# Analyses of non-coding RNAs generated from the Epstein–Barr virus W repeat region.

Walter N. Moss

Department of Molecular Biophysics and Biochemistry, Howard Hughes Medical Institute, Yale University School of Medicine, New Haven, Connecticut 06536, USA
walter.moss@yale.edu

**Keywords:** Epstein–Barr virus, EBV, RNA, non-coding RNA, RNA structure, stable intronic sequence

**Abstract.** *Introduction*: The Epstein–Barr virus (EBV) W repeats are transcribed during viral lytic reactivation, a highly oncogenic form of latency known as latency III, and a rare type of latency (Wp-restricted) that is observed in ~15% of endemic Burkitt's lymphomas. The W repeats encode an EBV protein, EBNA-LP, whose message is produced by splicing out a short (81 nt) and long (2791 nt) intron. We previously discovered a stable intronic sequence (sis)RNA (ebv-sisRNA-1), which is the third most abundant EBV small RNA generated in latency III and comprises all 81 nt of the short intron. We also previously found that regions within the long intron form stable and conserved RNA structures, including a massive (586 nt) hairpin loop.

*Methods*: To identify RNAs that potentially interact with ebv-sisRNA-1, *in-silico* bimolecular RNA folding and sequence comparison were performed, examining all human and EBV micro (mi)RNAs from the miRBase database. RNA-Seq data from human B cells expressing a latency III program were re-analyzed to identify novel transcripts corresponding to the W repeats. RNA structure modeling was performed using a fold-and-align strategy coupled to manual comparative sequence analysis.

*Results*: ebv-sisRNA-1 is not likely to form conserved stable hybrids with EBV miRNAs. It may, however, interact with several human miRNAs. Analysis of RNA-Seq data indicates that the entire long W repeat intron is also likely a sisRNA (ebv-sisRNA-2). Stable RNA structure, rather than being localized to particular regions, is likely to span the entire ebv-sisRNA-2 sequence. We propose a model for this conserved fold that is supported by consistent and compensatory (structure-preserving) mutations.

*Conclusions*: The EBV W repeats generate two structured latent sisRNAs. Ebv-sisRNA-1 is predicted to interact with host miRNAs. Ebv-sisRNA-2 is modeled to have wide-spread and exceptionally thermodynamically stable RNA structure that is evolutionarily conserved in related herpesviruses. These results suggest important functions for both sisRNAs in type III latency.

## 1 Introduction

Epstein–Barr virus (EBV) is a human herpesvirus (~170 kbp double-stranded DNA) that infects as many as 95% of adults [1]. Primary EBV infection causes infectious mononucleosis and results in lifelong latent infection of B lymphocytes. In ways that are not yet fully understood, latent EBV infections are associated with a number of human malignancies: e.g. in autoimmune disorders such as lupus [2] and rheumatoid arthritis [3] and in cancers such as Burkitt's [4] and Hodgkin's lymphomas [5], and nasopharyngeal carcinomas [6]. EBV latency proceeds via several distinct "programs" named latency III through 0, which express a diminishing set of proteins. Non-coding (nc)RNAs, however, are transcribed throughout latency, including subsets of as many as 50 viral micro (mi)RNAs [7, 8], a virus-encoded small nucleolar RNA (v-snoRNA1) [9], and the two Epstein–Barr virus-encoded small RNAs (EBER1/2; [10]). The functions of these RNAs are active areas of research. The EBERs, for example, are the most abundant RNAs in EBV-infected cells (~107 transcripts per cell; [10]), are conserved in sequence and structure between EBV and related herpesviruses [11], and associate with host proteins to form ribonucleoprotein (RNP) complexes [12-15], yet a precise function for these RNAs remains to be found.

Beyond these RNAs little was known about other EBV ncRNAs or the roles of RNA structure in EBV virology until recently. A survey for novel non-coding and structured RNAs was recently completed on EBV [11]. This analysis combined a bioinformatic scan of the EBV genome with RNA-Seq. The bioinformatics analysis utilized the program RNAz [16], which scans genome alignments for fragments that contain evolutionarily conserved and unusually thermodynamically stable RNA secondary structures. Conserved and thermodynamically stable RNAs are the hallmarks of functional ncRNAs [17] and play important roles in many viruses [18-21]: e.g. translation [22], transcript stability/abundance [23-25], alternative splicing [26-28]; and viruses, in general, appear to be enriched for RNA structure [18, 19, 29, 30]

which may serve a variety of functions. The survey of EBV predicted a high likelihood of conserved and stably-folded RNA being generated from ~30% of the EBV genome [16]. Homology searching, using the INFERNAL package [31, 32], revealed that much of this structure is also present in the closely-related Macacine herpesvirus 4 (MHV4, a.k.a rhesus lymphocryptovirus) and, in some cases could extend to more distant herpesviruses.

A particular "hot-spot" for RNA structure was found in intronic areas of a repetitive region of the EBV genome known as the W repeats. The W repeats are transcribed during lytic reactivation, in a rare type of latency (Wp-restricted latency) observed in ~15% of endemic Burkitt's lymphomas [33], and during a highly oncogenic type of latency (type III; [34]). W repeat transcripts occur at the 5' end of the primary transcript for the Epstein–Barr virus nuclear antigens (EBNAs) and are, themselves, spliced to generate EBNA-LP (Fig. 1). Splicing of the W repeat transcript releases long (2791 nt) and short (81 nt) introns. Interestingly, analysis of small RNA-Seq data from latency III expressing human B cells found that all 81 nt of the short intron are stable and highly abundant in latency III (the third most abundant EBV RNA after the

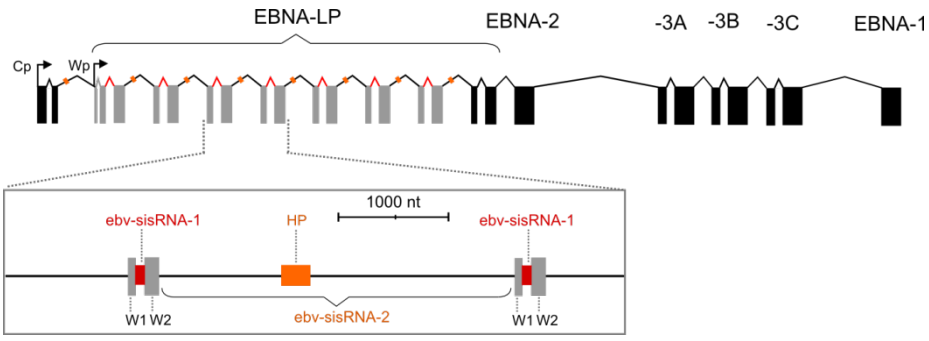



**Fig. 1.** Cartoon of the region encoding the Epstein–Barr virus nuclear antigens (EBNAs). The W repeats occur within the region encoding EBNA-LP. Transcription initiation occurs from an upstream C promoter (Cp) or W promoter (Wp) during latency III. Promoters are shown with bent arrows; introns are shown with crooked lines and exons with solid bars. The coding exons (W1 and W2) in the W repeats that encode EBNA-LP are shown with grey boxes. The short intron that generates ebv-sisRNA-1 is colored red; the long intron generating ebv-sisRNA-2 is shown in black with an orange box that represents the location of the long hairpin (HP).

EBERs; [11]). Thus this short intron is a stable intronic sequence (sis)RNA ([35]; ebv-sisRNA-1). In this current work we use bioinformatics to predict potential interactors of ebv-sisRNA-1, provide evidence for the long intron also being a sisRNA (ebv-sisRNA-2), and provide a structure model for ebv-sisRNA-2.

## 2 Methods

### 2.1 Prediction of ebv-sisRNA-1 miRNA interactions

All human and EBV miRNA sequences were obtained from the miRBase miRNA database [36]. Each mature miRNA sequence was co-folded vs. ebv-sisRNA-1 in silico, using the program RNAduplex [37] with the default temperature (37 °C) and energy model [38, 39]. A Perl script was written to filter out hybrids without perfect seed pairing between the miRNA and ebv-sisRNA-1. Results were further filtered to select miRNAs known to be expressed in B cells, the host for EBV. Hybrids were re-calculated with the program RNAstructure to check predictions and analyze potential suboptimal folds [40]. Conservation of base pairing was analyzed by comparing complementarity between the host miRNA sequence and ebv-sisRNA-1 homologs sequenced in herpesviruses infecting other primates: rhesus lymphocryptovirus, baboon herpesvirus, and pongine herpesviruses.

### 2.2 Analysis of ebv-sisRNA-2

Ribosome-depleted RNA-Seq data from latency III expressing (JY cell line) cells were aligned against the hg19 human genome and Akata strain EBV genome [41] simultaneously. Reads were aligned using bowtie2 [42]. Alignments were visualized using the Integrated Genomics Viewer (IGV) [43] after converting output files into BedGraph format and visually inspected to identify transcripts in the W repeat region.

Sequences for ebv-sisRNA-2 were obtained from EBV strains and related herpesviruses used in ref. [11]. Sequences were folded and aligned using LOCARNA-P [44]. LOCARNA-P, structure-based alignments were then used as input for the

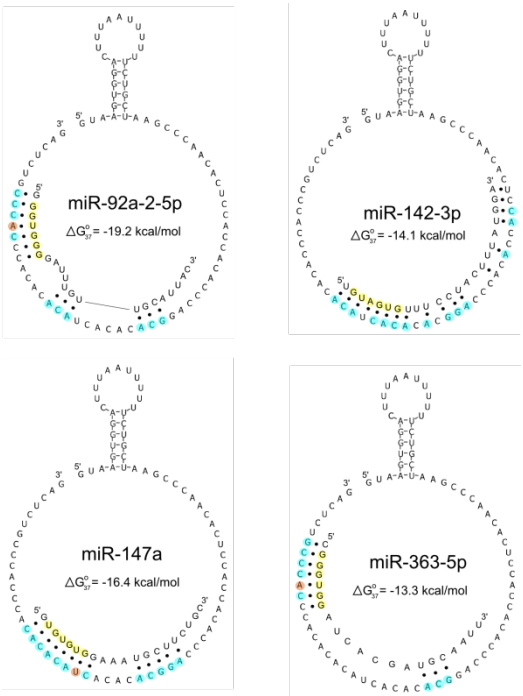
program RNAalifold [45], which finds consensus structures for homologous RNAs and identifies consistent and compensatory mutations. Initial alignments and models were manually refined to maximize conservation of base pairing and structure-preserving mutations. Reliability plots from partition function data were then generated.

Z-scores for the EBV genome and ebv-sisRNA-2 sequences were calculated using a Perl script that used a 600 nt sliding window (30 nt step size) to scan each sequence. Extracted fragments were "mutagenized" in silico to build a set of 50 randomized sequences. Free energies for the native and mutant sequences were predicted using the RNAfold program [46] and used in the calculation of the thermodynamic z-score according to the following equation: $z\text{-score} = (\Delta G° - \mu)/\sigma$. Here, $\Delta G°$ is the predicted native sequence folding free energy at 37 °C, $\mu$ is the average folding free energy of the randomized sequences, and $\sigma$ is the standard deviation of the set of free energies.

## 3 Results

### 3.1 ebv-sisRNA-1 is predicted to interact with human miRNAs

The program RNAduplex was used to predict the minimum folding free energy duplex structure between ebv-sisRNA-1 and all known EBV and human miRNAs (2622) comparisons. Of these predicted hybrids only 171 (6.5%) are predicted to have perfect base pairing in the miRNA seed region, which is the key interaction for active miRNA-mediated gene silencing. The predicted hybrids exhibit a range of energies (Fig. 2) with 649 hybrids in the most stable quartile. Of these very stable putative hybrids, only 51 also have perfect seed pairing (1.9%) to ebv-sisRNA-1. This list can be further "pruned" by considering only those known to be expressed in B cells, which leaves four putative hybrids: miR-92a-2-5p, miR-147a, and miR-142-3p, and miR-363-5p (Fig. 2). Each is able to form a stable hybrid that is rich in base pairs ~100% conserved in EBV. Conservation of base pairing also occurs in all sequenced




homologous lymphocryptovirus introns: seed interactions are 100% conserved and

**Fig. 2.** Putative ebv-sisRNA-1 miRNA hybrid structures. Predicted intermolecular pairing is shown with black "dots". 100% conserved hybridized nucleotides in ebv-sisRNA-1 are blue. When a mutation preserves pairing, it is colored orange. Paired seed nt in the miRNAs are colored yellow.

single compensating mutations are present in seed helices for all hybrids but that for miR-142-3p (Fig. 2). Interestingly, each of these predicted miRNA-interactors are also suggested to play roles in various cancers [47-50].

The interaction between ebv-sisRNA-1 and miR-142-3p is particularly interesting. This miRNA is down-regulated in EBV-associated NK/T cell lymphomas [49] and is also known to target and suppress an EBV lytic gene product (BHRF1; [8]). Intriguingly, despite its roles as a lytic protein, BHRF1 expression is driven from the W repeats in Wp-restricted latency (found in ~ 15% of Burkitt's lymphoma's; [33]). Thus, if ebv-sisRNA-1 is able to bind to and affect the levels of miR-142-3p, this may play important roles in regulating virus host interactions (e.g. the targeting of BHRF1) and, possibly, in oncogenesis. Precedent for herpesvirus small ncRNAs influencing miRNA abundance is seen in the HSURs, where HSUR 1 reduces levels of miR-27 [51]. It is noteworthy that miR-142-3p also binds HSURs 1 and 2, albeit with lower predicted stability than ebv-sisRNA-1 (8 bp in HSURs vs. 18 in ebv-sisRNA-1) and

with no apparent influence miR-142-3p abundance.

Additional experimental work is required to validate these predicted interactions and to determine if, like the HSURs, ebv-sisRNA-1 can affect the levels of endogenous host miRNAs: e.g. the interactions highlighted here or in the other miRNAs predicted to interact with ebv-sisRNA-1 with perfect seed pairing.

### 3.2 Novel sisRNA from the W repeats

Reads mapping to repeat regions are typically discarded in RNA-Seq analyses due to difficulties in quantifying transcripts from unique sites. However, when RNA-Seq data collected from EBV-infected human cells are aligned against the W repeats several important features are apparent (Fig. 3). Reads from EBV latency III are abundant and evenly cover the long W repeat intron. Indeed, intronic read peaks are equal to or greater in height than those mapping to the W1 exon (encoding EBNA-LP). This suggests that the long intron, as well as the short intron that generates ebv-sisRNA-1, is also a sisRNA (ebv-sisRNA-2). Likewise, as read density evenly covers the intron right up to the intron/exon boundaries, ebv-sisRNA-2 may comprise all the excised intronic nucleotides: as is the case for ebv-sisRNA-1, where the free sisRNA is made up of all 81 nt of the intron.

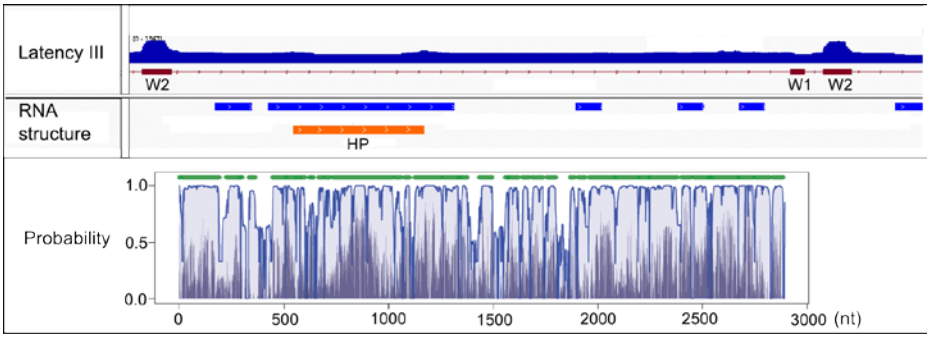



**Fig. 3.** Transcripts in the long W repeat intron. At the top are aligned RNA-Seq reads (in blue) from EBV-infected human B cells (JY cell line) expressing latency III. In red is a cartoon of the EBV genome covering ~1 W repeat. Intronic sequence is indicated with thin red arrowed lines and the W1 and W2 coding exons with solid boxes. Below this are blue "tracks" that show regions predicted to contain local RNA structure [11] and an orange track that shows the location of the long hairpin (HP) structure. At the very bottom of the figure is the LOCARNA-P reliability plot for the long intron. The blue line estimates reliability of predicted structure (scale of 0 – 1). The green bars estimate the extent of reliably predicted regions.

Previously, the long intron was found to be a "hot spot" for putative conserved and thermodynamically stable RNA secondary structure: predictions covered ~40% of the sequence (Fig. 3; [11]). This locally predicted structure was modeled and one region was found to fold into a very large, 586 nt, hairpin loop structure. In order to see if this local structure could be extended and if global, long-range, structure can exist in ebv-sisRNA-2, EBV and homologous sequences from other herpesviruses were folded and aligned simultaneously using LOCARNA-P. In contrast to initial studies of the long intron, which made use of an align-and-fold strategy (where primary nucleotide sequences were aligned and common structure predicted), simultaneous folding and aligning can better identify structural homologies in some cases [52].

LOCARNA-P also calculates base pair partition functions that can estimate the reliability of a predicted structure existing in vivo [53]. A reliability plot was generated from these data (Fig. 3), which shows that highly reliably predicted structure extends throughout ebv-sisRNA-1. This suggests that ebv-sisRNA-2 has global secondary structure that is stable and conserved, not only the local regions discovered previously. A complementary approach for estimating global RNA structure is to calculate a thermodynamic z-score. The z-score compares the native sequence folding free energy to sets of predicted free energies for dinucleotide randomized "mutant" sequences. The predicted difference in free energy is normalized by the standard deviation of the

set and thus represents how much more stable than random, is the native sequence: more negative z-scores indicate that the native sequence is much more stable than random. Compared to the EBV genome, the ebv-sisRNA-2 sequence has a much more negative mean z-score (-2.7 vs. -1.2). Comparing whisker plots for EBV vs. ebv-sisRNA-2, one sees that the ebv-sisRNA-2 sequence is shifted towards lower z-scores (Fig. 4) and thus, ebv-sisRNA-2 is more structured than expected for

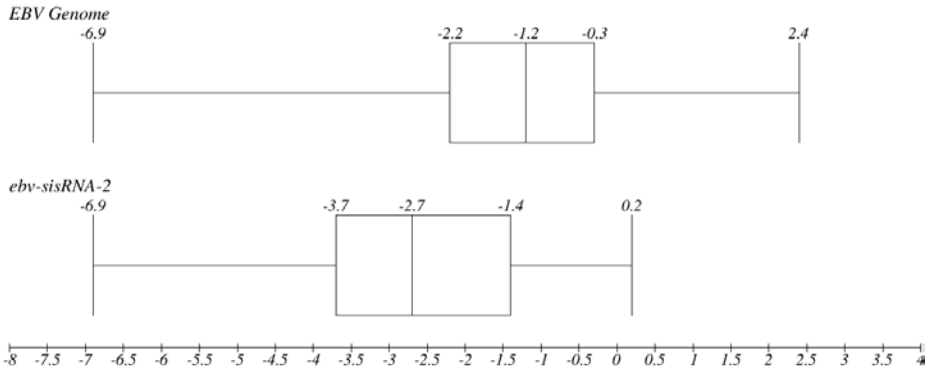



**Fig. 4.** Box plots for thermodynamic z-scores calculated for 600 nt windows covering the EBV genome and ebv-sisRNA-2 sequence. The ends of the plot (the whiskers) show the minimum and maximum values in the set, the boxed ends the first and third quartiles and the center line the median z-score value.

average EBV transcripts. The minimum value for both sequences is -6.9, which is the window that covers the long hairpin in ebv-sisRNA-2. Homologous sisRNAs from other herpesviruses showed similar shifts in their z-scores (data not shown).

To model what this global ebv-sisRNA-2 structure may look like, the LOCARNA-P alignment was used as input in the program RNAalifold, which predicts consensus structures given an alignment of homologous RNAs. This initial alignment/secondary structure was manually refined to increase base pair conservation. The final model is shown in Figure 5. Previously reported locally predicted structures are, in general, preserved. In particular the very large hairpin loop is retained in the global model (Fig. 5). Interestingly, secondary structure extends from either end of the long hairpin (interrupted by a short hairpin in a multibranch loop). This extended domain, 756 nt

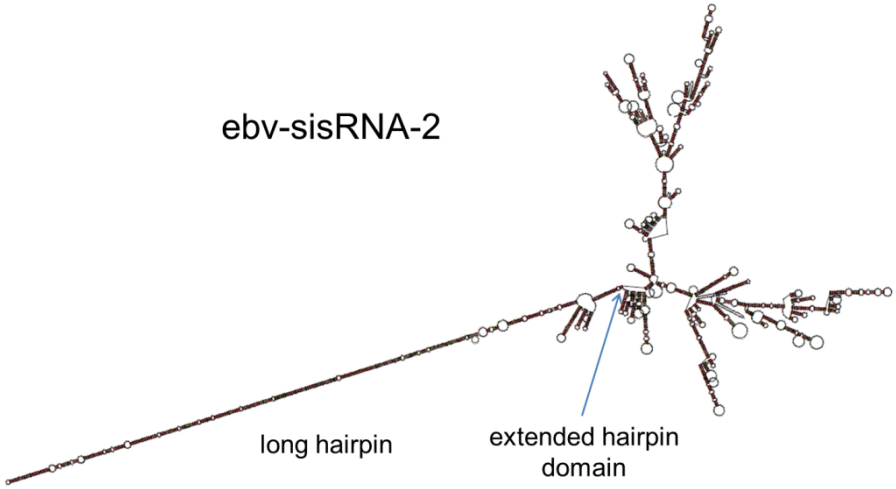



**Fig. 5.** Secondary structure model for ebv-sisRNA-2. Pairs are color-annotated with yellow indicating consistent mutations, green compensatory mutations, and red invariant.

long, corresponds to the previously predicted region of likely conserved structure (the long blue bar in Figure 3). Interestingly, this domain also corresponds to "humps" in the RNA-Seq reads in this region, which overlap with the 5' and 3' sequences that form the basal stem of this domain (Fig. 5). It is intriguing that this RNA-Seq artifact so closely follows the extent of this structural domain. Perhaps this highly structured fragment is even more resistant to degradation than the, already stable, remaining intronic sequence. Support for this model from compensatory and consistent mutations (double and single point mutations, respectively, that preserve base pairing) is highest in the long hairpin domain. However, such mutations are also observed, to a lesser degree, in the remaining model structure (Fig. 5). Further experimental work (e.g. using biochemical structure probing) will help to resolve these areas of the model.

There are a number of potential functions for RNA structure in ebv-sisRNA-2. Before being spliced, structure in the long W repeat intron may be influencing splicing: for example, by altering the accessibility or spatial organization of splicing regulatory elements [54] . The long hairpin structure was previously proposed to be the site of A-to-I (adenosine to inosine) editing [11], which occurs in long double-stranded RNA regions. A-to-I editing may play roles in splicing regulation by altering, abolishing or

creating splicing regulatory sites [55]. Post-splicing, the freed ebv-sisRNA-2 may mediate virus-host interactions important in the maintenance of latency. For example, other herpesviruses generate large (~2000 nt) stable introns known as the LATs (latency associated transcripts). Like ebv-sisRNA-2, the LATs are expressed in latency and play roles important in the maintenance of latency: for example, inhibiting cellular apoptotic pathways (e.g. caspase-8 and -9 induced apoptosis [56]). Similar roles may be played by ebv-sisRNA-2.

## 4 Conclusion

The W repeat region is an import component of the EBV genome. Two sisRNAs are generated from this region during a highly oncogenic form of latency (type III): ebv-sisRNA-1 and -2. We propose a hypothetical mode of action for ebv-sisRNA-1: the binding of host miRNAs. We highlight four potential interactions with miRNAs that can strongly bind ebv-sisRNA-1, have perfect seed pairing, and conservation of hybrid structure between EBV and related herpesviruses. We give evidence for ebv-sisRNA-2 being composed of most or all of the nucleotides in the long W repeat intron and provide a structure model for this sisRNA. These hypotheses will guide future studies aimed at understanding the roles for these exciting new RNAs in EBV virology.

## Acknowledgements

I sincerely thank Prof. Joan Steitz for her mentoring and support on this project. Also, I thank Prof. Erik Flemington (Tulane) and his lab members for providing RNA-Seq data for the analysis of ebv-sisRNA-2. This work was supported by grant CA16038 from the NIH. JAS is an investigator of the Howard Hughes Medical Institute.

## References


1. **Epstein-Barr Virus and Infectious Mononucleosis.** [http://www.cdc.gov/ncidod/diseases/ebv.htm.]
2. Parks C G, C ooper G S, H udson L L, D ooley M A, T readwell E L, S t C lair E W, Gilkeson G S, Pandey J P: **Association of Epstein-Barr virus with systemic lupus erythematosus: effect modification by race, age, and cytotoxic T lymphocyte-associated antigen 4 genotype**. *Arthritis and rheumatism* 2005, **52**(4):1148-1159.
3. Alspaugh M A, J ensen F C, R abin H , T an E M: **Lymphocytes transformed by Epstein-Barr virus. Induction of nuclear antigen reactive with antibody in rheumatoid arthritis**. *The Journal of experimental medicine* 1978, **147**(4):1018-1027.
4. Burkitt D: **A sarcoma involving the jaws in African children**. *The British journal of surgery* 1958, **46**(197):218-223.
5. Flavell KJ, Murray P G: **Hodgkin's disease and the Epstein-Barr virus**. *Molecular pathology : MP* 2000, **53**(5):262-269.
6. Burgos JS : **Involvement of the Epstein-Barr virus in the nasopharyngeal carcinoma pathogenesis**. *Med Oncol* 2005, **22**(2):113-121.
7. Chen S J, C hen G H, C hen Y H, L iu C Y, C hang K P, C hang Y S, C hen H C: **Characterization of Epstein-Barr virus miRNAome in nasopharyngeal carcinoma by deep sequencing**. *PloS one* 2010, **5**(9).
8. Riley K J, R abinowitz GS, Y ario TA , Luna J M, D arnell R B, S teitz J A: **EBV and human microRNAs co-target oncogenic and apoptotic viral and human genes during latency**. *The EMBO journal* 2012, **31**(9):2207-2221.
9. Hutzinger R , F eederle R , M razek J , S chiefermeier N , B alwierz P J, Z avolan M, Polacek N , D elecluse H J, H uttenhofer A : **Expression and processing of a small nucleolar RNA from the Epstein-Barr virus genome**. *PLoS pathogens* 2009, **5**(8):e1000547.
10. Lerner M R, A ndrews N C, M iller G , S teitz J A: **Two small RNAs encoded by Epstein-Barr virus and complexed with protein are precipitated by antibodies from patients with systemic lupus erythematosus**. *Proceedings of the National Academy of Sciences of the United States of America* 1981, **78**(2):805-809.
11. Moss W N, S teitz J A: **Genome-wide analyses of Epstein-Barr virus reveal conserved RNA structures and a novel stable intronic sequence RNA**. *BMC genomics* 2013, **14**:543.
12. Conrad NK, Fok V, Cazalla D, Borah S, Steitz JA: **The challenge of viral snRNPs**. *Cold Spring Harbor symposia on quantitative biology* 2006, **71**:377-384.
13. Fok V , M itton-Fry R M, G rech A , S teitz J A: **Multiple domains of EBER 1, an Epstein-Barr virus noncoding RNA, recruit human ribosomal protein L22**. *RNA* 2006, **12**(5):872-882.
14. Toczyski D P, M atera A G, Ward D C, S teitz J A: **The Epstein-Barr virus (EBV) small RNA EBER1 binds and relocalizes ribosomal protein L22 in EBV-infected human B lymphocytes**. *Proceedings of the National Academy of Sciences of the United States of America* 1994, **91**(8):3463-3467.

15. Lee N, Pimienta G, Steitz JA: **AUF1/hnRNP D is a novel protein partner of the EBER1 noncoding RNA of Epstein-Barr virus**. *RNA* 2012, **18**(11):2073-2082.
16. Gruber AR, Findeiss S, Washietl S, Hofacker IL, Stadler PF: **RNAz 2.0: improved noncoding RNA detection**. *Pacific Symposium on Biocomputing Pacific Symposium on Biocomputing* 2010:69-79.
17. Clote P, Ferre F, Kranakis E, Krizanc D: **Structural RNA has lower folding energy than random RNA of the same dinucleotide frequency**. *RNA* 2005, **11**(5):578-591.
18. Moss W N, Priore S F, Turner DH: **Identification of potential conserved RNA secondary structure throughout influenza A coding regions**. *RNA* 2011, **17**(6):991-1011.
19. Hofacker I L, Stadler P F, Stocsits R R: **Conserved RNA secondary structures in viral genomes: a survey**. *Bioinformatics* 2004, **20**(10):1495-1499.
20. Lodeiro MF, Filomatori CV, Gamarnik AV: **Structural and functional studies of the promoter element for dengue virus RNA replication**. *Journal of virology* 2009, **83**(2):993-1008.
21. Cordey S, Gerlach D, Junier T, Zdobnov EM, Kaiser L, Tapparel C: **The cis-acting replication elements define human enterovirus and rhinovirus species**. *RNA* 2008, **14**(8):1568-1578.
22. Pfingsten JS, Costantino DA, Kieft JS: **Structural basis for ribosome recruitment and manipulation by a viral IRES RNA**. *Science* 2006, **314**(5804):1450-1454.
23. Conrad NK, Shu MD, Uyhazi KE, Steitz JA: **Mutational analysis of a viral RNA element that counteracts rapid RNA decay by interaction with the polyadenylate tail**. *Proceedings of the National Academy of Sciences of the United States of America* 2007, **104**(25):10412-10417.
24. Mitton-Fry R M, D eGregorio S J, W ang J , S teitz T A, S teitz JA: **Poly(A) tail recognition by a viral RNA element through assembly of a triple helix**. *Science* 2010, **330**(6008):1244-1247.
25. Tycowski KT, Shu MD, Borah S, Shi M, Steitz JA: **Conservation of a triple-helix-forming RNA stability element in noncoding and genomic RNAs of diverse viruses**. *Cell reports* 2012, **2**(1):26-32.
26. Abbink T E, B erkhout B : **RNA structure modulates splicing efficiency at the human immunodeficiency virus type 1 major splice donor**. *Journal of virology* 2008, **82**(6):3090-3098.
27. Moss W N, De la-Moss L I, P riore S F, T urner D H: **The influenza A segment 7 mRNA 3' splice site pseudoknot/hairpin family**. *RNA biology* 2012, **9**(11):1305-1310.
28. Priore SF, Kierzek E, Kierzek R, Baman JR, Moss WN, Dela-Moss LI, Turner DH: **Secondary structure of a conserved domain in the intron of influenza A NS1 mRNA**. *PloS one* 2013, **8**(9):e70615.
29. Priore S F, Moss W N, T urner DH: **Influenza B virus has global ordered RNA structure in (+) and (-) strands but relatively less stable predicted RNA folding free energy than allowed by the encoded protein sequence**. *BMC research notes* 2013, **6**:330.
30. Priore SF, Moss WN, Turner DH: **Influenza A virus coding regions exhibit host-specific global ordered RNA structure**. *PloS one* 2012, **7**(4):e35989.

31. Nawrocki EP, Kolbe DL, Eddy SR: **Infernal 1.0: inference of RNA alignments**. *Bioinformatics* 2009, **25**(10):1335-1337.
32. Gardner PP: **The use of covariance models to annotate RNAs in whole genomes**. *Briefings in functional genomics & proteomics* 2009, **8**(6):444-450.
33. Kelly GL, Long HM, Stylianou J, Thomas WA, Leese A, Bell AI, Bornkamm GW, Mautner J, Rickinson AB, Rowe M: **An Epstein-Barr virus anti-apoptotic protein constitutively expressed in transformed cells and implicated in burkitt lymphomagenesis: the Wp/BHRF1 link**. *PLoS pathogens* 2009, **5**(3):e1000341.
34. Rezk SA, Weiss LM: **Epstein-Barr virus-associated lymphoproliferative disorders**. *Human pathology* 2007, **38**(9):1293-1304.
35. Gardner EJ, Nizami ZF, Talbot CC, Jr., Gall JG: **Stable intronic sequence RNA (sisRNA), a new class of noncoding RNA from the oocyte nucleus of Xenopus tropicalis**. *Genes & development* 2012, **26**(22):2550-2559.
36. Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, Enright AJ: **miRBase: microRNA sequences, targets and gene nomenclature**. *Nucleic acids research* 2006, **34**:D140-D144.
37. Lorenz R, Bernhart SH, Honer Zu Siederdissen C, Tafer H, Flamm C, Stadler PF, Hofacker IL: **ViennaRNA Package 2.0**. *Algorithms for molecular biology : AMB* 2011, **6**:26.
38. Mathews DH, Disney MD, Childs JL, Schroeder SJ, Zuker M, Turner DH: **Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure**. *Proceedings of the National Academy of Sciences of the United States of America* 2004, **101**(19):7287-7292.
39. Turner DH, Mathews DH: **NNDB: the nearest neighbor parameter database for predicting stability of nucleic acid secondary structure**. *Nucleic acids research* 2010, **38**(Database issue):D280-282.
40. Reuter JS, Mathews DH: **RNAstructure: software for RNA secondary structure prediction and analysis**. *BMC bioinformatics* 2010, **11**:129.
41. Lin Z, Wang X, Strong MJ, Concha M, Baddoo M, Xu G, Baribault C, Fewell C, Hulme W, Hedges D *et al*: **Whole-genome sequencing of the Akata and Mutu Epstein-Barr virus strains**. *Journal of virology* 2013, **87**(2):1172-1182.
42. Langmead B, Salzberg SL: **Fast gapped-read alignment with Bowtie 2**. *Nature methods* 2012, **9**(4):357-359.
43. Thorvaldsdottir H, Robinson JT, Mesirov JP: **Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration**. *Briefings in bioinformatics* 2013, **14**(2):178-192.
44. Will S, Joshi T, Hofacker IL, Stadler PF, Backofen R: **LocARNA-P: accurate boundary prediction and improved detection of structural RNAs**. *RNA* 2012, **18**(5):900-914.
45. Bernhart SH, Hofacker IL, Will S, Gruber AR, Stadler PF: **RNAalifold: improved consensus structure prediction for RNA alignments**. *BMC bioinformatics* 2008, **9**:474.
46. Gruber AR, Lorenz R, Bernhart SH, Neubock R, Hofacker IL: **The Vienna RNA websuite**. *Nucleic acids research* 2008, **36**(Web Server issue):W70-74.

47. Nilsson S , Moller C , J irstrom K , L ee A, B usch S , L amb R , L andberg G : **Downregulation of miR-92a is associated with aggressive breast cancer features and increased tumour macrophage infiltration**. *PloS one* 2012, **7**(4):e36051.
48. Caruso S, Bazan V, Rolfo C, Insalaco L, Fanale D, Bronte G, Corsini LR, Rizzo S, Cicero G, Russo A: **MicroRNAs in colorectal cancer stem cells: new regulators of cancer stemness?** *Oncogenesis* 2012, **1**:e32.
49. Motsch N, Alles J, Imig J, Zhu J, Barth S, Reineke T, Tinguely M, Cogliatti S, Dueck A, Meister G *et al*: **MicroRNA profiling of Epstein-Barr virus-associated NK/T-cell lymphomas by deep sequencing**. *PloS one* 2012, **7**(8):e42193.
50. Jonigk D , Izykowski N , M aegel L , S chormann E , M aecker-Kolhoff B , L aenger F , Kreipe H, Hu ssein K: **MicroRNA expression in Epstein-Barr virus-associated post-transplant smooth muscle tumours is related to leiomyomatous phenotype**. *Clinical sarcoma research* 2013, **3**(1):9.
51. Cazalla D , Y ario T , S teitz J A: **Down-regulation of a host microRNA by a Herpesvirus saimiri noncoding RNA**. *Science* 2010, **328**(5985):1563-1566.
52. Mathews DH, M oss W N, T urner DH: **Folding and finding RNA secondary structure**. *Cold Spring Harbor perspectives in biology* 2010, **2**(12):a003665.
53. Mathews DH: **Using an RNA secondary structure partition function to determine confidence in base pairs predicted by free energy minimization**. *Rna* 2004, **10**(8):1178-1190.
54. Warf MB, Berglund JA: **Role of RNA structure in regulating pre-mRNA splicing**. *Trends Biochem Sci* 2010, **35**(3):169-178.
55. Lev-Maor G, Ram O, Kim E, Sela N, Goren A, Levanon EY, Ast G: **Intronic Alus influence alternative splicing**. *PLoS genetics* 2008, **4**(9):e1000204.
56. Henderson G , Peng W , J in L , Perng GC, N esburn A B, W echsler S L, J ones C : **Regulation of caspase 8- and caspase 9-induced apoptosis by the herpes simplex virus type 1 latency-associated transcript**. *Journal of neurovirology* 2002, **8 Suppl 2**:103-111.