Gene 552 (2014) 116–125

Contents lists available at ScienceDirect

Gene

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

Annotation of the *Daphnia magna* nuclear receptors: Comparison to *Daphnia pulex*

Elizabeth J. Litoff [a], Travis E. Garriott [a], Gautam K. Ginjupalli [b], LaToya Butler [a], Claudy Gay [a], Kiandra Scott [a], William S. Baldwin [a,b,*]

[a] Biological Sciences, Clemson University, United States
[b] Environmental Toxicology Program, Clemson University, United States

ARTICLE INFO

Article history:
Received 24 April 2014
Received in revised form 9 September 2014
Accepted 13 September 2014
Available online 17 September 2014

Keywords:
Phylogenetics
Crustacea
Arthropod
Evolution

ABSTRACT

Most nuclear receptors (NRs) are ligand-dependent transcription factors crucial in homeostatic physiological responses or environmental responses. We annotated the *Daphnia magna* NRs and compared them to *Daphnia pulex* and other species, primarily through phylogenetic analysis. *Daphnia* species contain 26 NRs spanning all seven gene subfamilies. Thirteen of the 26 receptors found in *Daphnia* species phylogenetically segregate into the NR1 subfamily, primarily involved in energy metabolism and resource allocation. Some of the *Daphnia* NRs, such as RXR, HR96, and E75 show strong conservation between *D. magna* and *D. pulex*. Other receptors, such as EcRb, THRL-11 and RARL-10 have diverged considerably and therefore may show different functions in the two species. Curiously, there is an inverse association between the number of NR splice variants and conservation of the LBD. Overall, *D. pulex* and *D. magna* possess the same NRs; however not all of the NRs demonstrate high conservation indicating the potential for a divergence of function.

1. Introduction

Daphnia species are an important aquatic bioindicator species and invertebrate toxicology model (Heckmann et al., 2008). Daphnids are commonly studied zooplankton because of their importance to aquatic ecosystems, ability to contend with environmental challenges, amenability to culture, short life-cycle, and parthenogenic reproduction (Baldwin and LeBlanc, 1994; Thomson et al., 2009). They occupy a wide array of environments with habitats ranging drastically in size, permanence, salinity, nutrient levels, and UV exposure (Colbourne et al., 1997). *Daphnia magna* are one of the most commonly used test species for aquatic toxicity tests. The *Daphnia pulex* genome has already been fully sequenced (Colbourne et al., 2011), and a concerted effort is in progress to sequence the highly studied related cladoceran, *D. magna* (Lehman et al., 1995). These genomic models will offer a way of interpreting molecular modifications as well as convergence of adaptive traits associated with specific habitats that vary between the different species of daphnids (Shaw et al., 2008).

The regulation of physiological pathways that maintain proper metabolism and homeostasis within higher organisms continues to be a key concept in biological research. In order to maintain homeostasis cells must be able to acclimate to external and internal cues such as xenobiotics, nutrients, hormones, and other environmental cues. Nuclear receptors (NRs) are a key set of transcription factors that induce acclimation and maintain homeostasis by responding to chemical cues. Because NRs are considered so important in physiology they have been called "the Rosetta stone of physiology" (Evans, 2005). Once activated, NRs translocate to the nucleus (if not already found in the nucleus), bind DNA at specific response elements and initiate transcription. This provides for transcriptional regulation of specific proteins involved in a vast array of diverse physiological functions such as reproduction, embryonic development, cell differentiation, resource allocation, and the maintenance of homeostasis (Chawla et al., 2001; King-Jones and Thummel, 2005; Kretschmer and Baldwin, 2005).

Most NRs consist of five modules: A/B, C, D, E and F. The C module serves as the DNA-binding domain (DBD) and is responsible for binding to the response element on a target gene and is highly conserved among orthologs of different species (Hernandez et al., 2009). The A/B module binds to coactivators. The D subunit constitutes the hinge region and often controls nuclear translocation once the receptor is activated by a ligand. The E module, which is moderately conserved among orthologs of different species, is the ligand-binding domain (LBD) that controls ligand-mediated activation of the nuclear receptor (Hernandez et al.,

Abbreviations: NR, nuclear receptor; CDD, conserved domain database; DBD, DNA binding Domain; LBD, Ligand binding Domain; Pfam, Protein family; ML, Maximum Likelihood; EcR, Ecdysone Receptor; KNR, Knirp; THRL, Thyroid hormone receptor-like; RARL, Retinoid Acid Receptor-like; PPAR, Peroxisome Proliferator Activated Receptor; HR3, Hormone Receptor 3; HR4, Hormone Receptor 4; HR38, Hormone Receptor 38; HR39, Hormone Receptor 39; HR78, Hormone Receptor 78; HR96, Hormone Receptor 96; HR97, Hormone Receptor 97; HNF4, Hepatocyte Nuclear Factor 4; NR2E6, Nuclear Receptor 2E6; RXR, Retinoid X Receptor; DSF, Dissatisfaction; PNR, Photoreceptor cell-specific nuclear receptor; TLL, Tailless; SVP, Seven-up; ERR, Estrogen Related Receptor; FTZF1, Fushi Tarazu Factor 1.

* Corresponding author at: Clemson University, Biological Sciences, 132 Long Hall, Clemson, SC 29634, United States.
E-mail address: Baldwin@clemson.edu (W.S. Baldwin).

http://dx.doi.org/10.1016/j.gene.2014.09.024
0378-1119/

2009). The exact function of the last subunit, F, is currently unknown. The NR superfamily is categorized into seven main subfamilies based on their structural similarities (Committee, 1999).

Some NRs are highly specific and the LBD will only bind to a select molecule or group of molecules; others are much more promiscuous. Promiscuous NRs bind to a wide range of different molecules and, depending on the molecule, activate the transcription of a wide range of proteins. Some of these promiscuous NRs are involved in inducing phase I–III responses following exposure to toxicants (Kliewer et al., 1998; Kawamoto et al., 1999; Wei et al., 2000; King-Jones et al., 2006; Karimullina et al., 2012). It has been hypothesized that specificity/promiscuity comes into play when examining the evolution of nuclear receptors due to natural selection. In the most primitive species, the genomes contain significantly fewer NRs. Either there are less NR-mediated pathways to regulate or NRs regulate more pathways by responding to more ligands in primitive species. Recently it has been postulated that with less nuclear receptors, the responsibilities of each individual receptor increase and must be able to activate a larger array of pathways (Bridgham et al., 2010; Eick et al., 2012). As evolution of species has progressed, the quantity of nuclear receptors has grown and the receptors have become increasingly specific. This observation lends itself to the idea of specificity through selection in order to make the most optimal form of the receptor (Eick et al., 2012). Through selection, new specialized nuclear receptors seem to have evolved based on the greater efficiency to carry out pathways as well as increase the regulation of these pathways in response to specific cues. Interestingly, the *D. pulex* genome has nearly half the receptors of humans, but more genes to regulate (Colbourne et al., 2011).

Many NRs are a conduit between internal and external environmental conditions. For example, chronic stress that may be physical or emotional increases adrenocorticotropic hormone and glucocorticoid release, and in turn glucocorticoid receptor (GR) activity. The GR responds by regulating behavior, the immune system, metabolism, growth, and reproduction (O'Connor et al., 2014). Overall, NRs induce the proper physiological responses by responding to chemical cues and transcriptionally regulating pathways that help individuals respond to current conditions.

The purpose of this study was to annotate the *D. magna* NRs and compare them to *D. pulex* and other species, primarily through phylogenetic analysis. Because daphnid species are globally distributed zooplankton, we chose to annotate and compare the NRs of two common model species, *D. magna* and *D. pulex* with the hope of providing insight into how their respective NRs evolved while under different habitats and external pressures.

2. Materials and methods

2.1. Identification and genomic characterization of D. magna nuclear receptors

Identification of *D. magna* NRs was performed using a Basic Local Alignment Search Tool (BLAST) (Altschul et al., 1990) with each of the previously annotated 25 NRs from *D. pulex* (Thomson et al., 2009) against the assembled *D. magna* genome (http://server7.wfleabase.org:8091/gbrowse/cgi-bin/gbrowse/daphnia_magna2/). Each of the BLAST search hits was compared to the NCBI database using BLASTp to confirm its status as an NR. The DBD and the LBD of each receptor were identified at this time using the conserved domain database (CDD) (Marchler-Bauer et al., 2007). Protein family (pfam) Zf-C4 (pfam00105) was used to identify the confines of the DBD, and Hormone recep (pfam00104) was used to identify the confines of the LBD to maintain consistency in domain identification. The DBD and LBDs were compared between *D. magna* and *D. pulex* using clustalw (http://www.ebi.ac.uk/Tools/clustalw2/) and percent identity reported. In addition, the DBD and LBD sequences were used during phylogenetic analysis (see below).

2.2. Gene structure

D. magna homologs identified were mapped to the *D. magna* genome project's browser (http://server7.wfleabase.org/genome/Daphnia_magna_prerelease/) in order to identify the gene structure (position, length, exons, introns, and intron phase). After protein translation, translation start and stop sites were determined, and protein-coding exons estimated from the gene models and number of nucleotides determined in each intron and exon.

2.3. Phylogenetics

Phylogenetic analysis was performed using analysis methods described previously (Thomson et al., 2009; Hannas et al., 2010). All non-daphnid sequences used for phylogenetic analysis were derived from the NCBI database. The *D. pulex* sequences are predicted protein sequences from the fleabase dataset, and the *D. magna* sequences are predicted protein sequences from the *D. magna* genome browser. *D. magna* was compared to *D. pulex* and to nuclear receptors from other species available in GenBank such as *Drosophila melanogaster*, *Homo sapiens*, *Ciona intestinalis*, *Ixodes scapularis*, *Nasonia vitripennis*, *Bombus terrestris*, *Apis mellifera*, *Aedes aegypti*, *Metaseiulus occidentalis* and *Caenorhabditis elegans* (Additional file 1).

Phylogenetic analysis was performed using only the highly conserved DBD and moderately conserved LBD of each receptor. These domains were identified using the conserved domain database CDD (Marchler-Bauer et al., 2007). Zf-C4 (pfam00105) was used to identify the boundaries of the DBD, and Hormone recep (pfam00104) was used to identify the boundaries of the LBD of each receptor. ClustalX default parameters were used to align the domains (Thompson et al., 1997). Trees were constructed using Bayesian Inference (BI) with MrBayes software version 3.1.2 (Ronquist and Huelsenbeck, 2003) on Bioportal (Kumar et al., 2009). Phylogenetic trees were constructed using the "mixed-model" approach in which the Markov chain Monte Carlo sampler explores nine different fixed-rate amino acid substitution models implemented in MrBayes. We used 4 chains with runs of 5 million generations, chains sampled every 100 generations, and a burnin of 10,000 trees with the *WAG* model (Whelan and Goldman, 2001). The *C. elegans* NHR-1 receptor was used as the outgroup.

Maximum parsimony and distance parameters were used to provide additional support for the phylogenetic relationships observed. Distance parameters were measured using PAUP 4.0b10 with default characteristics (mean character difference and among site rate variation), and full heuristic searches. Branch support was measured by bootstrap analysis with 1000 replicates. Parsimony was constructed using PAUP version 4.0b10 with heuristic searches, tree-bisection-reconnection, topological constraints not enforced, and multiple tree option in effect with an initial maximum tree setting at 100,000. Branch support was measured by bootstrapping with 10,000 replicates. Trees were visualized with FigTree (http://tree.bio.ed.ac.uk/software).

Phylogenetic analysis was also confirmed with Maximum Likelihood (ML) using MEGA 6.0 (Tamura et al., 2013). "Find Best Model" was used to determine the parameters for Maximum Likelihood. In turn, the analysis was performed using the Bootstrap method with 500 replications, and the LG model was used with Gamma distributed rates among sites (2). Tree inference options included SPR level 3, BIONJ with a very strong branch filter. For consistency with BI, a WAG model with gamma distributed rates among sites was also attempted. This model showed very little difference when compared to the LG model (data not shown).

3. Results and discussion

3.1. Nuclear receptor groups in Daphnia

Analysis of the *D. magna* genome found 26 NRs. Previously, 25 NRs were found in *D. pulex*. The *D. magna* BLAST searches found an additional