Developmental and Comparative Immunology 35 (2011) 959–974

Contents lists available at ScienceDirect

Developmental and Comparative Immunology

journal homepage: www.elsevier.com/locate/dci

Review

Invertebrate immune diversity

Julie Ghosh [1], Cheng Man Lun [1], Audrey J. Majeske [1], Sandro Sacchi [1], Catherine S. Schrankel [1,2], L. Courtney Smith [*,1]

Department of Biological Sciences, George Washington University, Washington, DC United States

ARTICLE INFO

Article history:
Available online 21 December 2010

Keywords:
Penaeidins
Crustins
FREPs
Fuhc
Fester
Uncle fester
Dscam
PGRP
VCBP
Sp185/333
Innate immunity
Immune diversification
Gene family evolution

ABSTRACT

The arms race between hosts and pathogens (and other non-self) drives the molecular diversification of immune response genes in the host. Over long periods of evolutionary time, many different defense strategies have been employed by a wide variety of invertebrates. We review here penaeidins and crustins in crustaceans, the allorecognition system encoded by *fuhc*, *fester* and *Uncle fester* in a colonial tunicate, Dscam and PGRPs in arthropods, FREPs in snails, VCBPs in protochordates, and the *Sp185/333* system in the purple sea urchin. Comparisons among immune systems, including those reviewed here have not identified an immune specific regulatory "genetic toolkit", however, repeatedly identified sequences (or "building materials" on which the tools act) are present in a broad range of immune systems. These include a Toll/TLR system, a primitive complement system, an LPS binding protein, and a RAG core/Transib element. Repeatedly identified domains and motifs that function in immune proteins include NACHT, LRR, Ig, death, TIR, lectin domains, and a thioester motif. In addition, there are repeatedly identified mechanisms (or "construction methods") that generate sequence diversity in genes with immune function. These include genomic instability, duplications and/or deletions of sequences and the generation of clusters of similar genes or exons that appear as families, gene recombination, gene conversion, retrotransposition, alternative splicing, multiple alleles for single copy genes, and RNA editing. These commonly employed "materials and methods" for building and maintaining an effective immune system that might have been part of that ancestral system appear now as a fragmented and likely incomplete set, likely due to the rapid evolutionary change (or loss) of host genes that are under pressure to keep pace with pathogen diversity.

Contents

Abbreviations: 4DSC, four-disulfide core; aa, amino acids; AMPs, antimicrobial peptides; BAC, bacterial artificial chromosome; CRD, cysteine-rich domain; EGF, epidermal growth factor; FBG, fibrinogen; FNIII, fibronectin type III; GNBP1, Gram-negative binding protein 1; Ig, immunoglobulin; IgSF, immunoglobulin superfamily; indels, insertions/deletions; LPS, lipopolysaccharide; NACHT, NAIP/CIITA/HET-E/TP1; NITR, novel immune-type receptors; PAMPs, pathogen-associated molecular patterns; PGNs, peptidoglycans; PRD, proline-rich domain; PRR, pattern recognition receptor; *R*, resistance; RGD, arginine-glycine-aspartic acid; SCR, short consensus repeat; SEPs, secretory-excretory products; SNPs, single nucleotide polymorphisms; TCRs, T-cell receptors; TEPs, thioester-containing proteins; TIR, Toll/IL-1 receptor; TM, transmembrane; WAP, whey acidic protein.

* Corresponding author at: Tel.: +1 202 994 9211; fax: +1 202 994 6100.
E-mail address: csmith@gwu.edu (L.C. Smith).

[1] All authors contributed equally and are listed in alphabetical order.

[2] Current address: Department of Immunology, University of Toronto, Toronto Canada.

0145-305X/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.dci.2010.12.009

1. Introduction

Host immunity, by its very nature, is an ongoing arms race against invading pathogens, with the onus on pathogens to outwit host immunity, and on the host to successfully eliminate invaders. This long-term host-pathogen co-evolutionary history involves high rates of mutation and/or variation in microbes that generally have short generation times versus long-lived hosts with corresponding low mutation rates [1], which encounter and must defend against many different pathogen types. On the host side, animals employ pattern recognition receptors (PRRs) that generically recognize the pathogen-associated molecular patterns (PAMPs) of the important classes of microbes, including lipopolysaccharide (LPS), peptidoglycans (PGNs), double stranded RNA, and β-glucans among others. On stimulation, these PRRs activate downstream signaling pathways that result in either the release of antimicrobial molecules, such as those known for *Drosophila* [2], and/or the stimulation of a cell-mediated phagocytic response [3]. Although these host responses are broadly successful, pathogens have evolved adaptations, such as altered surface molecules [4] to enable them to evade detection or immune attack. Consequently, for survival, the host requires a repertoire of receptor and effector molecules that is capable of detecting foreign molecules more diverse than basic PAMPs, and of mounting an effective response. Because the precise pathogen adaptations that are encountered are necessarily unpredictable, the best option for the host – and the one that is characteristic of immune systems across the animal kingdom – is to find ways to generate random or near random diversification and expansion of immune receptors so that the largest possible range of pathogens, regardless of the type of evasive adaptation evolved, will be detectable by the immune system as foreign.

There are many different strategies among animals to expand the repertoire of immune responsiveness. Although the somatic rearrangements involved in diversification of the T cell receptor genes (TCRs) and immunoglobulins (Igs) in mammals are well known and the assembly of the variable lymphocyte receptor (VLR) genes in the agnathans [5] is an intriguing alternative solution to immune diversification, these mechanisms are not observed outside of the vertebrates. Invertebrates employ other approaches that include (i) the presence of large gene families either within individuals or within populations that encode a wide variety of protein isoforms, (ii) genomic instability within the families of similar genes that promote unequal crossovers, gene conversion, gene duplication/deletion and paralogous mispairing, all of which promote sequence diversification, (iii) a variety of modifications to mRNAs including alternative splicing, RNA editing and low fidelity RNA polymerases, and (iv) a broad array of modifications to proteins either during or after translation. Several examples of immune genes and gene families in invertebrates that show sequence diversification are described below.

2. Antimicrobial peptides (AMPs) in crustaceans

AMPs are found in a wide variety of living organisms, including bacteria, fungi, plants, and animals, and are an important aspect of the innate immune response [6]. Most AMPs are known to be immunomodulators that are active against Gram-negative and Gram-positive bacteria, yeast, fungi, parasites, enveloped viruses and tumor cells, and some AMPs kill pathogens directly *in vitro* [6–9]. AMPs are typically small cationic amphipathic molecules that range in size from 15 to 200 amino acids (aa) in length, but are rarely larger than 30 kDa [10]. They are classified into three major groups based on aa sequence, secondary structure and functional properties. There has been particular interest in identifying AMPs in shrimp because of a growing number of diseases that affect this economically important group, and the notion that understanding shrimp AMPs may lead to therapeutic applications to curb the loss of shrimp production from infections [11]. Two major shrimp AMPs are the cysteine-rich penaeidins and crustins. Each shows sequence diversity and is comprised of multiple classes (penaeidins) and types (crustins) that are synthesized mostly in hemocytes and are released into the hemolymph in response to infection [6].

2.1. Penaeidins

There are four classes of penaeidins in penaeid shrimp and each class has several isoforms (Fig. 1) [12–14]. Penaeidins are small peptides of 5–7 kDa with an N-terminal signal peptide region followed by a proline-rich domain (PRD) and C-terminal cysteine-rich domain (CRD) containing six cysteine residues ([15,16]; see PenBase http://www.penbase.immunaqua.com for all (>200) penaeidins). The N-terminal PRD is longer than the CRD and is free of disulfide bonds, thus making it less rigid, whereas the C-terminal CRD is more conserved across classes and is stabilized by three disulfide bonds [15,17,18]. The sequence diversity within the PRD among different penaeidin classes is likely the source of variation in anti-microbial responses [17,19] based on evidence that the CRD