Life Science Informatics Publications

Research Journal of Life Sciences, Bioinformatics, Pharmaceutical and Chemical Sciences

Journal Home page http://www.rjlbpcs.com/

RJLBPCS

ISSN 2454-6348

Original Research Article **DOI: 10.26479/2018.0406.24**

GENOME WIDE ANALYSIS OF MICROSATELLITE REPEATS IN THE *PARAMYXOMAVIRIDAE* FAMILY VIRUSES: AN *INSILICO* APPROACH

KiranKumar Burranboina[1*], Akash Swain[1], Arnika Swain[1], Ajay kumar sahu[1], Anadi, Vishwanath T[2], Shilpa BR[3]

1. Department of Biotechnology and Microbiology, Bangalore City College, Bengaluru, Karnataka, India.
2. Department of Microbiology, Maharani's Science College for Women, Bengaluru, Karnataka, India.
3. Department of Biotechnology, REVA Institute of Science and management, Bengaluru, Karnataka,India.

ABSTRACT: Microsatellites also known as simple sequence repeats (SSRs) present in across coding and non-coding regions of prokaryotes, eukaryotes and viruses, Mainly used as neutral genetic marker. Paramyxoma viruses are the enveloped, single-strand RNA viruses in the large family of *Paramyxomaviridae.* These viruses have emerged to infect humans and animals previously act as unidentified zoonoses. And have been previously recognized as bio-medically and veterinary important. In this study we analysed whole genome paramyxoma viral evolutionary tree and screened microsatellite repeats in 20 paramyxoma viral genomes along with emerging viruses, a total of 1336 SSRs and revealed a total of 76 cSSRs distributed across all the genomes. Among all paramyxoma viruses dinucleotides more prevalence followed mononucleotide and trinucleotides. Microsatellites are sequences of mono-, di-, tri-, tetra- and pentanucleotide units which are widely distributed in the genome of species and these repeats are the most common choice for molecular genetic studies. Among *Paramyxoviridae* family, *Pneumovirinae* viruses shows more divergence and high rate of Simple sequence repeats than *Paramyxovirinae* viruses. Analysis of microsatellite content in whole genome sequences would conjointly facilitate comprehensive studies on the direct role of microsatellites in genome organization and evolution of simple sequence repeats, a model adaptation to divergent host and genome evolution.

KEYWORDS: *Paramyxoviridae* family, Microsatellites, Phylogenetic analysis, Emerging viruses.

Corresponding Author: Dr. KiranKumar Burranboina* Ph.D.
Department of Biotechnology and Microbiology, Bangalore City College,
Bengaluru, Karnataka, India. Email Address: kirankumarq@gmail.com

1. INTRODUCTION

Paramyxoviridae is a diverse family of viruses which consist of many human, animal, and zoonotic pathogens. Paramyxoviruses are known to cause systemic, respiratory and neurological diseases in a wide variety of animals which also includes humans [1,24] . These are enveloped, negative-stranded RNA viruses that are classified into two subfamilies: Paramyxovirinae and Pneumovirinae. Within the subfamily of Paramyxovirinae, there are currently seven genera: Aquaparamyxovirus, Avulavirus, Ferlavirus, Henipavirus, Morbillivirus, Respirovirus and Rubulavirus. Among the members of the Paramyxovirinae, measles virus (MeV), mumps virus and human parainfluenza virus 1–4 (HPIV-1–4) are the most well-known human paramyxoviruses that can cause outbreaks of respiratory/systemic infections [2,3,4]. In the last few decades, some novel paramyxoviruses belonging to the subfamily of Paramyxovirinae have emerged in animals and humans, causing severe illness and death. These are best epitomize by Hendra virus (HeV) [5,6] and Nipah virus (NiV) [7,8]. Natural reservoir for both viruses were subsequently found to be the Fruit bats [9,10,6]. Further outbreak in human is mainly occurred as a result of transmission from horses and pigs respectively [7,5]. A novel paramyxovirus species around 66 were identified throughout world survey of 5,000 bat specimans [25]. Among the members of paramyxovirus family Nipah virus is a newly discovered virus and this virus was responsible for a viral encephalitis outbreak in Malaysia in October 1998 and ended around midsummer 1999 [11]. Nipah virus was confirmed to be the cause of both pig and human disease. As a new human pathogen Nipah virus emerged under changing ecological conditions [12]. Nipah virus infection was confirmed during 2001-2004 in Bangladesh, but patients might have been infected directly by the consumption of fruits contaminated by bats [13]. Another emerging virus is Hendra virus, characterized by fever and respiratory distress in horses and is transmissible to human occurred uncommonly in Northeastern Australia. The virus is a lethal zoonotic agent capable of causing natural disease in human and horses and experimentally induced disease in cat, mice and guinea-pigs [14]. All paramyxoviruses are enveloped and have diameter of 150 to 300nm and are highly pleomorphic and a genome size of 15-16 kb. They entirely replicate in the cytoplasm. The virion has a nucleocapsid core containing the RNA genome and three nucleocapsid associated proteins: an RNA-binding protein, a phosphoprotein and a large protein. The matrix protein occupies space between the core and the envelope. The envelope is covered with spikes consisting of one glycoprotein that is involved in the cell attachment (haemagglutinin–neuraminidase, haemagglutinin or protein G) and another glycoprotein involved in the fusion of the viral envelope with the plasma membrane of cells [15,16]. Tandem repeats (TRs) have two subcategories i.e. Microsatellites (1 to 10 nucleotides) and minisatellites (> 10 nucleotides). They along with the predominant interspersed repeats make up genomic repetitive regions. Microsatellites, also known as Simple sequence repeats (SSR) are highly polymorphic, co-dominant markers that are specifically very important for the construction of genetic maps, comparative mapping, population genetics surveys, and paternity analyses [17]. Their repeat

number, length, and motif size does influence microsatellite mutability. For an instance, more the number of repeats, higher the mutability will be [18]. Moreover, variations in copy number due to strand slippage and unequal recombination highlight the instability of the microsatellites [19]; which makes them a predominant source of genetic diversity and a crucial player in viral genome evolution. Microsatellites represent non-randomly distributed 1 to 6 nucleotide units which are widely distributed in the genome of prokaryotes, eukaryotes and viruses. They support in coping up with the sudden changes in the environment. (Microsatellite Marker Analysis). Simple Sequence Repeats have many applications. They are used in variety of fundamental and applied fields of life and medical science as they are highly polymorphic, relatively small size and rapid detection protocols. The microsatellite mutation rate is estimated at 10^{-2} –10^{-6} per locus per generation [20], which is several orders of magnitude greater than that of regular non-repetitive DNA (10^{-9}) [21]. Due to their ceaseless mutational degree,hypermutability skills and widespread length variations of microsatellites useful in ascertain the driving forces of evolution by using powerful molecular techniques [26]. The evolutionary and biological significance of microsatellites is yet to understood well. However a closer investigation of microsatellite distributions and their association with other genomic features may help to elucidate the exact functions of microsatellites in the genome and their role in evolution. In the present study, we have investigated the occurrence and abundance of microsatellites across the genomes of Paramyxoma viruses. Such studies not only enhance our knowledge of microsatellite dynamics in the virus genomes but also provide important leads to understand the complex functional roles that are associated with these simple repeats involved in the infection biology of these viruses.

2. MATERIALS AND METHODS

2.1 Analysis of Paramyxoviridae Genomes

NCBI was used to obtain *Paramyxoviridae* family viral genomes and was analysed under FASTA. Phylogenetic analysis tree of the viruses falling under paramyxoviridae was constructed using Mega 6 software packages (Tamura K, 2013) [22] The nucleotide size ranges from 13335 bp (Acc No. NC_004148.2) to 18252 bp (Acc No. FJ513078.1). Table 1 presents the accession numbers and salient features of *paramyxovirdae* family genome.

2.3 Microsatellite Identification and Analysis

IMEx software is used to perform microsatellite identification. Advance-mode of IMEx was used to identify simple and compound microsatellite repeats; the parameters used are types of repeat: perfect; repeat size: all minimum repeat numbers: 6, 3, 3, 3, 3, 3; maximum distance allowed between any two SSRs (DMAX): 10. The parameters are defined earlier in this paper [23, 27].

2.4 Statistical Analysis

Microsoft Office Excel packages 2013 were used to perform the statistical analysis. Linear regression was used to detect the correlation between the relative abundance and relative density of microsatellites with genome size.

Table 1: An overview of the paramyxoviral genomes

Sl. No	Sp. ID	Genus	Species	Acc No.	Genome size (K bp)	GC%
Paramyxovirinae						
1	PV1	Henipavirus	Hendra virus (HeV)	JN255805	18.234	39.31
2	PV2		Nipah virus (NiV)	FJ513078	18.252	38.51
3	PV3	Respirovirus	Sendai virus (SeV)	AB855653	15.385	46.21
4	PV4	Rubulavirus	Mumps virus (MuV)	JF727651	15.384	42.53
5	PV5		Mapuera virus (MPRV)	NC_009489	15.486	44.67
6	PV6	Morbilivirus	Measles virus (MeV)	FJ161211	15.894	47.32
7	PV7		Rinderpest Virus (RPV)	LC057619	15.885	47.74
8	PV8		Canine distemper virus (CDV)	AB490680	15.69	43.08
9	PV9		Peste des petits ruminants virus (PPRV)	KP868655	15.954	48.14
10	PV10	Ferlavirus	Fer-de-lance virus (FDLV)	KJ956408	15.378	43.33
11	PV11	Avulavirus	Newcastle disease virus (NDV)	NC_002617	15.186	46.18
12	PV12		Avian paramyxovirus 2 (APMV-2)	EU338414	14.904	47
13	PV13		Avian paramyxovirus 3 (APMV-3)	EU782025	16.182	41.82
14	PV14		Avian paramyxovirus 4 (APMV-4)	NC_019531	15.048	46.26
15	PV15	Aquaparamyxovirus	Atlantic salmon paramyxovirus (ASPV)	EF646380	16.965	45.88
Pneumovirinae						
16	PV16	Metapneumovirus	Human metapneumovirus	NC_004148	13.335	36.42
17	PV17		Avian metapneumovirus	NC_007652	14.071	41.35
18	PV18	Pneumovirus	Bovine respiratory syncytial virus	NC_001989	15.14	33.58
19	PV19		Human respiratory syncytial virus	NC_001781	15.225	33.56
20	PV20		Murine pneumonia virus	NC_006579	14.885	40.21

3. RESULTS AND DISCUSSION

3.1 Phylogenetic Analysis

The evolutionary history was inferred by using the Maximum Likelihood method based on the Jukes-Cantor model. The tree with the highest log likelihood (-1269316.1502) is shown. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pair wise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site analysis. The involved 20 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 103903 positions in the final

dataset. Evolutionary analyses were conducted in MEGA6.

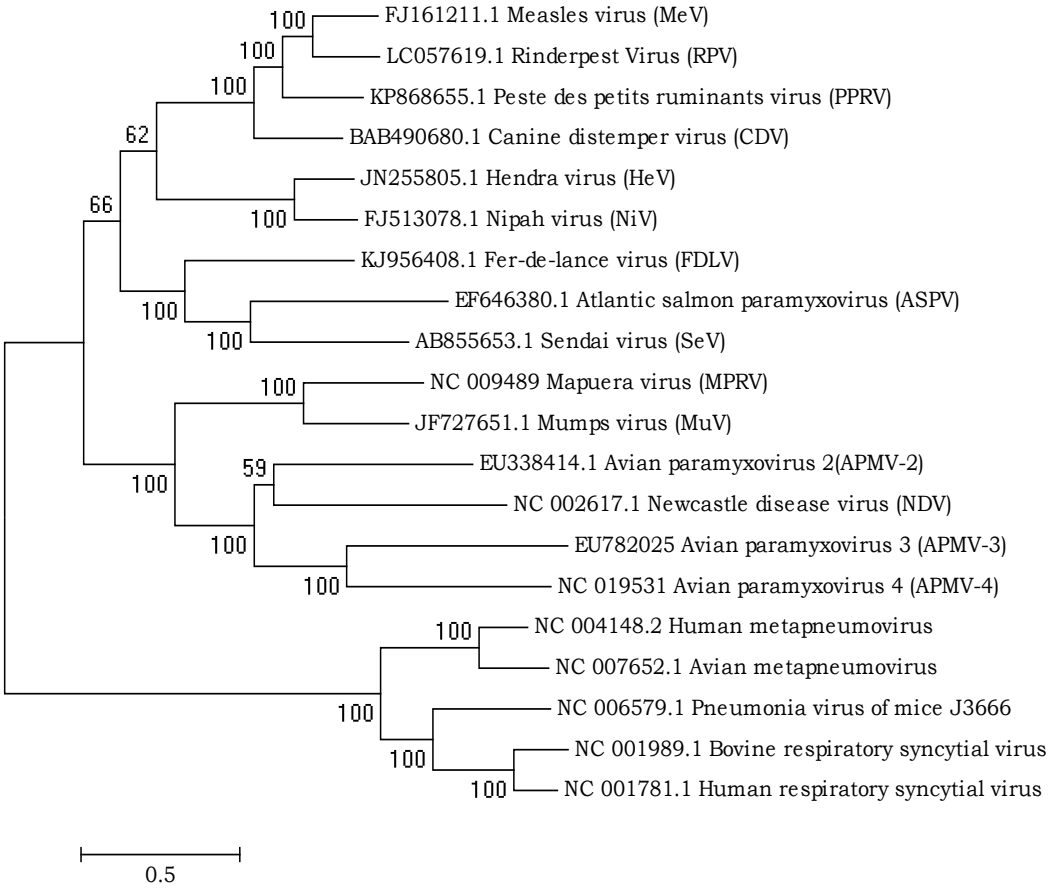


Figure 1: Molecular Phylogenetic analysis by Maximum Likelihood method

3.2 Analysis of Simple and Compound microsatellites

Genome-wide scan of Paramyxomaviral genomes IMEx revealed a total of 1336 SSRs distributed across all the species with varying incident frequencies ranging from 89 in PV18 (NC_001989.1) to 41 in PV14 (Acc No- NC_019531)(Table 2, Fig.2a). Further, the relative abundance values lied between a minimum of 2.7 for PV7 to a maximum of 6.2 bp/kb for PV17 (Table 2, Fig.3a). Comparatively, relative density of SSRs varied from 41.4 bp/kb in P17 to 17.8 bp/kb for PV7(Table 2, Fig.4a). IMEx scan for Paramyxomaviral genomes revealed a total of 76 cSSRs (Table 2, Fig.2b). In contrast to the ubiquitous presences of cSSRs, ranging from 9 in PV19 (Acc No- X69198) to 1 in PV14. Further, the relative abundance values lied between a minimum of 0.12 for PV8 to a maximum of 0.59 bp/kb for P19 (Table 2, Fig.3b). Comparatively, relative density of SSRs varied from 1 bp/kb in P12 to 10 bp/kb for P19(Table 2, Fig.4b). Transcription regions of microsatellites characteristics in rice, Arabidopsis and two mammals gradient in microsatellite density along the direction of transcription. And high frequency in 5'-flanking regions of microsatellites of plant genes potentially act as factors in regulating gene expression [29]. The evolution of some mutated viral strains may exibit potential changes from native strain because microsatellites are mainly homopolymer nucleotide repeats these are instable and may improve frameshift mutations that provide phenotypic changes [30].

Microsatellies are genetic markers and these are choice to study variety of population genetics and high quality microsatellite markers can be discoverd in non-model organisms. A large number of polymorphic microsatellite markers are helpful for fine scale genetic population structures [31]. Mutation rate is depend, among others, on the length of repeat motif and replication slippage and recombinatorial misalignment mutation mechanisms, are both creating variability in microsatellite loci [32]. The Applications of four different sets of di-, tri-, and tetranucleotide loci of a number of human populations have shwon mutation rates inversely related to their motif sizes of the non-disease-causing microsatellite loci [33].

Table 2: Summary of the microsatellites, Relative abundance and density of various simple repeat sequences detected in selected paramyxomaviral genomes

SlNo	Species ID	SSR	RA	RD	cSSR	cRA	cRD	cSSR%
Paramyxovirinae								
1	PV1	71	3.893825	24.89854	5	0.274213	4.22288	7.042254
2	PV2	75	4.109139	27.01074	5	0.273943	5.150121	6.666667
3	PV3	58	3.769906	23.20442	5	0.324992	5.264868	8.62069
4	PV4	61	3.965159	24.70099	2	0.130005	2.535101	3.278689
5	PV5	64	4.132765	25.89436	3	0.193723	3.551595	4.6875
6	PV6	52	3.271675	21.3288	5	0.314584	5.662514	9.615385
7	PV7	43	2.706956	17.81555	2	0.125905	2.392194	4.651163
8	PV8	61	3.887827	24.8566	2	0.12747	1.657106	3.278689
9	PV9	71	4.450295	28.77021	5	0.313401	6.456061	7.042254
10	PV10	53	3.446482	21.91442	2	0.130056	2.015867	3.773585
11	PV11	65	4.280258	27.1961	4	0.263401	3.687607	6.153846
12	PV12	69	4.62963	30.59581	1	0.067096	1.006441	1.449275
13	PV13	64	3.955012	27.12891	5	0.308985	9.393153	7.8125
14	PV14	41	2.724615	17.07868	1	0.066454	1.129718	2.439024
15	PV15	76	4.479811	28.41144	5	0.294724	6.601827	6.578947
Pneumovirinae								
16	PV16	74	5.549306	37.72028	3	0.224972	5.249344	4.054054
17	PV17	88	6.253998	41.43273	2	0.142136	2.203113	2.272727
18	PV18	89	5.878468	37.45046	3	0.198151	3.104359	3.370787
19	PV19	86	5.648604	36.84729	9	0.591133	10.77176	10.46512
20	PV20	75	5.038629	32.18005	7	0.470272	9.33826	9.333333

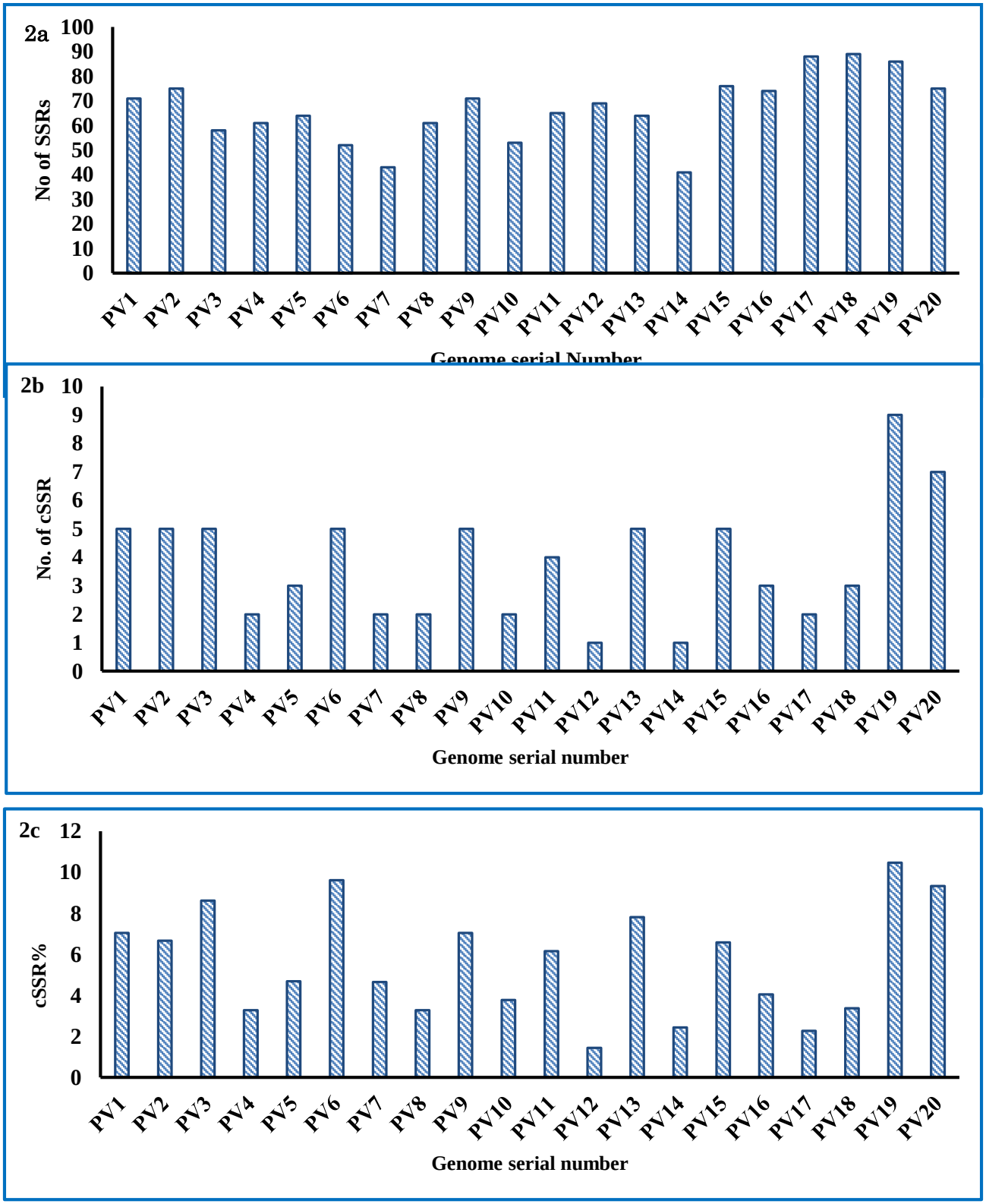


Figure 2: Analysis and Distribution of simple sequence repeats: 2a. Distribution of SSRs; 2b. Distribution of cSSRs; 2c. cSSR% across the *Paramyxomaviridae* family

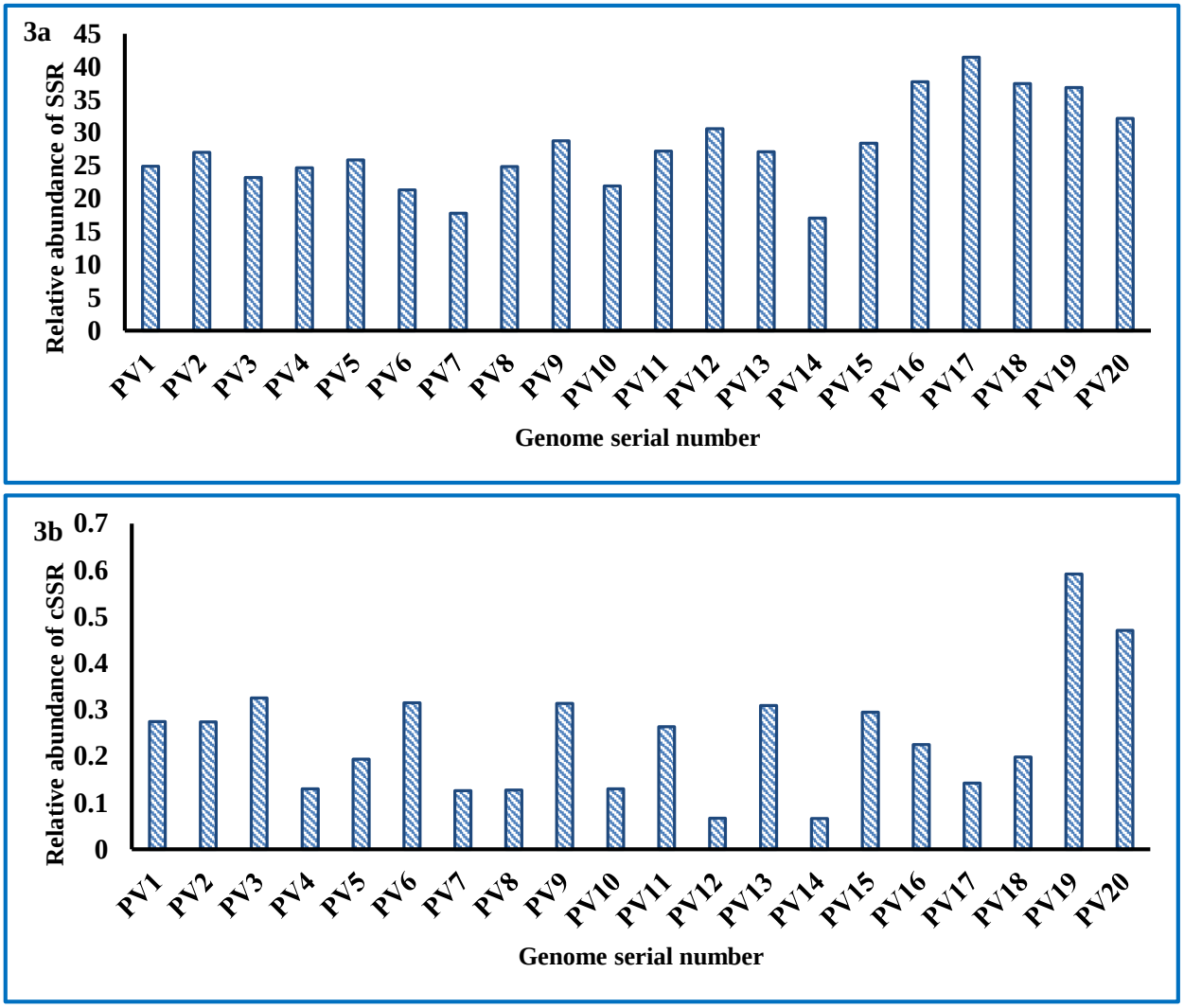
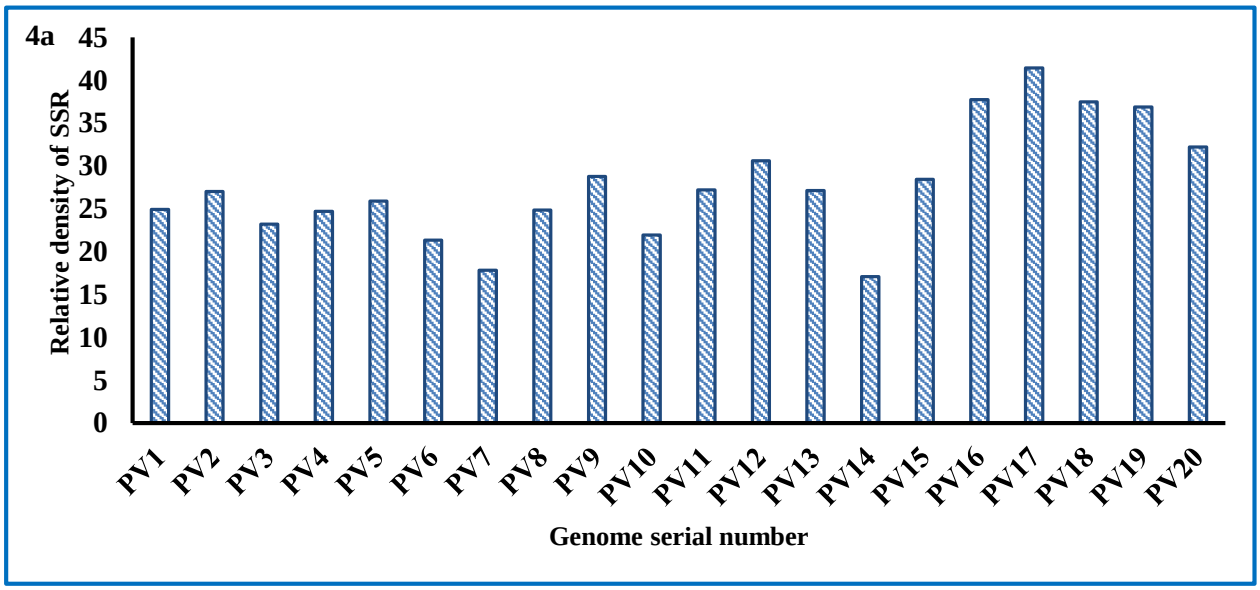


Figure 3: Relative Abundance: 3a. Simple Sequence Repeats 3b. Compound simple sequence repeats, Present per kilo base of genome.



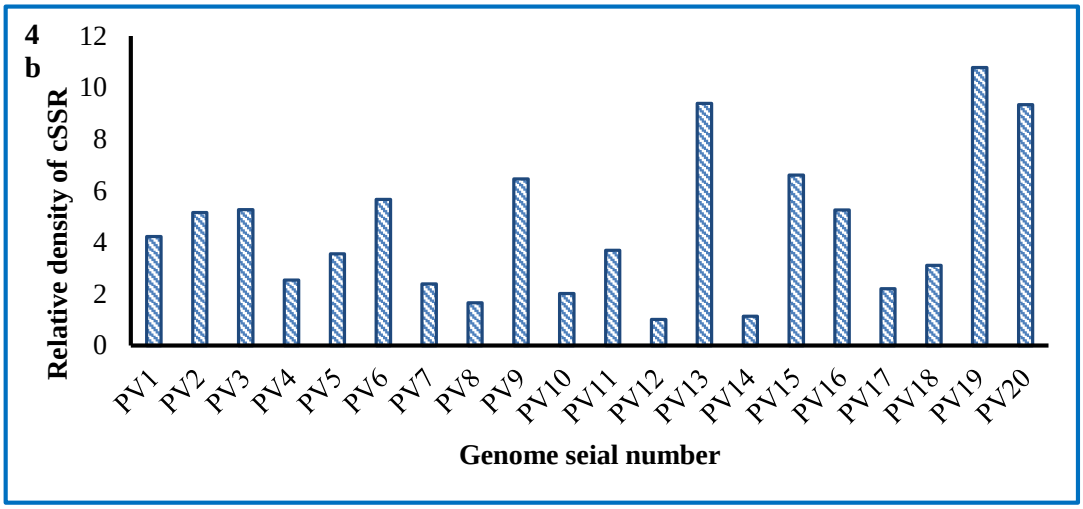


Figure 4: Relative density: Total length covered by 4a. Simple Sequence Repeats 4b. Compound simple sequence repeats, Per kilo base of genome

3.3 Diversity of Extracted SSR motifs

Tandem repeat units are often lost or gained during replication because of their instibility and their complexity and which is also called satellite repeats [28]**.** Mononucleotide repeats were observed in all the paramyxoma viral genomes analysed with poly(A/T) repeats being more prevalent than poly G/C repeats. Human respiratory syncytial virus shows higher mononucleotides than other viruses. Among all di nucleotide repeats were more around 830 dinucleotides were observed and shows more slippage rate may occur leads to mutation in viral genomes. Further maximum number of 61 dinucleotide observed in Atlantic salmon paramyxovirus (ASPV), Rinderpest Virus (RPV) shown 24 dinucleotide repeats. Tri nucleotides repeat motif observed in all paramyxoma viruses except Sendai virus. Avian metapneumovirus (12), Nipah virus (NiV)(10) and Human metapneumovirus (10) tri nucleotides were observed. Where as in tetra nucleotide repeats observed only in Avian paramyxovirus 3 (APMV-3), pentanucleotide were observed only one in Avian paramyxovirus 2 and Human metapneumovirus. Hexanucleotide observed only in Avian paramyxovirus 3 Not observed in all viruses.

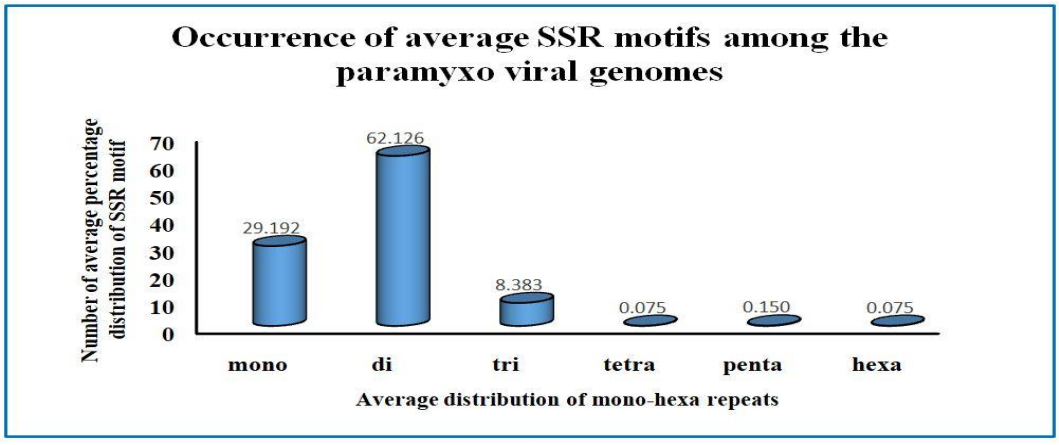


Figure 5: Occurrence of average SSR motifs among the paramyxoma viral genomes.

4. CONCLUSION

Paramyxoviridae family of viruses have emerged to infect humans previously act as a unidenttified zoonoses and the emerging disease threats such as Nipah virus and Hendra virus infections. These viruses have been previously recognized as biomedically and veterinarily important. The emerging and re-emerging viruses in the diverse family of paramyxoviridae can also be identified through comparative analysis of these repeat sequences. Therefore the distribution of different microsatellites in 20 different genera of paramyxoviridae genomes is studied among Paramyxoviridae family, Pneumovirinae viruses shows more divergence and high rate of Simple sequence repeats than Paramyxovirinae viruses. More detailed study of microsatellites in large enveloped RNA family of paramyxoma viruses along with emerging viral genomes may be useful for understanding complex biological features such as changes in virulence, and their emergence as new epidemics.

CONFLICT OF INTEREST

The authors have no conflict of interest associated with the study.

REFERENCES

1. Lau S, Woo P, Wu Y, Wong A, Wong B, Lau C, Fan R, Cai J, Tsoi H, Chan K, Yuen K. J. Gen. Identification and characterization of a novel paramyxovirus, porcine parainfluenza virus 1, from deceased pigs. Virol. 2013;94(10):2184-2190.
2. Lau SK, To WK, Tse PW, Chan AK, Woo PC, Tsoi HW, Leung AF, Li KS, Chan PK et al.. Human parainfluenza virus 4 outbreak and the role of diagnostic tests. J Clin Microbiol. 2005;43, 4515–4521.
3. Lau SK, Li KS, Chau KY, So LY, Lee RA, Lau YL, Chan KH, Lim WW, Woo PC & Yuen KY. Clinical and molecular epidemiology of human parainfluenza virus 4 infections in Hong Kong: subtype 4B as common as subtype 4A. J Clin Microbiol. 2009;47, 1549–1552.
4. Virtue ER, Marsh GA & Wang LF. Paramyxoviruses infecting humans: the old, the new and the unknown. Future Microbiol. 2009; 4,537–554.
5. Selvey LA, Wells RM, McCormack JG, Ansford AJ, Murray K, Rogers RJ, Lavercombe PS, Selleck P & Sheridan JW. Infection of humans and horses by a newly described morbillivirus. 1995; Med J Aust 162, 642–645.
6. Young PL, Halpin K, Selleck PW, Field H, Gravel JL, Kelly MA. & Mackenzie JS. Serologic evidence for the presence in Pteropus bats of a paramyxovirus related to equine morbillivirus. Emerg Infect Dis. 1996;2, 239–240.
7. Chua KB, Goh KJ, Wong KT, Kamarulzaman A, Tan PS, Ksiazek TG, Zaki SR, Paul G, Lam SK. & Tan CT. Fatal encephalitis due to Nipah virus among pig-farmers in Malaysia. Lancet. 1999; 354, 1257–1259.
8. Chua KB, Bellini WJ, Rota PA, Harcourt BH, Tamin A, Lam SK, Ksiazek TG, Rollin PE, Zaki SR. et al. Nipah virus: a recently emergent deadly paramyxovirus. Science. 2000; 288, 1432–1435.

9. Enserink M. Emerging diseases. Malaysian researchers trace Nipah virus outbreak to bats. Science. 2000; 289, 518–519.
10. Halpin K, Young PL, Field HE. & Mackenzie JS. Isolation of Hendra virus from pteropid bats: a natural reservoir of Hendra virus. J Gen Virol. 2000; 81, 1927–1932.
11. Stuart TN, Arikawa J and Kawaoka Y. Emerging viral disease. Proc. Natl. Acad. Sci. USA., 2000;97: 12411-12412.
12. Webster RG. and Hulse DJ. Microbial adaptation and change: Avian influenza. Rev. Sci. Tech. Off. Int. Epiz. 2004; 23: 453-465.
13. Desselberger U. Emerging and Re-emerging infectious disease. Journal of infection. 2000; 40: 3-15.
14. Manojkumar R and Mrudula V. Emerging Viral Disease of Zoonotic Importance-Review. International Journal of Tropical Medicine. 2006; 1(4): 162-166.
15. Lamb RA, Collins PL, Kolakofsky D, Melero JA, Nagai Y, Oldstone MBA, Pringle CR. & Rima BK. Family Paramyxoviridae. In Virus Taxonomy. Classification and Nomenclature of Viruses. Seventh Report of the International Committee on Taxonomy of Viruses. San Diego: Academic Press. 2000; 549–561.
16. Lamb RA & Kolakofsky D. Paramyxoviridae: the viruses and their replication. In Fields Virology, 4th edn, Philadelphia: Lippincott Williams & Wilkins. 2001; (1) 1305– 1340.
17. Dimitry A. Chistiakov, Bart Hellemans , Filip A.M. Volckaert. Microsatellites and their genomic distribution, evolution, function and applications: A review with special reference to fish genetics. USA OH; 2005; 45267-0529.
18. Mrazek J, Guo X, Shah A. Simple sequence repeats in prokaryotic genomes. Proc Natl Acad Sci U S A. 2007;104(20):8472–7.
19. Toth G, Gaspari Z, Jurka J. Microsatellites in different eukaryotic genomes: survey and analysis. Genome Res. 2000;10(7):967–81.
20. Ellegren H. Microsatellite mutations in the germline: implications for evolutionary inference. Trends Genet. 2000; 16, 551–558.
21. Li, WH. Molecular Evolution. Sinauer Associates, Sunderland, MA, 1997;177–213.22.
22. Tamura K, Stecher G, Peterson D, Filipski A, and Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Molecular Biology and Evolution. 2013; 30: 2725-2729.
23. Kiran Kumar Burranboina, Sunil Abraham, Kumar Kalavathi Murugan, Manjunath Reddy Gundallahalli Bayyappa, Revanaiah Yogisharadhya. Genome Wide Identification and Analysis of Microsatellite Repeats in the Largest DNA Viruses (Poxviridae Family): An Insilico Approach. ARRB; 2018. 22(1): 1-11.
24. Aguilar HC, Lee B. Emerging paramyxoviruses: molecular mechanisms and antiviral strategies. Expert Rev Mol Med. 2011;13:e6.

25. Zeltina A, Bowden TA, Lee B. Emerging Paramyxoviruses: Receptor Tropism and Zoonotic Potential. PLoS Pathog, 2016; 12(2): e1005390.
26. Saeed AF, Wang R, Wang S. Microsatellites in Pursuit of Microbial Genome Evolution. Front Microbiol. 2016;6:1462.
27. Suresh B.Mudunuri and H.A.Nagarajaram. IMEx:Imperfect Microsatellite Extractor.Bioinformatics. 2007; 23(10):1181-1187.
28. Jansen A, Gemayel R, Verstrepen KJ. Unstable microsatellite repeats facilitate rapid evolution of coding and regulatory sequences. Genome Dyn. 2012;7:108-25.
29. Fujimori S Washio T Higo K Ohtomo Y Murakami K et. al. A novel feature of microsatellites in plants: a distribution gradient along the direction of transcription. A novel feature of microsatellites in plants: a distribution gradient along the direction of transcription. 2003; 554 (1-2) pp: 17-22.
30. Deback C, Boutolleau D, Depienne C, Luyt CE, Bonnafous P, Gautheret-Dejean A, Garrigue I, Agut H. Utilization of Microsatellite Polymorphism for Differentiating Herpes Simplex Virus Type 1 Strains. Journal of Clinical Microbiology. 2009; 47 (3) 533-540.
31. Nikolic N, Duthoy S, Destombes A, Bodin N, West W, Puech A, et al. Discovery of Genome-Wide Microsatellite Markers in Scombridae: A Pilot Study on Albacore Tuna. PLoS ONE. 2015; 10(11): e0141830.
32. Alexander Renwick, Leslea Davison, Heidi Spratt, J. Patrick King and Marek Kimmel. DNA Dinucleotide Evolution in Humans: Fitting Theory to Facts. GENETICS. 2001.159(2). 737-747.
33. Ranajit Chakraborty, Marek Kimmel, David N. Stivers, Leslea J. Davison, Ranjan Deka. Relative mutation rates at di-, tri-, and tetranucleotide microsatellite loci. Proceedings of the National Academy of Sciences. 1997; 94 (3) 1041-1046.