

Summary

In the vertebrate evolutionary tree, amphibians occupy an essential position as ancient tetrapods. They form a transitional group between swimming fishes and land-dwelling amniotes, and are therefore key representatives of the fin-to-limb transition. In addition, amphibians exhibit many unique features, such as very diverse body plans and limb types between frogs, salamanders and caecilians, or a wide range of genome sizes between and within the three orders. This work deals with the genomic aspects that may have contributed to these characteristics, by studying the evolution of two clustered gene families that are likely to have played a role in the diversification of amphibians. Both gene families, *Hox* genes and TIL protein-encoding genes, have a different evolutionary history. *Hox* genes are essential for body patterning in all animals (except sponges) and are therefore highly conserved. TIL protein-encoding genes on the other hand exist only in amphibians and are thought to play a role in host defense, either as toxin, or as part of the innate immune system.

First, a PCR survey for *Hox* genes involved in limb development was conducted in two species of the three amphibian orders, which confirmed the presence of four clusters in all amphibians. Moreover, not every gene that was expected based on the *Hox* complement of other vertebrates was recovered, suggesting that lineage-specific losses occurred in amphibians. Next, a BAC library of a caecilian was constructed in order to obtain the *HoxD* cluster for comparative analyses with other vertebrates, including the identification of conserved, non-coding sequences or putative regulatory elements. At least one regulatory element present in other vertebrates was lost in amphibians. Strikingly, the caecilian posterior *HoxD* cluster has experienced the loss of one gene and the accumulation of transposable elements and other repeats, resulting in a considerable expansion of the cluster. This may exert an influence on the long-range enhancer regulating the expression of the posterior *Hoxd* genes during limb development, and may be correlated with the elongation of the caecilian body and absence of limbs.

The TIL protein-encoding genes were identified in the skin of distantly related frog species and in a salamander. These genes encode a protein domain that is also found in serine protease inhibitors, venom and anticoagulant proteins of (blood-feeding) arthropods and nematodes. In vertebrates, this domain is present in large, structural, multi-domain proteins, but the TIL protein architecture only exists in amphibians. All but a few amphibian TIL protein-encoding genes form a single clade, together with trypsin inhibitors from *Bombina* frog species. This gene family is characterized by extensive gene duplications and rapid diversification, which may be related to their putative toxin or immune function. Each in a different way, both gene families have contributed to the evolutionary diversification of amphibians.

Samenvatting

In de evolutieve boom van de gewervelde dieren nemen de amfibieën een essentiële positie in als oude viervoeters. Ze vormen de overgang tussen de zwemmende vissen en de landbewonende amnioten, en daarom zijn ze de belangrijkste vertegenwoordigers van de transitie van vinnen naar poten. Daarenboven bezitten amfibieën een aantal unieke kenmerken, zoals een lichaamsplan en ledematen die verschillen tussen kikkers, salamanders en wormsalamanders, of een brede waaier aan genoomgroottes, zowel tussen als binnen de drie ordes. De genomische aspecten die mogelijk hebben bijgedragen tot deze kenmerken worden in dit werk behandeld door middel van de studie van de evolutie van twee geclusterde genfamilies. Beide genfamilies, *Hox*-genen en TIL-proteïne coderende genen, hebben mogelijk een rol gespeeld in de diversificatie van amfibieën, maar hebben een verschillende evolutieve geschiedenis. *Hox*-genen spelen een essentiële rol in de embryonale ontwikkeling van alle dieren (behalve sponzen) en zijn daardoor sterk geconserveerd. TIL-proteïne coderende genen daarentegen bestaan enkel in amfibieën, en ze zijn mogelijk betrokken bij de afweer, hetzij als toxine, hetzij als een onderdeel van het aangeboren immuunsysteem.

Eerst werd er in twee soorten van de drie amfibieënordes op basis van PCR gezocht naar de *Hox*-genen die betrokken zijn bij de ontwikkeling van ledematen. De resultaten hiervan bevestigden de aanwezigheid van vier *Hox*-clusters in alle amfibieën. Bovendien werd niet elk gen dat verwacht was op basis van het aantal *Hox*-genen in andere gewervelden teruggevonden, wat erop wijst dat sommige genen zijn verloren gegaan specifiek in amfibieën. Vervolgens werd een BAC-genoombank van een wormsalamander geconstrueerd om de *HoxD*-cluster te bekomen voor vergelijkende analyses met andere gewervelden, waarbij ook geconserveerde, niet-coderende sequenties en dus potentiële regulatorische elementen werden geïdentificeerd. Ten minste één regulatorisch element dat aanwezig is in andere gewervelden is verloren gegaan in amfibieën. Maar opvallender is het verlies van een gen en de opeenstapeling van mobiele

elementen en andere *repeats* in de posterieure *HoxD*-cluster van de wormsalamander, waardoor de cluster aanzienlijk verlengd is. Dit heeft mogelijk een invloed op de *enhancer* die de expressie van de posterieure *Hoxd*-genen reguleert tijdens de ontwikkeling van de ledematen, en mogelijk is dit gecorreleerd met de verlenging van het wormsalamanderlichaam en de afwezigheid van ledematen.

TIL-proteïncoderende genen werden geïdentificeerd in de huid van ver verwante kikkersoorten en in een salamander. Deze genen coderen een proteïnedomein dat ook aanwezig is in serine protease-inhibitoren, gif en antistollingsproteïnen van (bloedzuigende) geleedpotigen en rondwormen. In gewervelde dieren is dit domein aanwezig in grote, structurele multidomeinproteïnen, maar de architectuur van TIL-proteïnen komt enkel voor in amfibieën. Nagenoeg alle TIL-proteïncoderende genen van amfibieën vormen een enkele clade, samen met trypsine-inhibitoren van *Bombina*-kikkersoorten. Deze genfamilie wordt gekenmerkt door extensieve genduplicaties en snelle diversificatie, wat mogelijk gerelateerd is aan hun vermeende toxine- of immuunfunctie. Op een verschillende manier hebben beide genfamilies bijgedragen tot de evolutionaire diversificatie van amfibieën.