alignment of the human dopamine receptor 4 (DRD$_4$) with zebrafish (Danio rerio), coelacanth (Latimeria chalumnae) and spotted gar (Lepisosteus oculatus) dopamine receptor 4rs (DRD$_{4rs}$) showing the numbers corresponding to the
Figure 10. An evolutionary hypothesis regarding the origin of dopamine receptor genes in vertebrates. According to this model, the vertebrate ancestor possessed two dopamine receptors, one of each class. However, after the two rounds of whole genome duplications (WGD) that occurred in the ancestor of the group each ancestral gene (DRD$_{1\text{anc}}$ and DRD$_{2\text{anc}}$) gave rise to four genes in each class of receptors. In the case of the DRD$_1$ class of dopamine receptors three out of the four genes originated as a product of the WGDs were retained in the genome of the vertebrate ancestor. In the gnathostome ancestor, the DRD$_{1\text{C/E}}$ gene underwent a duplication event that gave rise to the actual DRD$_{1\text{C}}$ and DRD$_{1\text{E}}$ genes. Thus, the gnathostome ancestor that existed between 615 and 473 mya had a repertoire of four DRD$_1$ genes: DRD$_1$, DRD$_5$, DRD$_{1\text{C}}$ and DRD$_{1\text{E}}$. In the case of the DRD$_2$ group of receptors, the vertebrate WGDs originated four genes, three of which were maintained in the genome of extant species (DRD$_2$, DRD$_3$ and DRD$_4$). In the ancestor of gnathostomes, the DRD$_2$ gene underwent a duplication event that gave rise to an extra copy of the DRD$_2$ gene (DRD$_{2\text{L}}$). Similarly, the DRD$_4$ gene also underwent a duplication event that gave rise to an extra copy of the gene (DRD$_{4\text{RS}}$). Thus, the ancestor of gnathostome vertebrates possessed a repertoire of five DRD$_2$ genes: DRD$_2$, DRD$_{2\text{L}}$, DRD$_3$, DRD$_4$ and DRD$_{4\text{RS}}$. 

human DRD$_4$ sequence (NP_000788). Conserved and non-conserved amino acids shown in (A) are indicated with red and green dots respectively. Non-conserved amino acids within the region are also shown in green fonts.
References


Comparative Endocrinology 250. DOI: 10.1016/j.ygcen.2017.06.006.


Sato Y., Nishida M. 2010. Teleost fish with specific genome duplication as unique models of vertebrate evolution. *Environmental Biology of Fishes* 88:169–188. DOI: 10.1007/s10641-


Singh PP., Arora J., Isambert H. 2015. Identification of Ohnolog Genes Originating from Whole Genome Duplication in Early Vertebrates, Based on Synteny Comparison across Multiple Genomes. PLOS Computational Biology 11:e1004394. DOI: 10.1371/journal.pebi.1004394.


Yamamoto K., Mirabeau O., Bureau C., Blin M., Michon-Coudouel S., Demarque M., Vernier P.


Figure 1 (on next page)

Maximum likelihood tree depicting evolutionary relationships among dopamine receptors in vertebrates.

Maximum likelihood tree depicting evolutionary relationships among dopamine receptors in vertebrates. Numbers on the nodes correspond to maximum likelihood ultrafast bootstrap support values. Human ADRA$_{1A}$, ADRA$_{1B}$, and ADRA$_{1D}$ sequences were used as outgroups.
Figure 2 (on next page)

Patterns of conserved synteny in the chromosomal regions that harbor the DRD₁ class of dopamine receptors.

Patterns of conserved synteny in the chromosomal regions that harbor the DRD₁ class of dopamine receptors.
Figure 3 (on next page)

Patterns of conserved synteny in the chromosomal regions that harbor the DRD₂ class of dopamine receptors.

Patterns of conserved synteny in the chromosomal regions that harbor the DRD₂ class of dopamine receptors.
Figure 4 (on next page)

Maximum likelihood trees depicting evolutionary relationships among DRD$_2$ and DRD$_2l$ dopamine receptors in vertebrates.

Maximum likelihood trees depicting evolutionary relationships among DRD$_2$ and DRD$_2l$ dopamine receptors in vertebrates. Numbers on the nodes correspond to maximum likelihood ultrafast bootstrap support values. This tree topology does not represent novel phylogenetic analyses; they are the DRD$_2$/DRD$_2l$ clade that was recovered from Fig. 1.
Figure 5 (on next page)

Phyletic distribution of dopamine receptor genes in vertebrates.

Phyletic distribution of dopamine receptor genes in vertebrates.
<table>
<thead>
<tr>
<th></th>
<th>D₁-like family</th>
<th>D₂-like family</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DRD1</td>
<td>DRD5</td>
</tr>
<tr>
<td>Mammals</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Birds</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Crocodiles</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Turtles</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Squamates</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Amphibians</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Coelacanths</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Holostean fish</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Teleost fish</td>
<td>✓²X</td>
<td>✓²X</td>
</tr>
<tr>
<td>Cartilaginous fish</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Cyclostomes</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
Figure 6 (on next page)

Alignment of the human dopamine receptor 2 (DRD₂) with zebrafish (Danio rerio), coelacanth (Latimeria chalumnae) and spotted gar (Lepisosteus oculatus) dopamine receptor 2l (DRD₂l).

Shaded regions denote transmembrane domains according to UniProt. Dopamine binding sites, agonist and antagonist binding sites were predicted with theoretical and computational techniques (Yashar et al., 2004) and experimental evidence (Shi & Javitch, 2002). Amino acids in the third intracellular loop conferring G protein subunit Gαi specificity (Senogles et al., 2004) are indicated by orange asterisks.
Figure 7 (on next page)

Maximum likelihood trees depicting evolutionary relationships among DRD$_4$ and DRD$_{4rs}$ dopamine receptors in vertebrates

Maximum likelihood trees depicting evolutionary relationships among DRD$_4$ and DRD$_{4rs}$ dopamine receptors in vertebrates. Numbers on the nodes correspond to maximum likelihood ultrafast bootstrap support values. This tree topology does not represent novel phylogenetic analyses; they are the DRD$_4$/DRD$_{4rs}$ clade that was recovered from Fig. 1.
DRD4

DRD4rs

Human
Human
Gibbon
Gorilla
Orangutan
Macaque
Olive baboon
Harvet
Bushbaby
Pig
Ole
Elephant
Squirrel
Mouse
Rat
Wallaby
American alligator
Chinese alligator
Zebrafish
Chicken
Turkey
Duck
Chinese turtle
Anole lizard
Taiwan habu
Burmese python
Clawed frog
Amazon molly
Platyfish
Tetraodon
Medaka
Fugu
Cave fish
Zebrafish
Cave fish
Zebrafish
Cod
Fugu
Tetraodon
Tetraodon
Stickleback
Tilapia
Spotted gar
Elephant shark
Whale shark
Coelacanth
Amazon molly
Platyfish
Tilapia
Medaka
Cod
Cave fish
Zebrafish
Spotted gar
Coelacanth
Sea lamprey

2.0

99

91

100

100

98

93

28

20
Alignment of the human dopamine receptor 4 (DRD$_4$) with zebrafish (*Danio rerio*), coelacanth (*Latimeria chalumnae*) and spotted gar (*Lepisosteus oculatus*) dopamine receptor 4rs (DRD$_{4rs}$).

Alignment of the human dopamine receptor 4 (DRD$_4$) with zebrafish (*Danio rerio*), coelacanth (*Latimeria chalumnae*) and spotted gar (*Lepisosteus oculatus*) dopamine receptor 4rs (DRD$_{4rs}$). Shaded regions denote transmembrane domains according to UniProt. Dopamine binding sites (red dots) were determined by site directed mutagenesis (Cummings et al., 2010) and homology to DRD$_2$. Antagonist binding sites and hydrophobic pocket-including selectivity region were obtained from mutagenesis studies (Cummings et al., 2010) and from the crystal structure of the receptor coupled to the antagonist nemonapride (Wang et al., 2017). Non-conserved amino acids in the nemonapride binding pocket are labeled with green asterisks. Binding sites for the selective agonist UCSF924 are also shown (light blue dot).
Figure 9 (on next page)

Structural details of the human DRD₄ binding site to the antipsychotic drug nemonapride (in blue)

Structural details of the human DRD₄ binding site to the antipsychotic drug nemonapride (in blue). (A) Conserved amino acids within 4Å of the drug molecule are shown with functional groups (in red). Non-conserved amino acids (in green) were changed (inset A’) to the residue present in the fish species: F91Y and V193I. Mutagenesis was simulated choosing the rotamer with the highest probability (B). Partial alignment of the human dopamine receptor 4 (DRD₄) with zebrafish (Danio rerio), coelacanth (Latimeria chalumnae) and spotted gar (Lepisosteus oculatus) dopamine receptor 4rs (DRD₄rs) showing the numbers corresponding to the human DRD₄ sequence (NP_000788). Conserved and non-conserved aminoacids shown in (A) are indicated with red and green dots respectively. Non-conserved aminoacids within the region are also shown in green fonts.
nemonapride

Human DRD4: VLPLF...ALMAMDVMLCT....CRLERDYVVYSSVCS...WTPFFVHITQALCPACSVPPRLVSAVTLWGY
Zebrafish DRD4rs: VLPLY...ALMTDVMLCT....CRLEDNQFVVYSSVCS...WTPFFVHVTKALCESDYGTPLTISVVTWLY
Coelacanth DRD4rs: VLPLY...ALMTDVMLCT....CRLEDNFFVVYSSAC...WTPFFVHVTKVLCACNIGPTLISVVTWGY
Spotted gar DRD4rs: VLPLY...ALMTDVMLCT....CRLEDDNFIYYSSVCS...WTPFFVHVTKVCDSIIGPTLISVVTWGY
An evolutionary hypothesis regarding the origin of dopamine receptor genes in vertebrates.

According to this model, the vertebrate ancestor possessed two dopamine receptors, one of each class. However, after the two rounds of whole genome duplications (WGD) that occurred in the ancestor of the group each ancestral gene (DRD$_{1anc}$ and DRD$_{2anc}$) gave rise to four genes in each class of receptors. In the case of the DRD$_1$ class of dopamine receptors three out of the four genes originated as a product of the WGDs were retained in the genome of the vertebrate ancestor. In the gnathostome ancestor, the DRD$_{1C/E}$ gene underwent a duplication event that gave rise to the actual DRD$_{1C}$ and DRD$_{1E}$ genes. Thus, the gnathostome ancestor that existed between 615 and 473 mya had a repertoire of four DRD$_1$ genes: DRD$_1$, DRD$_5$, DRD$_{1C}$ and DRD$_{1E}$. In the case of the DRD$_2$ group of receptors, the vertebrate WGDs originated four genes, three of which were maintained in the genome of extant species (DRD$_2$, DRD$_3$ and DRD$_4$). In the ancestor of gnathostomes, the DRD$_2$ gene underwent a duplication event that gave rise to an extra copy of the DRD$_2$ gene (DRD$_{2l}$). Similarly, the DRD$_4$ gene also underwent a duplication event that gave rise to an extra copy of the gene (DRD$_{4rs}$). Thus, the ancestor of gnathostome vertebrates possessed a repertoire of five DRD$_2$ genes: DRD$_2$, DRD$_{2l}$, DRD$_3$, DRD$_4$ and DRD$_{4rs}$. 

Figure 10 (on next page)
Common ancestor of vertebrates

- DRD1
- DRD5
- DRD1C
- DRD1E

Common ancestor of gnathostomes

- DRD1
- DRD5
- DRD1C
- DRD1E

2 Whole Genome Duplications

- D1anc
- D2anc

- DRD2
- DRD2l
- DRD4
- DRD4rs
- DRD3
- DRD3rs