Biochimie xxx (2015) 1–11

Contents lists available at ScienceDirect

Biochimie

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

Review

Infectious long non-coding RNAs

Konstantina Katsarou [a], A.L.N. Rao [b], Mina Tsagris [c], Kriton Kalantidis [a, c, *]

[a] Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology-Hellas, P.O. Box 1527, GR-71110, Heraklion, Crete, Greece
[b] Department of Plant Pathology and Microbiology, University of California, Riverside, CA, 92521-01222, USA
[c] Department of Biology, University of Crete, Heraklion, Crete, Greece

ARTICLE INFO

Article history:
Received 5 January 2015
Accepted 7 May 2015
Available online xxx

Keywords:
Viroids
RNA satellites
Hepatitis D virus
lncRNA

ABSTRACT

Long non protein coding RNAs (lncRNAs) constitute a large category of the RNA world, able to regulate different biological processes. In this review we are focusing on infectious lncRNAs, their classification, pathogenesis and impact on the infected organisms. Here they are presented in two separate groups: 'dependent lncRNAs' (comprising satellites RNA, Hepatitis D virus and lncRNAs of viral origin) which need a helper virus and 'independent lncRNAs' (viroids) that can self-replicate. Even though these lncRNA do not encode any protein, their structure and/or sequence comprise all the necessary information to drive specific interactions with host factors and regulate several cellular functions. These new data that have emerged during the last few years concerning lncRNAs modify the way we understand molecular biology's 'central dogma' and give new perspectives for applications and potential therapeutic strategies.

1. Introduction

In 1956 Francis Crick developed the 'central dogma' of molecular biology in which genetic information flows in a unidirectional way from DNA to mRNA and then to protein [1]. However, in the past three decades distinct types of non protein coding RNAs (ncRNA) of different sizes and shapes have been reported making this landmarking theory incomplete to some extent. These ncRNAs have been implicated in regulation of different biological and physiological processes. Based on their size they can be separated into three families: 1) short ncRNA (between 17 and 30 nt of length), 2) middle-size ncRNAs varying between 30 and 200 nt and finally 3) long ncRNAs (lncRNA), with RNAs over 200 nt (Fig. 1). lncRNAs may actually be more important than initially thought. In fact, in the GENCODE 21 release of 2014, it is suggested that the human genome contains at least 15.877 lncRNA genes and 9.534 short ncRNA genes (information in Ref. [2] and http://www.gencodegenes.org/). Their functional role remains mostly elusive, however there are cases implicating lncRNAs in gene expression (at transcriptional, post-transcriptional and epigenetic level) and thus numerous cellular processes [3].

Since the lncRNA field is relatively new, so far it has not been clearly categorized, affecting results interpretation. lncRNAs can be separated on the basis of their production pathway, way of action, overall structure, or influence on the organism (Fig. 1). Starting from the first, depending on their production it has been proposed to categorize lncRNAs as intergenic lncRNA, intronic lncRNA, sense lncRNAs, antisense lncRNA and sense-overlapping lncRNA [3,4]. Depending on their action they can be divided into *cis*-acting lncRNA controlling the expression of genes neighboring their transcription start site, or *trans*-acting lncRNA producing an activation or repression of genes at distant loci (reviewed in Ref. [5]). They can also act as competing endogenous RNAs (ceRNA). These particular lncRNAs are known as 'miRNA sponges' and their suggested role is to protect from degradation specific mRNAs, by targeting and sequestrating miRNAs, resulting in increase translation of these mRNAs [6]. lncRNAs can also be categorized by their shape, into linear/structured (lincRNA) and circular (circRNA). The first group comprises RNAs like ribosomal RNAs and RNAse P, important RNAs for proper cellular function. Endogenous circRNAs were reported for the first time in HeLa cells in 1979, but at the time, were considered as cryptic viral RNAs [7]. The first circRNA of exogenous origin was viroid, whose circularity was confirmed by electron microscopy in 1973 [8–10]. Till recently, circRNAs were considered to be a peculiarity of the viral world. However in recent years, with the advent of new sequencing technologies, numerous circRNAs have been discovered in several cells and tissues. Till today, 2000

* Corresponding author. Department of Biology, University of Crete, Heraklion, Crete, Greece.
E-mail address: kriton@imbb.forth.gr (K. Kalantidis).

http://dx.doi.org/10.1016/j.biochi.2015.05.005
0300-9084/

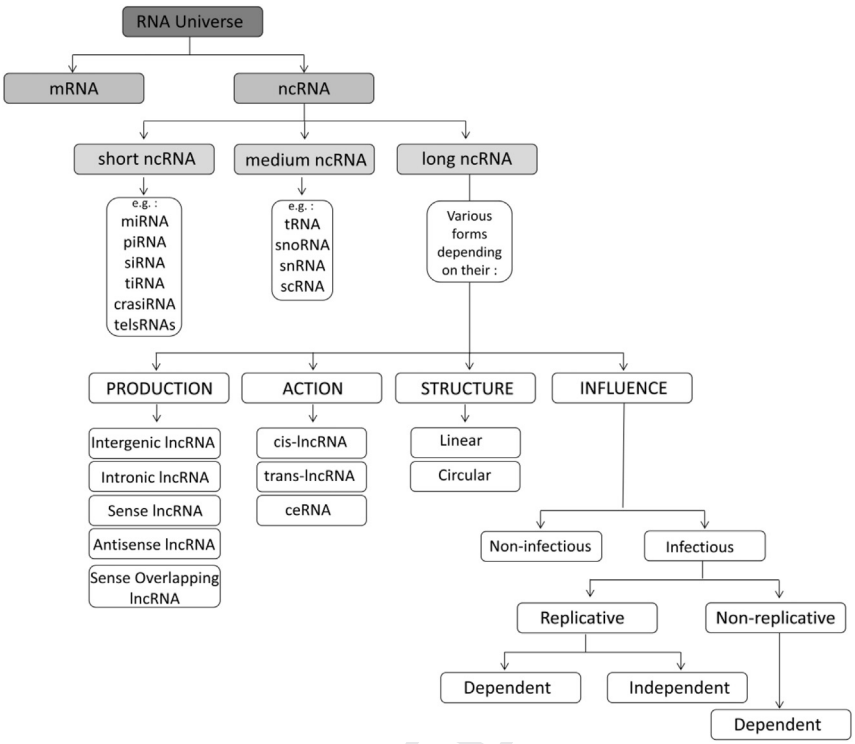

Fig. 1. Classification of long non-coding RNAs. Abbreviations: mRNA: messenger RNAs, ncRNA: non-coding RNA, miRNA: micro RNA, piRNA: piwi RNA, siRNA: small interfering RNA, tiRNA: transcription initiation RNA, crasiRNA: Centromere repeat associated small interacting RNA, telsRNAs: telomere-specific small RNA, tRNA: transfert RNA, snoRNA: small nucleolar RNA, snRNA: small nuclear RNA, scRNA: small cytoplasmic RNA, cis-lncRNA: cis-acting long non coding RNA, trans-lncRNA: trans-acting long non coding RNA, ceRNA: competing endogenous RNA.

circRNAs in humans, 1900 in mice, 700 in nematodes, and a significant number in other organisms have been identified, raising questions about their biological significance [11–13].

Finally, we propose to divide lncRNAs into non-infectious and infectious (Fig. 1). Infectious lncRNAs in their turn can be divided into non-replicative and replicative, since some of these lncRNAs are able to reproduce. This last category comprises infectious entities that can be separate into dependent and independent on the basis of their ability to self-replicate (viroids) or use components of a helper virus (satellite RNAs, HDV) for their infectivity. This review will focus on *infectious lncRNAs* (viroids, satellite RNA, *Hepatitis D virus)* but also some lncRNAs of viral origin, and discuss important aspects of their infectivity and common actions.

2. Independent infectious lncRNA

2.1. Viroids

- General aspects

Viroids are naked, circular, long non-coding RNA pathogens that range in size from 246 to 467 nt, causing plant diseases of economic importance [14]. They were reported for the first time in 1971 in the spindle tuber disease of potato [8]. Since then, thirty-two different viroid species have been identified and separated into two families (*Pospiviroidae* and *Avsunviroidae*) and eight genera (*Avsunviroid*, *Pelamoviroid*, *Elaviroid* for *Avsunviroidae* and *Pospiviroid*, *Hostuviroid*, *Cocadviroid*, *Apscaviroid, Coleviroid* for *Pospiviroidae*) [15,16]. However, new viroids have been identified (ex. *Dahlia latent viroid*) and await approval by the International committee on taxonomy of viruses (ICTV) [17]. Since their first discovery, viroid infections have been reported in every continent with different economic impact depending on their host and the local phytosanitary measures [14,18]. In the 2014 quarantine pest list of the European and Mediterranean plant protection organization (EPPO) there are three viroids of the *Pospiviroidae* family present: *Coconut cadang–cadang viroid*, *Chrysanthemun stunt viroid* and *Potato tuber spindle viroid* (PSTVd) and one, *Tomato apical stunt viroid*, in the alert list [19].

Viroid symptoms in host plants vary depending on their genomic RNA sequence/structure, the host and the environment. They range from mild to very severe, and affect either the entire plant or different organs such as leaves, fruits, flowers, roots and reserve organs [14]. Hosts are herbaceous, crop and woody plant species, as well as ornamentals [20,21]. Some of these hosts can act as reservoirs for potential future infections, imposing selective evolutionary pressure. The most recent example is the Potato Spindle Tuber Viroid case (PSTVd-family: *Pospiviroidae*) which was found mostly in asymptomatic ornamental plants (*Solanaceae*, *Scrophulariaceae* and *Asteracae* families) and nevertheless was able to infect tomato and potato [22,23]. However, viroid adaptation in these hosts created minor sequence or structural changes that leaded to dramatic effects on symptom expression [24–27].

Viroid spread is attributed to mechanical means and is facilitated by harvesting and cultural operations. However, transmission by seed, pollen, aphids and bumble bees have been proposed, with the last two remaining controversial [28,29]. However, an additional transmission mode is plausible, since it has been proposed that upon co-infection of PSTVd with *Potato leafroll virus* (family: *Luteoviridae*, genus: *Polerovirus*) transencapsidation of the viroid is possible, transforming temporarily the viroid into a 'virus', facilitating its spread and reducing the possibilities to control it [30,31].

- Molecular aspects

Pospiviroidae members, contains a rod-like or quasi rod-like secondary structure with five distinct domains, C (central), P

Please cite this article in press as: K. Katsarou, et al., Infectious long non-coding RNAs, Biochimie (2015), http://dx.doi.org/10.1016/j.biochi.2015.05.005