ijlpr 2021; doi 10.22376/ijpbs/lpr.2021.11.5.L81-87

International Journal of Life science and Pharma Research

Research Article

Bioinformatics for health care

In-silico evaluation of 'Mirror Repeats' In HIV Genome

Sandeep Yadav, Usha Yadav and Dinesh C. Sharma*

School of Life Sciences, Starex University, Gurugram (India)

Abstract: The repetitive sequences played an important role in the characterization of both prokaryotic & eukaryotic organisms. Various different patterns of repetitive sequences have also been identified in organisms. Among all the repeat sequences. Mirror Repeats (MR`s) play an important role in various types of neurological disorders. These MR`s have also been reported for structure determination of genomes, triplex DNA formation & various other genome functions. We have followed a distinguished method referred to as FPCB (FASTA PARALLEL COMPLEMENT BLAST) for the identification of MR`s. The above said method used to identify MR's in both types of HIV viruses (HIV-1 & HIV-2). Present investigation reported that MR's are frequently distributed in all the regions of the genomes of both types. As a result, 232 & 248 total numbers of MR`s identified in both the HIV-1 & HIV-2 genome respectively. In addition, it was also revealed that the majority of the identified sequences are imperfect. The maximum length of MR`s in HIV-1 is of 47 nucleotides (NTD`s), however in case of HIV-2, it is of 49 nucleotides (NTD`s). Present investigation will be helpful for further development of a link between mirror repeats and host genome, which will be a new trend to block the viral integration as well as pathogenicity.

Keywords: HIV, Mirror Repeats (MR's), Triplex DNA, FASTA, Parallel Complement

***Corresponding Author**

Dinesh C. Sharma , School of Life Sciences, Starex University, Gurugram (India)

Received On 23 June, 2021
Revised On 6 August, 2021
Accepted On 13 August, 2021
Published On 3 September, 2021

Funding This research did not receive any specific grant from any funding agencies in the public, commercial or not for profit sectors.

Citation Sandeep Yadav, Usha Yadav and Dinesh C. Sharma , In-Silico Evaluation of 'Mirror Repeats' in HIV Genome.(2021).Int. J. Life Sci. Pharma Res.11(5), 81-87 http://dx.doi.org/10.22376/ijpbs/lpr.2021.11.5.L81-87

1. INTRODUCTION

DNA acts as a genetic material of most of the life forms except some viruses where RNA serves as a regulatory molecule. Various forms of DNA are found in nature. These forms are uncommon and differ from the wildest form of DNA (B-form) in a variety of parameters & characteristics.[1] Studies elucidate the reasons for the existence of non-B DNA forms because of the presence of various types of repeat sequences in the genomes. [2] These repeat sequences constitute major fractions of the genomes including human beings. [3] Their existence is still attracting the attention of researchers with big curiosity to solve the puzzles associated with their roles &functions in the genomes. The most common types of repeat sequences like satellite DNA, inverted repeats, direct repeats, VNTR's, palindrome sequences, transposable elements, etc. are present in the genomes of all living creatures. [4,5] These repeats acquired both coding and non-coding parts of genomes &found to be associated with variety of roles in the same.[6] Similar to these a peculiar type of repeat refers to as Mirror Repeat (MR) also being reported. It is defined as the repeat with bilateral symmetry or having a centre of symmetry on the same strand of genome sequence. For example, ATCTGGTCTA is a mirror repeat in which ATCTG part of the sequence shares homology with the rest part of the same sequence. Studies elucidate that Mirror repeats (MR's) are associated with variety of roles in the genomes like formation of non-B DNA structures or also known as H-form of DNA.[7] These mirror repeat (MR's) sequences carries out the folding of the genome using non Watson – Crick base pairing which ultimately lead the formation of triplex DNA.[8] Modified base analogs based studies also reveal the formation of triplex DNA & provide an insight of its importance at molecular level. [9,10] These triplex DNA or non-B DNA structures formed because of the presence of triplet repeats are found to be associated with various human diseases.[11,12,13] These structures found to be mutagenic in mammalian cells & induced genetic instability and replication arresting by introducing DSB's.[14,15] These are also responsible for structure determination of genes & involve in repairing mechanisms.[16] To identify variety of repeats including Inverted repeats, Direct repeats, MR's, SSR's, triplet repeats & other kinds present in genome sequences is a curiosity based research work which utilize the power of bioinformatics in modern time

.

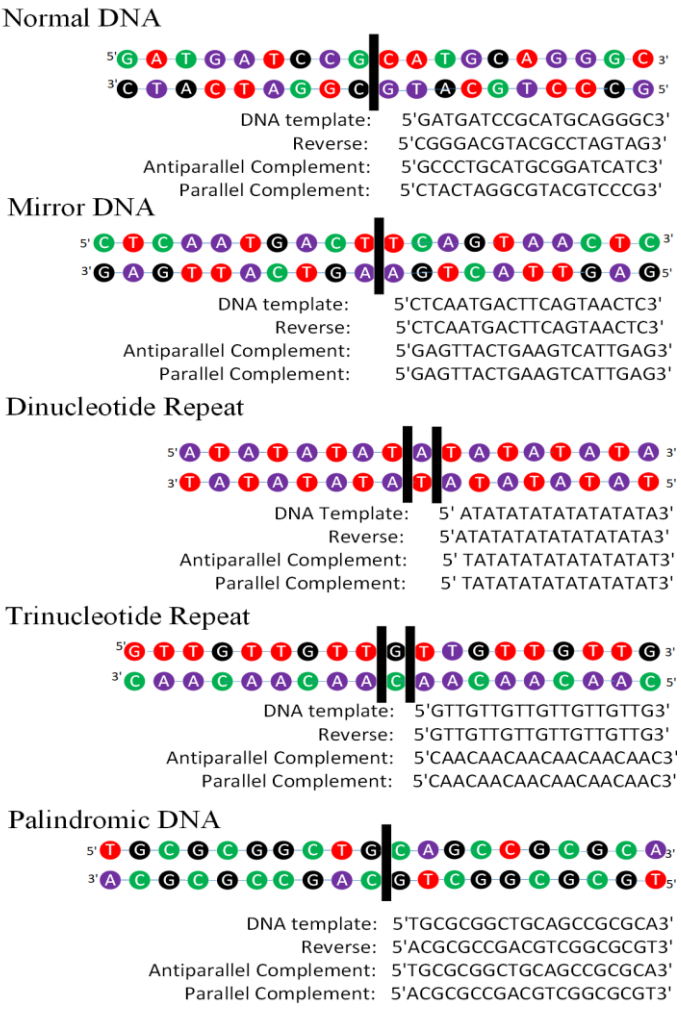

Fig 1- Depicted a comparative analysis between some of the types (not all) of repetitive sequences& normal DNA sequence. All types of repeats including Mirror repeat (MR), Dinucleotide repeat, Trinucleotide repeat and Palindrome sequence shows a unique pattern as compared to Normal DNA sequence. Along with the original sequences their reverse, parallel & anti-parallel sequences are also given in the figure. In case of MR, Di & Tri nucleotide sequence it has been found that their parallel & anti-parallel complement are same whereas this is not true in case of normal & palindrome sequence.

Various tools & approaches[17, 18], mathematical models[19] and bioinformatics based pipelines[20] are designed to identify them. The focus of our investigation is on identification of Mirror repeats (MR's) using a manual bioinformatics based approach refers to as FPCB (FASTA PARALLEL COMPLEMENT BLAST).This approach requires some simple steps and the uses of freely accessible public domain databases to identify MR's. We choose viral genomes for the investigation because of their smaller size; it is easy to handle the sequences while doing manual bioinformatics based analysis. For that purpose genome sequences of both types of HIV viruses (HIV-1 & HIV-2) were taken from the NCBI database for MR's identification.

2. MATERIALS AND METHODS

The genome sequence of HIV-1 (NC_001802.1) & HIV-2 (NC_001722.1) were taken from NCBI [21]. The whole genome sequence was divided into regions of 1000 base pairs each to ensure the identification of maximum as well as shorter MR's also. At one time analysis 1000 bps region represent your query sequence while performing the steps of MR Identification. The manual approach (FPCB) [22] involves the following steps to identify MR's sequences (Fig.1):

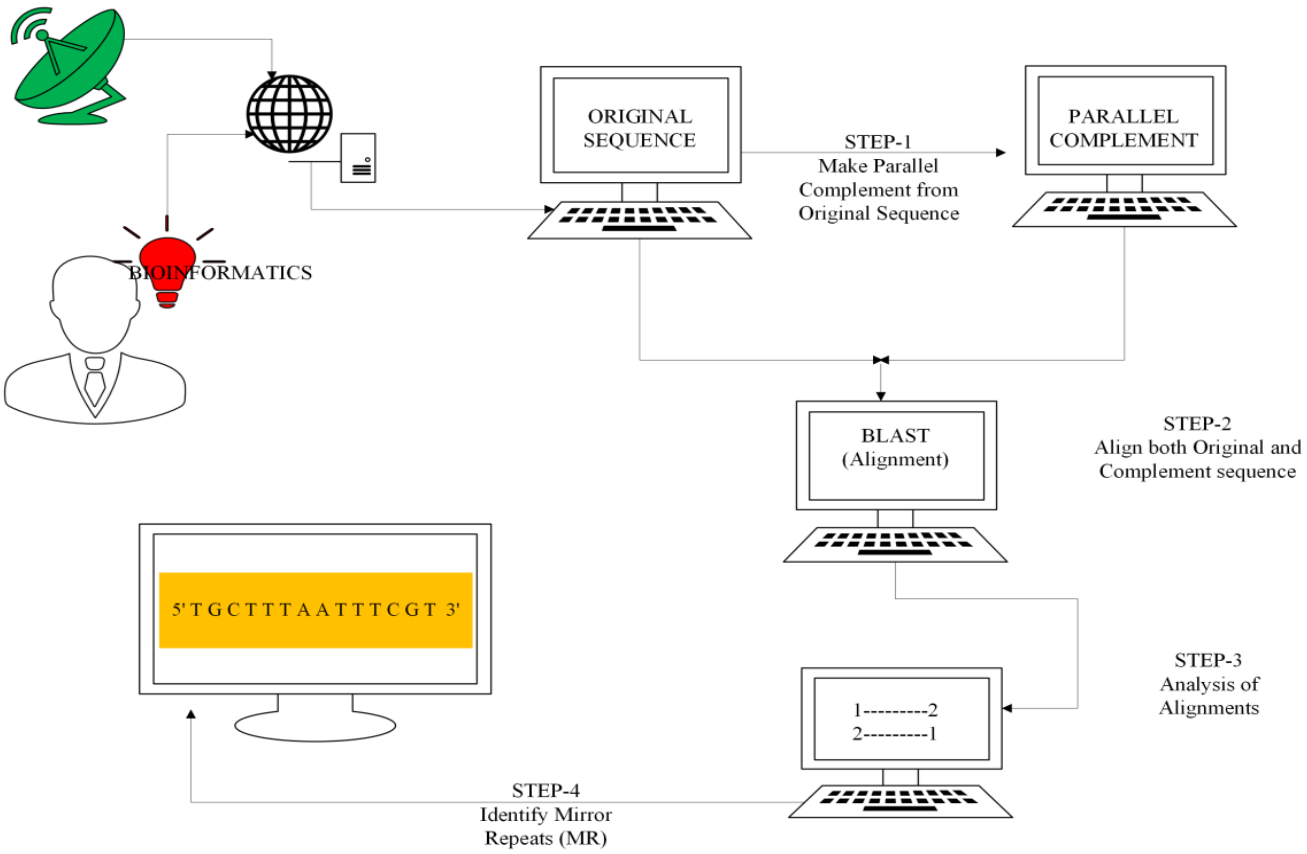

Fig 2- Stepwise Pictorial view of FPCB: involve downloading the query sequence from a public domain database, followed by making its parallel complement and the final step of BLAST analysis to identify Mirror repeats (MR's).

2.1 Steps involve in FPCB as follow

2.1.1 Downloading sequence in FASTA format

The sequences for the genome of interest (Query sequence) (HIV-1 & HIV-2 genome sequences in our case) were downloaded in FASTA format using the following link of NCBI database-http://www.ncbi.nlm.nih.gov

2.2 Making Parallel complement

Once downloaded, the FASTA format of query sequences were pasted into Reverse Complement program to make its parallel complement (subject sequence) using the following link- http://www.bioinformatics.org/sms/rev

2.3 BLAST Analysis

Both FASTA format of query sequence and its parallel complement (subject sequence) were aligned for BLAST homology search (with some selective parameters including word size-7, Expected frequency -1000 in our case) using BLAST tool:
http://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=MegaBlast&PROGRAM=blastn&BLAST_PROGRAMS=megaBlast&PAGE_TYPE=BlastSearch&BLAST_SPEC=blast2seq&QUERY=&SUBJECTS

2.4 Mirror Repeats (MR's) Identification

If the position number of the query and subject sequence is exactly reversed then it will be a mirror repeat (MR).

3. RESULTS AND DISCUSSION

3.1 Distribution of MR's in HIV-1 & HIV-2 genome

It was known to all that genome is the collections of all the genetic informations.[23] Variations in the genome result into the formation of new strains in living world including viruses .[24] Because of the varied genomic features microbial species are adapt to the changeable environment.[25] The presence of various types of repeat in the genome involve in evolution as well as diversification of viruses are also reported.[26-28] The present study is also focused on a unique repeat type whose study carried out on HIV genome using bioinformatics based approach. The both type of HIV have ≤10 kb sized genome,[29] which are closely relative on the basis of unique open reading frame sequences.[30] During the present investigation the genome sequences of both types of HIV virus were divided into 1000 bps regions (user defined criteria) to ensure maximum MR's identification as well as to cover shorter MR's as stated in the methodology. The present investigationhas identified the MR`s within every 1000 bps. As the result total no of 232 & 248 MR's sequences were identified by FPCB method in both HIV-1 & HIV-2 genome respectively. These MR's are frequently & randomly distributed in every region of the genome/s (Fig. 3 & 4). Lang worked on various genes of the HIV and successfully identified MR`s in *gag* gene. As compared to the Lang`s work, present investigation identified over 200 MR`s in both type of the HIV.[31]

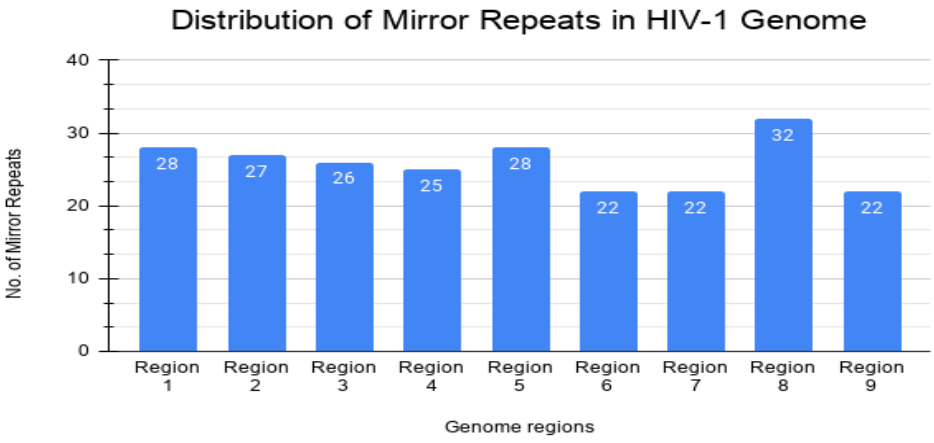

Fig: 3- Represent genome- region wise distribution of MR's in HIV-1 genome. The maximum & minimum no of MR's reported in the region 8 (32) & region 6, 7, 9 (22) respectively.

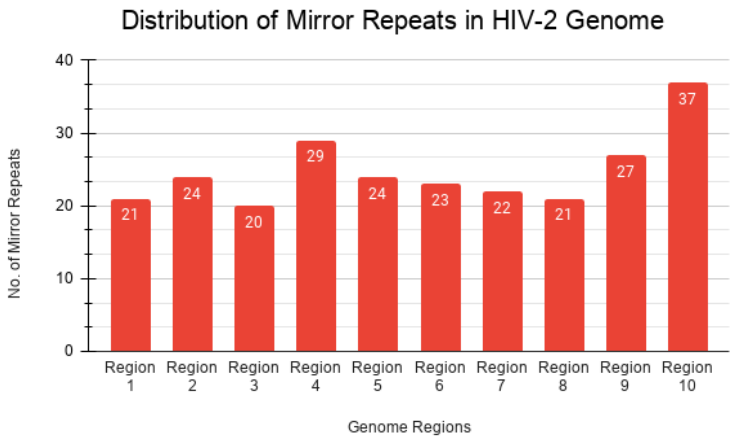

Fig: 4 - Represent genome region wise distribution of MR's in HIV-2 genome. The maximum & minimum no of MR's reported in the region 10 (37) & region 3 (20) respectively.

3.2 Length or sequence size based distribution of MR's

During the present investigation sequences of varied length have been identified, however it is noted here that abundance of perfect mirror sequences is relatively low as compared to the imperfect mirror sequences. The result identified that most of the MR sequences are of shorter length & interestingly distributed to all the regions of the genome in both HIV virus types (Table-1). The another study, it was also notified that mirror repeat sequences are distributed randomly in the genome with varied length and sequences [32]

Table.1. Represent the size/length and no. of Mirror repeat sequences in the analyzed genome regions of HIV-1 & HIV-2:

S.N.	Genome/s Information	Genome Regions (bps)	No. of mirror sequences in size range				Total MR's
			(7-5bps)	(16-5bps)	(26-5bps)	(>35bps)	
	HIV-1/ (NC_001802.1) Size- 9181 bps	1-1000	28	0	0	0	28
		1001-2000	25	1	1	0	27
		2001-3000	23	2	0	1	26
		3001-4000	21	4	0	0	25
		4001-5000	26	1	0	1	28
		5001-6000	19	2	0	1	22
		6001-7000	20	2	0	0	22
		7001-8000	30	1	1	0	32
		8001-9181	20	1	0	1	22
	HIV-2/ (NC_001722.1) Size- 10359 bps	1-1000	19	1	1	0	21
		1001-2000	19	4	0	1	24
		2001-3000	15	1	1	3	20
		3001-4000	25	2	1	1	29
		4001-5000	22	2	0	0	24
		5001-6000	20	2	1	0	23
		6001-7000	19	2	1	0	22
		7001-8000	20	0	1	0	21
		8001-9000	27	0	0	0	27
		9001-10359	35	1	1	0	37

Result mentioned number of mirror repeats in both types of HIV. The study has been carried out on every 1000 bps intervals of HIV-1 &HIV-2 genome.

3.3 Genome location of identified MR's

Both shorter & larger Mirror repeat sequences were identified (see Table-1). Among them some of the larger sequences (not all) & their locations in the genome (from query start site to end site) is mentioned below in the Table-2. The middle base pairs of larger Mirror repeat (MR) sequences are not completely mentioned here (only some start point bases & end point bases were given, rest portion denoted by dashes (-------) which represent middle base pairs). The maximum length of MR`s in HIV-1 is of 47 nucleotides (NTD`s), however in case of HIV-2, it is of 49 nucleotides (NTD`s). For complete sequences & all other shorter & larger sequence related details like their genome location, length, perfect/imperfect nature, MR numbers etc. see supplementary data files of HIV-1 & HIV-2. The present study also supported by the previous one, in which the Lang & coworkers also identified varied ntds sequences and the 63 ntds reported as maximum length of MR`s in HIV, however 15 ntds reported as minimum length.[31]-The mirror repeat profiling of the genome of an organism is the novel way to characterize the organism and set a new classification strategies[33]. On the basis of obtained results it is noted here that mirror repeats are the essential part of the genome and further they will also be used for various diagnostic as well as therapeutic research which is stated by the previous studies also.[34, 35, 36]Some in vivo study shows that triple helical forming sequences (MR's) play important roles in various molecular level process, the same is applicable on HIV also.[37] The earlier studies related to mirror repeats have confirmed their applications in molecular level therapy.[38]

Table.2. Represent the genome locations of some longer Mirror repeat sequences in the analyzed genome regions of HIV-1 & HIV-2:

S.N.	Genome/s Information	Genome Regions (bps)	Mirror Repeats (MR's) location in genome		Length (bps)
			Query Start Site	Query End site	
	HIV-1/ (NC_001802.1) Size- 9181 bps	1-1000	103-TGTGTGCCCGTCTGT-	117	15
		1001-2000	265- ACAGG---------ATGACA	-289	26
		2001-3000	199- TAGAA---------AAAAAT	-244	46
		3001-4000	879- AAATA---------ATAAA	-901	23
		4001-5000	594- ACAGA--------AGACA	-629	37
		5001-6000	711- GACAA--------AAGAG	-755	46
		6001-7000	311- ATAATA---------TACTA	-335	25
		7001-8000	273- AAAGA----------AGAAA	-298	26
		8001-9181	686- ACCAC----------CACCA	-729	47
	HIV-2/ (NC_001722.1) Size- 10359 bps	1-1000	196- TAGTA---------ATGAT	-227	33
		1001-2000	151- GAAAA--------AAAAG	-186	39
		2001-3000	767- CCAAA--------ACACC	-805	40
		3001-4000	925- AAAAG-------GAAAA	-973	49
		4001-5000	679-CCCAT--------TACCC	-698	20

	5001-6000	911- GAGAA--------AACAG -935	27
	6001-7000	204- GGACC---------CCAGG -229	26
	7001-8000	633- TAAGA--------AGGAT -667	35
	8001-9000	292- CGCGGTCCCTGACGC -306	15
	9001-10359	700- TAGTA-------CATGAT-731	32

Result mentioned shows the specific genome locations of selected mirror sequences

4. CONCLUSIONS

The roles & functions of repetitive sequences is still a curiosity based puzzle whose complete information is still untouched. Our present study concluded that both shorter and longer mirror repeats have frequently been distributed in the HIV genome. There were 232 and 248 MR`s of varied length identified in the genome of both HIV-1 and HIV-2, respectively. Among them larger MR's frequency is low in both cases (HIV-1 & HIV-2). It was also concluded here that frequency and distribution of shorter mirror repeats in the genome of HIV is maximum. The present study shall provide a futuristic approach in the context of diagnostics & therapeutics, development of new drug or therapy or in drug designing of HIV and also helpful to understand the pathogenesis mechanisms. This study will also be helpful in the development of new bioinformatics tools or software for Mirror repeats (MR's) identification.

5. ACKNOWLEDGEMENT

The authors of this article want to acknowledge Mr. Mohinder Singh Ji, Hon`ble Chancellor, Starex University, Gurugram for providing facilities to do research work.

6. AUTHORS CONTRIBUTION STATEMENT

The whole research idea as well as experimental setup was carried out by Mr. Sandeep Yadav. Ms. Usha Yadav worked together with Mr. Sandeep Yadav to find the new and applied approaches of the current research. Dr. Dinesh C Sharma, corresponding author helped in the modification and representation of the data for the development of the manuscript. All the authors read and approved the final version of the manuscript.

7. FUNDING ACKNOWLEDGEMENT

This research did not receive any specific grant from any funding agencies in the public, commercial or not for profit sectors.

8. CONFLICT OF INTEREST

Conflict of interest declared none.

9. REFERENCES

1. Ghosh A, Bansal M. A glossary of DNA structures from A to Z. ActaCrystallogr D BiolCrystallogr. 2003 Feb 10;59(4):620–6.doi:10.1107/S0907444903003251
2. Brazda V, Fojta M, Bowater RP. Structures and stability of simple DNA repeats from bacteria. Biochem J. 2020 Jan 3;477(2):325–39.doi:10.1042/BCJ20190703
3. de Koning APJ, Gu W, Castoe TA, Batzer MA, Pollock DD. Repetitive elements may comprise over Two-Thirds of the human genome. PLoS Genet. 2011 Dec 4;7(12):1–13.doi:10.1371/journal.pgen.1002384
4. Jurka J, Kapitonov V V., Kohany O, Jurka M V. Repetitive sequences in complex genomes: Structure and evolution. Annu Rev Genomics Hum Genet. 2007 May 21;8:241–59. doi:10.1146/annurev.genom.8.080706.092416
5. Mehrotra S, Goyal V. Repetitive Sequences in Plant Nuclear DNA: Types, Distribution, Evolution and Function. Genomics, Proteomics Bioinforma. 2014 July 3;12(4):164–71.doi:10.1016/j.gpb.2014.07.003
6. Shapiro JA, Von Sternberg R. Why repetitive DNA is essential to genome function. Biol Rev CambPhilos Soc. 2005 Sep 29;80(2):227–50.doi:10.1017/S1464793104006657
7. Mirkin SM, Frank-kamenetskii MD. H-DNA AND RELATED STRUCTURES. Annu Rev BiophysBiomolStruct. 1994 June;23:541–76. doi:10.1146/annurev.bb.23.060194.002545
8. Belotserkovskii BP, Veselkov AG, Filippov SA, Dobrynin VN, Mirkin SM, Frank-Kamenetskii MD. Formation of intramolecular triplex in homopurine-homopyrimidine mirrors repeats with point substitutions. Nucleic Acids Res. 1990 Oct 13;18(22):6621–4.doi:10.1093/nar/18.22.6621
9. Pant P, Fisher M. DNA triplex with conformationally locked sugar disintegrates to duplex: Insights from molecular simulations. BiochemBiophys Res Commun. 2020 Aug 25;532(4):662–7. doi:10.1016/j.bbrc.2020.08.097
10. Taniguchi Y, Magata Y, Osuki T, Notomi R, Wang L, Okamura H, et al. Development of novel C-nucleoside analogues for the formation of antiparallel-type triplex DNA with duplex DNA that includes TA and dUA base pairs. Org Biomol Chem. 2020 Mar 17;18(15):2845–51.doi:10.1039/D0OB00420K
11. LeProust EM, Pearso CE, Sinden RR, Gao X. Unexpected formation of parallel duplex in GAA and TTC trinucleotide repeats of Friedreich's ataxia. J Mol Biol. 2000 July 24;302(5):1063–80. doi:10.1006/jmbi.2000.4073
12. Bissler JJ. Triplex DNA and Human diseases. Front Biosci. 2007 May 1;12:4536–46.doi:10.2741/2408
13. Zhang J, Fakharzadeh A, Pan F, Roland C, Sagui C. Atypical structures of GAA/TTC trinucleotide repeats underlying Friedreich's ataxia: DNA triplexes and RNA/DNA hybrids. Nucleic Acids Res. 2020 Aug 4;48(17):9899–917.doi:10.1093/nar/gkaa665
14. Wang G, Vasquez KM. Naturally occurring H-DNA-forming sequences are mutagenic in mammalian cells. ProcNatlAcadSci U S A. 2004 Aug 10;101(37):13448–53.doi:10.1073/pnas.0405116101
15. Wang G, Vasquez KM. Non-B DNA structure-induced genetic instability. Mutat Res – FundamMolMech Mutagen. 2006 Mar 3;598(1–2):103–

19.doi:10.1016/j.mrfmmm.2006.01.019
16. Usdin K, House NCM, Freudenreich CH. Repeat instability during DNA repair: Insights from model systems. Crit Rev BiochemMol Biol. 2015 Dec 12;50(2):142–67.doi:10.3109/10409238.2014.999192
17. Gurusaran M, Ravella D, Sekar K. RepEx: Repeat extractor for biological sequences. Genomics. 2013 July 12;102(4):403–8.doi:10.1016/j.ygeno.2013.07.005
18. Shortt JA, Ruggiero RP, Cox C, Wacholder AC, Pollock DD. Finding and extending ancient simple sequence repeat-derived regions in the human genome. Mob DNA. 2020 Feb 4;11(1):1–12.doi:10.1186/s13100-020-00206-y
19. Reza Dawoudi M. Mathematical Modeling Approaches to Understanding Severe Acute Respiratory Syndrome Coronavirus 2 (SARSCoV-2) DNA Sequences Linked Coronavirus Disease (COVID-19) for Discovery of Potential New Drugs. Open Access J Biomed Sci. 2020 May 27;2(2):316–20. doi:10.38125/OAJBS.000173
20. Negm S, Greenberg A, Larracuente AM, Sproul JS. RepeatProfiler: A pipeline for visualization and comparative analysis of repetitive DNA profiles. MolEcolResour. 2021 Jan 4;21(3):969–81. doi:10.1111/1755-0998.13305
21. National Center for Biotechnology Information (NCBI)[Internet]. Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information; [1988] – [cited 2021 April 22]. Available from: http://www.ncbi.nlm.nih.gov
22. Vikash B, Swapni G, Sitaram M, Kulbhushan S. FPCB: a simple and swift strategy for mirror repeat identification. Preprint arXiv: 2013 Dec 13;1312.3869. https://arxiv.org/abs/1312.3869v1
23. Beji S, Fontaine V, Devaux R, Thomas M, Negro SS, Bahrman N, Siol M, Aubert G, Burstin J, Hilbert JL, Delbreil B. Genome-wide association study identifies favorable SNP alleles and candidate genes for frost tolerance in pea. BMC genomics. 2020 Dec;21(1):1-21. doi:10.1186/s12864-020-06928-w
24. Nasir A, Caetano-Anollés G. A phylogenomic data-driven exploration of viral origins and evolution. Science advances. 2015 Sep 1;1(8):e1500527. doi: 10.1126/sciadv.1500527
25. Koskella B, Vos M. Adaptation in natural microbial populations. Annual Review of Ecology, Evolution, and Systematics. 2015 Dec 4;46:503-22. doi:10.1146/annurev-ecolsys-112414-054458
26. Berkhout B, Jeang KT. Functional roles for the TATA promoter and enhancers in basal and Tat-induced expression of the human immunodeficiency virus type 1 long terminal repeat. Journal of virology. 1992 Jan;66(1):139-49.doi:10.1128/jvi.66.1.139-149.1992
27. Deng Z, Wang Z, Lieberman PM. Telomeres and viruses: common themes of genome maintenance. Frontiers in oncology. 2012 Dec 31;2:201.doi:10.3389/fonc.2012.00201
28. Meshram CD, Agback P, Shiliaev N, Urakova N, Mobley JA, Agback T, Frolova EI, Frolov I. Multiple host factors interact with the hypervariable domain of chikungunya virus nsP3 and determine viral replication in cell-specific mode. Journal of virology. 2018 Aug 15;92(16):e00838-18.doi:10.1128/JVI.00838-18
29. Li G, Piampongsant S, Faria NR, Voet A, Pineda-Peña AC, Khouri R, Lemey P, Vandamme AM, Theys K. An integrated map of HIV genome-wide variation from a population perspective. Retrovirology. 2015 Dec;12(1):1-8. doi: 10.1186/s12977-015-0148-6
30. Shibata R, Miura T, Hayami M, Ogawa K, Sakai H, Kiyomasu T, Ishimoto A, Adachi A. Mutational analysis of the human immunodeficiency virus type 2 (HIV-2) genome in relation to HIV-1 and simian immunodeficiency virus SIV (AGM). Journal of virology. 1990 Feb;64(2):742-7. doi: 10.1128/JVI.64.2.742-747.1990.
31. Lang DM. Imperfect DNA mirror repeats in the gag gene of HIV-1 (HXB2) identify key functional domains and coincide with protein structural elements in each of the mature proteins. Virology journal. 2007 Dec;4(1):1-3. doi:10.1186/1743-422X-4-113
32. Cox R, Mirkin SM. Characteristic enrichment of DNA repeats in different genomes. Proceedings of the National Academy of Sciences. 1997 May 13;94(10):5237-42. doi:10.1073/pnas.94.10.5237
33. Ussery DW, Wassenaar TM, Borini S. Word Frequencies and Repeats. InComputing for Comparative Microbial Genomics 2009 (pp. 137-150). Springer, London. doi:10.1007/978-1-84800-255-5_8
34. Tateishi-Karimata H, Sugimoto N. Chemical biology of non-canonical structures of nucleic acids for therapeutic applications. Chemical Communications. 2020;56(16):2379-90. doi:10.1039/C9CC09771F
35. Jain A, Wang G, Vasquez KM. DNA triple helices: biological consequences and therapeutic potential. Biochimie. 2008 Aug 1;90(8):1117-30. doi:10.1016/j.biochi.2008.02.011
36. Wilson WD, Tanious FA, Mizan S, Yao S, Kiselyov AS, Zon G, Strekowski L. DNA triple-helix specific intercalators as antigene enhancers: unfused aromatic cations. Biochemistry. 1993 Oct 1;32(40):10614-21.doi:10.1021/bi00091a011
37. Buske FA, Mattick JS, Bailey TL. Potential in vivo roles of nucleic acid triple-helices. RNA biology. 2011 May 1;8(3):427-39.doi:10.4161/rna.8.3.14999
38. Dey I, Rath PC. A novel rat genomic simple repeat DNA with RNA-homology shows triplex (H-DNA)-like structure and tissue-specific RNA expression. Biochemical and biophysical research communications. 2005 Feb 4;327(1):276-86. doi:10.1016/j.bbrc.2004.12.015