Eukaryotic genomes contain numerous copies of endogenous viral elements (EVEs), most of which are considered endogenous retrovirus (ERV) sequences. Over the past decade, non-retroviral endogenous viral elements (nrEVEs) derived from ancient RNA viruses have been discovered. Several functions have been proposed for these elements, including antiviral defense. This review summarizes the current understanding of nrEVEs derived from RNA viruses, particularly endogenous bornavirus-like elements (EBLs) and endogenous filovirus-like elements (EFLs). EBLs are one of the most extensively studied nrEVEs. The EBL derived from bornavirus nucleoprotein (EBLN) is thought to function as a non-coding RNA or protein that regulates host gene expression or inhibits virus propagation. Ebolavirus and marburgvirus, which are filoviruses, induce severe hemorrhagic fever in humans and nonhuman primates. Although the ecology of filoviruses remains unclear, bats are believed to be potential reservoirs. Based on the knowledge from EBLs, it is postulated that EFLs in the bat genome help to maintain the balance between filovirus infection and the bat’s defense system, which may partially explain why bats act as potential reservoirs. Further research into the functions of nrEVEs could reveal novel antiviral systems and inspire novel antiviral approaches.

Key words: EVE, nrEVE, bornavirus, filovirus, antiviral
sequences have the potential to cis-regulate gene expression [10, 11]. For example, LTRs are the predominant promoters of IL2RB (interleukin-2 receptor B) and NOS3 (nitric oxide synthase 3) in the placenta [12]. In addition, some ERV-derived RNAs are known to play a role as functional non-coding RNAs (ncRNAs). For example, long ncRNAs (lncRNAs) derived from human endogenous retrovirus H are essential for human embryonic stem cell identity [13].

In 2010, Horie and Honda et al. [14] discovered another type of viral elements related to a non-retroviral RNA virus, i.e., endogenous bornavirus-like elements (EBLs), in the genomes of several mammalian species, including humans, non-human primates, rodents, and elephants (Fig. 1). After the first report of EBLs, non-retroviral endogenous viral elements (nrEVEs) derived from a variety of viral families, including filoviruses [15-17], paroviruses [18, 19], circoviruses [19], and hepadnaviruses [20, 21], were identified in vertebrate and invertebrate genomes. EBLs appear to have originated from reverse transcription and integration of ancient bornavirus mRNA into the genomes of anthropoids (Fig. 1, (1) and (2)). Germline insertions of bornavirus sequences have been inherited by descendant hosts thereafter, some of which may exert biological functions, such as antiviral responses. During host evolution, animals with EBLs have been selected by surrounding conditions, such as infection with ancient bornavirus-related viruses (Fig. 1, (3)).

This review summarizes what is currently known about EBLs to provide an overview of nrEVE functions, because EBLs are one of the most extensively studied nrEVEs. We then discuss the potential roles of endogenous filovirus-like elements (EFLs) in filovirus ecology. Finally, we suggest possible application of nrEVE function to novel antiviral approaches.

EBLs Provide Concepts How nrEVEs Exert Their Functions

Borna disease virus 1 (BoDV-1) is an enveloped, 90-130 nm spherical virus that contains a linear, negative-stranded RNA genome and belongs to the family Bornaviridae [22]. The genomic structure of BoDV-1 comprises five genes arranged sequentially in the following order: 3′-nucleoprotein (N)-small accessory protein (X)/phosphoprotein (P)-matrix protein (M)-glycoprotein (G)-large protein (L)-5′ (Fig. 2). The genomic RNA and the N, P, and L proteins form viral ribonucleoprotein (RNP) complexes, which are viral units for viral replication and transcription [23, 24]. Unlike most other RNA viruses, which replicate in the cytoplasm, BoDV-1 replicates in the nuclei of infected...
cells. Therefore, the nucleocytoplasmic trafficking of BoDV-1 proteins plays a critical role in the BoDV-1 replication cycle [25].

EBLs are derived from ancient bornavirus mRNAs, as described above (Figs. 1 and 2). EBLs derived from N, M, G, and L (EBLNs, EBLMs, EBLGs, and EBLLs, respectively) have been identified (Fig. 2) [26], with the EBLNs being the most extensively investigated to date. Several functions of EBLN-derived proteins have been proposed based on the investigations of EBLNs. An EBLN in the genome of the thirteen-lined ground squirrel (Ictidomys tridecemlineatus), itEBLN, encodes an open reading frame (ORF) having 77% amino acid sequence identity with the current BoDV-1 N [27]. The itEBLN protein has been shown to be associated with BoDV-1 RNPs and to inhibit BoDV-1 replication [28, 29]. A recent study found that the Homo sapiens EBLN-2 (hsEBLN-2) protein interacts with apoptosis-associated mitochondrial proteins, such as HCLS1-associated protein X-1 (HAX-1) and apoptosis-inducing factor mitochondria-associated 1 (AIFM), and is involved in cell viability [30]. Another study has reported on EBLN proteins located on the rough endoplasmic reticulum in African elephants, although the precise function of these proteins remains unclear [31].

While itEBLN retains an intact ORF homologous to the N protein, most EBLNs have lost cognate ORFs. There is no evidence of selection to maintain the ORF of EBLNs in primates [32]. Nonetheless, RNAs are expressed from hsEBLN (hsEBLN-1 to -7) in at least one tissue. These observations suggest that most EBLNs have either lost their function or have a function unrelated to protein-coding, such as that of ncRNAs [33]. The hsEBLN-1 transcript suppresses the gene expression of its neighboring COMMD3 (COMM Domain Containing 3) gene [34]. Because the COMMD3 gene encodes a protein that interacts with and inhibits the NF-kB pathway [35], the presence of EBLN in this locus may downregulate the expression of the COMMD3 gene, thereby potentiating the NF-kB pathway to counteract invading pathogens. Furthermore, some EBLNs in the rodent and primate genomes (e.g., Mus musculus EBLN (mmEBLN)-3 to -5) are located within P-element-induced wimpy testis (PIWI)-interacting RNA (piRNA)-generating loci, “piRNA clusters,” and are expressed as piRNAs [36]. EBLN piRNAs may confer the host antiviral activity against viruses related to bornaviruses.

Based on the above-mentioned functions of EBLNs, the potential roles of nrEVEs can be deduced by analogy (Fig. 3). nrEVEs containing coding sequences (CDSs) can be transcribed and translated into proteins with various functions, including antiviral defense (e.g., itEBLN [28]) and apoptosis regulation (e.g., hsEBLN-2 [30]). If nrEVEs do not contain CDSs, they may act as regulatory DNA elements or ncRNAs. nrEVE-derived ncRNAs play significant roles in the regulation of host gene expression (e.g., hsEBLN-1 to the neighboring COMMD3 gene [34]) and possibly in the antiviral response (e.g., mmEBLN-3 to -5 [36]). In the next section, we use this concept to propose a potential role of endogenous filovirus-like elements (EFLs) in filovirus ecology.

**EFLs May Play a Role in Filovirus Ecology**

Filoviruses, including ebolavirus and marburgvirus, are enveloped, filamentous viruses that contain a linear,
negative-stranded RNA genome (approximately 19 kb in length), and are classified in the family Filoviridae. The genomic structure of a filovirus comprises seven genes arranged sequentially in the following order: 3’-nucleoprotein (NP)-viral protein (VP) 35-VP40-glycoprotein-VP30-VP24-RNA polymerase (L)-5’ (Fig. 4) [37]. Filovirus NP forms nucleocapsid-like structures, whereas VP30 functions as a transcription activator [38-40]. They interact with viral RNA together with the polymerase complex consisting of VP35 and L proteins and form viral RNPs, viral replication units [37]. The L protein provides the RNA-dependent RNA polymerase activity of the complex and VP35 is a polymerase complex cofactor that affects the mode of RNA synthesis [37]. VP35 also functions as a type I interferon (IFN) antagonist by interrupting interferon regulatory factor 3 (IRF3) [41, 42]. The VP40 and VP24 proteins act as matrix proteins located on the inner side of the membrane [37]. VP40 plays a key role in the budding process, whereas VP24 is presumed to be involved in nucleocapsid formation and assembly [43]. Marburgvirus VP40 and ebolavirus VP24 also antagonize IFN [44-46]. The family Filoviridae contains six genera: Ebolavirus, Marburgvirus, Cuevavirus, Dianlovirus, Striavirus, and Thamnovirus [47]. Among these viruses, ebolavirus and marburgvirus cause severe hemorrhagic fever in humans and nonhuman primates. Outbreaks of filovirus infection in humans, i.e., Ebola virus disease and Marburg virus disease, frequently occur in Central Africa. The 2014-2016 Ebola virus outbreak in West African countries (e.g., Guinea, Liberia, and Sierra Leone) was the largest outbreak to date [48].

The ecology of filoviruses, including zoonotic transmission events followed by subsequent human-to-human transmission, is still largely unknown. Epidemiologic evidence suggests that bats are the primary natural reservoir [49-53]. Some insectivorous bats and Egyptian rousette bats (Rousettus aegyptiacus) are susceptible to filovirus infection and exhibit seroconversion without any symptoms [52, 54, 55]. Therefore, bats seem to have co-evolved with filoviruses to tolerate infection; they allow viral replication only to levels sufficient for transmission, concurrently

---

**Fig. 3** Categorization of the functions of non-retroviral endogenous viral elements (nrEVEs). nrEVEs containing coding sequences (CDSs) can be transcribed and translated into proteins. If nrEVEs do not contain CDSs, nrEVEs may act as either regulatory DNA elements or RNA elements. nrEVEs located within PIWI-interacting RNA (piRNA) loci produce piRNAs that may confer antiviral activity against related exogeneous viruses. Even if nrEVEs are located within non-piRNA loci, they may transcribe into functional non-coding RNA (ncRNA).

**Fig. 4** Schematic view of the filovirus genome, mRNAs, and endogenous filovirus-like elements (EFLs). The dashed lines connect viral mRNAs to homologous EFLs (i.e., EFLNP, EFL35, and EFLL) in the host genomes. The sizes of the rectangles indicating nrEVEs do not reflect their actual sizes.
mounting a subdued antiviral immune response that controls clinical disease by minimizing proinflammatory responses [56]. Several researchers have focused on the bat immune system to determine how bats acquire disease tolerance to filoviruses [57-61]. For example, bats have evolved unique mechanisms to limit virus-induced proinflammatory responses [57-62]. Since canonical immune systems in bats are partially suppressed but bats can control filovirus infection, bats are thought to be equipped with an as-yet-unknown non-canonical means of controlling viral replication [59, 60, 63].

EFLs derived from ancient filovirus NP, VP35, and L genes (EFLNPs, EFL35s, and EFLLs, respectively) have been identified in mammalian genomes, such as those of microbats, tarsiers, wallabies, and opossums (Fig. 4) [15-17]. Among these EFLs, bat genomes contain EFLNPs and EFL35s. Myotis bat EFL35 retains an intact ORF that potentially encodes a protein of ~280 amino acids homologous to VP35 [16]. Previous studies using Myotis lucifugus EFL35 (mlEFL35) have shown that the mlEFL35 protein retains the ability to inhibit IFN-β production [64, 65]. These results suggest that EFLs might also retain other VP35-like properties, such as incorporation into viral RNPs, and thereby suppress viral replication, similarly to the above-mentioned itEBLN function (see section 2). If this is the case, EFL-derived immunity might play roles in noncanonical means of establishing disease tolerance to filovirus infection [66, 67]. Therefore, future investigations of EFL functions could provide insights into how bats establish disease tolerance to filovirus infection and act as potential reservoirs.
Conclusions

nrEVEs have been reported to play important roles in host defenses. Virus infections generally induce canonical immune responses, such as the production of type I IFN and cytokines (Fig. 5A). Although an appropriate level of immunity is required to combat infectious agents, excessive responses lead to host damage. As we summarized in this review, some nrEVE proteins and RNAs might directly suppress the replication of nrEVE-related exogenous viruses (Fig. 5B). In addition, nrEVEs might function as regulatory DNA sequences of host gene expression that leads to an antiviral state (Fig. 5B).

Over the past decades, bats have been reported to be important reservoirs and vectors of zoonotic viruses causing emerging infectious diseases, including Ebola and Marburg virus diseases, severe acute respiratory syndrome (SARS), Nipah and Hendra viral infections, and rabies [68]. The etiological agent of coronavirus disease 2019 (COVID-19) [69], SARS coronavirus 2, shares many genetic similarities with bat-borne beta-coronavirus [70, 71], suggesting that bats play a key role as coronavirus reservoirs. A large number of nrEVEs have been discovered in bat genomes [15]. This may explain, at least in part, why bats act as reservoir hosts for many viruses. Thus, we can hypothesize that bats have developed a system to assimilate viral sequences into their genomes and renovate them to combat related viruses without enhancing their canonical immune systems. As a result, bats may readily acquire disease tolerance to various viruses. Further research will be needed to verify this hypothesis.

Finally, it is conceivable that activation of nrEVEs could lead to the enforcement of antiviral activity without excessive activation of canonical immunity that may damage the host (Fig. 5C). To develop such a novel approach for the antiviral state, ways to enhance the antiviral activity of nrEVEs, e.g., the way inducing nrEVE expression, should be explored.

Acknowledgments. This study was supported in part by JSPS KAKENHI (grant numbers JP19K07592 and JP22K06027 to H.O., and JP22K19436, JP18H02664 and JP21H02738 to T.H.) and grants from the Okayama Medical Foundation (H.O.), Okayama University Assignment of Research Support Staff (H.O.), Takeda Science Foundation (T.H.), Naito Foundation (T.H.), SEI Group CSR Foundation (T.H.), Ryobi Teien Memory Foundation (T.H.), and Suzuken Memorial Foundation (T.H.).

References


