

# Identification of Common Carp Innate Immune Genes with Whole-Genome Sequencing and RNA-Seq Data

Yanju Zhang<sup>1\*</sup>, Elia Stupka<sup>2</sup>, Christiaan V. Henkel<sup>3</sup>, Hans J. Jansen<sup>3</sup>, Herman P. Spaijk<sup>3,4</sup> and Fons J. Verbeek<sup>1\*</sup>

<sup>1</sup>Section Imaging and Bioinformatics, Leiden Institute of Advanced Computer Science (LIACS), Leiden University, The Netherlands

<sup>2</sup>Bioinformatics group, UCL Cancer Institute, London, UK

<sup>3</sup>ZF-screens B.V., Leiden, The Netherlands

<sup>4</sup>Institute of Biology Leiden, Leiden University, The Netherlands

## Summary

The common carp is a candidate model system for immunology research. Using next-generation sequencing technology, we have generated a huge amount of sequence reads from the carp genome and transcriptome. Currently, our aim is to identify carp genes involved in the development of the innate immune response, particularly TIR domain-containing genes, from a preliminary genome assembly. To achieve this, we developed a comprehensive gene identification pipeline. This analysis allowed us to estimate that the carp has 39 TIR domain-containing transcript isoforms and genes.

## 1 Introduction

Common carp (*Cyprinus carpio*) is one of the most important freshwater cultured fish species that has been widely used in fish biology research [1]. A single female is capable of producing up to a few hundred thousand eggs that can be efficiently fertilized *in vitro*. Since the innate immune response is already active in developing embryos, common carp can be a relevant model for studying its mechanisms. The innate immune response is the first line of defence against infectious diseases and cancer by identifying and killing pathogens and detrimental cells, and relies heavily on signalling by pattern recognition receptors. The best-studied pattern recognition receptors of the vertebrate innate immune system are the Toll-like receptors (TLRs). All the TLRs, some Interleukin receptors (IL-Rs) and downstream adaptor proteins contain a Toll/Interleukin-1 receptors (TIR) domain, a highly conserved functional unit mediating the protein-protein interactions between the receptors and the adaptors.

TIR domain-containing genes therefore play important roles in immunity signalling pathways. In zebrafish (*Danio rerio*), this gene-family has been studied using microarray technology [2]. However, microarrays have a number of shortcomings, i.e. low sensitivity and specificity, low consistency across platforms, and, above all, they rely on a fixed definition of the transcriptome for their design.

---

\* To whom correspondence should be addressed. Email: {yanju|fverbeek}@liacs.nl





















