Neural replicators analysis of human papillomaviruses

Alexandr Alexandrovich Ezhov (ezhov@triniti.ru)
Troickij institut innovacionnyh i termoadernyh issledovanij https://orcid.org/0000-0002-7208-0623

Research Article

Keywords:

Posted Date: July 22nd, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1847539/v1

Abstract

An analysis of complete human papillomavirus genomes using neural replicators is presented. When studying the shape of replicator tables, which reflect the ability of neural networks of different sizes to generate self-reproducing descendants, as well as the structures of the patterns transmitted by them, it was found that this information reliably distinguishes between two subgenera of human papillomaviruses of the genus Alpha, where one subgenus consists of Alpha-2 and Alpha-4 species belonging to warts, and the second subgenus, other species related to diseases of the mucosa, in particular, to cervical cancer. The genera Beta and Gamma are also reliably distinguished from the genus Alpha, in particular, by a different form of transmitted patterns. More interestingly, the genus Nu seems to be a good fit as a type of subgenus containing both Alpha-2 and Alpha-4 species, a fact that has not previously been revealed by conventional analysis based on DNA sequence alignment and genome similarity assessment. In addition, analysis of the patterns transmitted by the replicator network clearly distinguishes porcupine sigma papillomavirus EdPV1 from human Nu, as well as from any other human papillomavirus, which calls into question the well-known argument for horizontal gene transfer between humans and other types of papillomavirus hosts.

Introduction

The use of artificial neural replicators for the analysis of nucleotide sequences of viroid RNA was presented in [1]. The main idea of the approach was an attempt to differentiate classes of viroids according to their "*pathogens*". These virtual pathogens can replicate using the information contained in the viral sequences. As candidates for the role of such artificial pathogens, artificial neural replicators were used in [1]. In this work, a specific representations of nucleotide sequences were introduced using two incomplete binary codes and it was found that the self-reproduction of neuronal replicators differs in the corresponding two cases. It also turned out that in many cases the patterns transmitted by neural replicators - *fuzzy motifs* - can have interesting symmetry and contain useful information for further analysis.

Here we present the results of the analysis of the genome sequences of both RNA and DNA viruses, with a focus on double-stranded human papillomaviruses. We start with a brief overview of neural replicator analysis (NRA), then briefly touch on the features of replicator tables (RT) for viral genome sequences of different sizes, and devote the main part of the article to the analysis of more than 200 types of human papillomavirus. In conclusion, a discussion of the results obtained is presented.

Neural Replicators Analysis

The basic artificial neural network model used in Neural Replicators Analysis (NRA) is the self-reproducible neural network (neural replicator) [1, 3, 4]. This model includes the mechanism of synchronously changing threshold of all neurons having binary states x_i (+ 1 or − 1) in the standard

Hopfield network [2]. The remarkable phenomena observed in such a system [1, 3, 4] is that in a chain of networks after few steps of transmission a special network arises in a chain which transmit further *just those patterns which it learned from its neighbor*. So, this network produces its *exact copy*, or is self-reproducible. The self-reproducible networks are absolutely transparent ones – they show as quasi attractors all learned patterns during the cycle of threshold growth. The model [1] suggests that neurons take binary values. Though many generalizations of this model permit to avoid this restriction just such code scheme was used for genomic analysis in a previous paper [1]. In this paper *non-traditional representation* of nucleotide sequences was used. Instead of four-letter genetic code *two binary code schemes* to represent these sequences were introduced. The first code (called WS code) combines the Watson-Crick pairs (AT) and (CG) and presents them as a weak (AT) pair encoded by "–1" and a strong (CG) pair encoded by "+1". The second keto-amino (KM) code combines a wobble pair (TG) encoded with "+1" and a less stable (AC) pair encoded with "–1". These two incomplete codes were used to construct sets of networks of different sizes *K* starting from 3 with the Hebbian interconnections calculated with the use of patterns generated by sliding the nucleotide sequence consisting of *N* nucleotides with a window having a length *K. N* resulting patterns (note, that their number does not depend on *K*) are used to form the Hebbian matrices of interconnections of the two parent fully connected Hopfield networks (for WS and KM coded patterns, correspondingly). For a sliding window of the same size the source parent network can generate a nontrivial replicator with a non-empty set of the patterns for transmission, or non-replicating network with empty set of patterns for transmission. This last network cannot generate descendants or, in other words, cannot breed. The last situation is rather common for some virus types, but, in general, RTs of viruses have non-trivial form. In [1] it was demonstrated that despite a wide range of RT forms some reasonable approximate categorization of two viroid families can be derived. Other interesting phenomenon is connected to the replicator transmitted patterns - *fuzzy motifs*. It was shown [1] that patterns transmitted by replicators contain additional information and often have interesting symmetries and periodicities [1]. In this paper it is demonstrated that these symmetries can be useful for differentiation of human papillomaviruses genera. Really in this case the sets of WS-coded transmitted patterns simplify to single pattern for the network of maximal size (the replicators of large sizes do not exist) and the analysis of this set becomes easily interpreted. More details about the different additional results of the application of NRA to the study of rather short viroid genomes are presented in [1].

Application To Virus Genomes

Obviously, the approach proposed in [1] and applied to the analysis of viroids can also be applied to virus genomes. Hepatitis delta virus (HDV) has the smallest DNA genome, closely resembling the RNA genome of viroids [5], and its replicator table has a form typical of viroids, as well as a small genome, such as that of narnaviruses or mitoviruses (see Fig. 1).

The replicator tables of other hepatitis viruses have different shapes, but, as we will further see, the RTs of hepatitis A, C, and E viruses have fairly common forms for viruses with modest and large genomes. The

main feature of these viruses is that they do not have replicators generated on KM, encoded DNA or RNA sequences (Fig. 2).

Let us pay attention to the RT of the hepatitis E virus (Fig. 2E). It also lacks replicators when its RNA sequence is represented by the WS code. For hepatitis A and C, these replicators exist, but only up to a certain maximum size of the neural network: five for hepatitis A and eight for hepatitis C. Also for hepatitis A, there are all replicators of smaller network dimensions (starting from 3). We will call such RTs monotonic. In contrast, for the hepatitis C virus, there is no replicator for network size 6. We will call corresponding RTs nonmonotonic and will further use asterisks to mark corresponding virus data.

Further we can forget about right columns of RT and use only maximal size of replicators corresponding to WS code for the analysis. It can be seemed that this information is rather poor, in particular, if we take into account that RT for the virus SARS 2 isolate 2019-nCoV (WHU01 29881 bp ssRNA(+)) is the same as the RT for hepatitis A. But in such a case we can use additional information related to the patterns which use replicator networks for information transmission. For example, for hepatitis A replicators of the size 5 are built on only one pattern: 11–111, while for SARS virus on two patterns: 11–111 and − 11 − 1 -11. As we will see the forms of RT and forms of patterns can give us interesting information about virus genomes similarities and also about their divergence.

Analysis Of Human Papillomaviruses

Here, we apply the approach described above to the analysis of human papillomavirus (HPV) genomic sequences. HPVs are small, non-enveloped, double-stranded DNA viruses belonging to the Papillomaviridae family. This family was designated as a separate family, Papillomaviridae, in the 7th ICTV report [6]. The taxonomy of these viruses is usually based on the study of the nucleotide sequences of the main viral capsid protein L1 [7]. HPV types belonging to different genera share less than 60% similarity within the main capsid protein L1 sequence of the genome. Different types of viruses within the genus have 60 to 70% similarity. The new HPV type has less than 90% similarity to any other HPV type. The papillomavirus nomenclature at the species level and above is determined by the papillomavirus study group of the International Committee on Taxonomy of Viruses [8]. Human papillomaviruses are classified into 5 genera - Alpha, Beta, Gamma, Mu and Nu, containing many species and types: the number of these types increases linearly with time for the beta genus and extremely rapidly for the gamma genus - the rate of detection of HPV types increases, mainly as the result of metagenomic sequencing [9]. Here we use species and types taxonomy data provided in [10] and the relevant NCBI and GenBank references are provided in the Materials section.

Instead of RTs, which in this case do not have replicators for CM-encoded sequences for a genome size of about 8000, we will use a convenient visualization of the situation, showing only replicators with WS-code. Next, we will use colors to mark the maximum size of the N_{max} replicator neural network generated using the WS-encoded genome sequence. Thus, the situation of the absence of replicators will be marked

in black, the presence of only a replicator of size 3 in purple, the presence of a maximum size 4 in blue, the maximum size 5 in light blue, the maximum size 6 in green, the maximum size 7 in yellow, the maximum size 8 in orange and the maximum size 9 in red (see Fig. 3). It turns out that this set of colors is sufficient to characterize all replicators of maximum size for all types of human papillomaviruses studied. We will also use one or two asterisks to denote cases with non-monotonic sets of replicators (up to one maximum size), when a replicator does not exist for one or two smaller sizes, respectively. We start with the genus Alpha papillomavirus and present the results of their study on Fig. 3.

Alpha papillomaviruses

Standard characteristics of this genus is that “Alpha HPVs preferentially infect the anogenital and oral mucosa, causing both malignant and benign neoplasms. Cutaneous lesions have also been observed”. [11]. One of the oncological disease connected with Alpha genus is the cervical cancer (note, however, that HPV belonging to Beta and Gamma genera are also considered as carcinogenic cofactors of cervical cancer [12]).

There are some interesting observations that can be seen in this picture. Firstly, the Alpha-2 and Alpha-4 species, which have a large size of replicators (up to 9), differ significantly from other species of the Alpha papillomavirus genus. It is noteworthy that, unlike the types of other species, it is the types of these two species that cause the formation of skin warts (as we found in [7] “The genus alpha also includes a few cutaneous HPV types (HPV2, 3, 7, 10, 27, 28, and 57), which cause common and plantar warts " In addition, most types of high oncogenic risk (HR) are characterized by the absence of replicators (black boxes). Thus, using NRA, we can recognize a clear division of the genus Alpha into two subgenera, which was not obtained by a method based on the study of the similarity of the nucleotide sequences of the main capsid protein L1 [7].

More information can be obtained by considering patterns which are transmitted by replicators of maximal size. In all cases they transmit single patterns which for all types of species Alpha-2, Alpha-4 and also Alpha-3 are *periodic* with period equals to 2 (see Fig. 4). With only one very interesting exception (which will be further discussed) such periodic transmitted patterns are typical only to the Alpha papillomavirus genus. However, for species Alpha-14 non-periodic patterns are transmitted by replicator of the size 5. In order to clarify the situation with Alpha-14 let us consider the Beta genus of human papillomavirus.

Beta papillomaviruses

Beta HPVs cause only skin lesions and exist in a latent form in the general population, but are activated under conditions of immunosuppression [11]. Beta HPV types under the influence of certain cofactors can also trigger a malignant process. Recent studies point to the role of human papillomavirus Beta types and HPV-associated inflammation in the development of squamous cell skin cancer (the second most common non-melanoma skin cancer after basal cell carcinoma). But Beta HPV infection appears to play an important role in initiating carcinogenesis, but not in tumor progression [13]. NRA shows that the

"coloristic" of Beta papillomaviruses (Fig. 5) differs from what we observe for Alpha papillomaviruses (Fig. 3). It is characterized by a predominance of replicators with a maximum size of 5 (blue boxes), a lack of larger replicators (such as those of Alpha-2 and Alpha-4), and a small number of types without replicators at all (Fig. 5). Even more remarkable, all 5-neuron replicators transmit a single pattern that is *identical* to the non-periodic pattern of HPV90 and HPV106 types Alpha-14 (Fig. 3). So, in terms of neural replicator analysis, should species Alpha-14 be moved to the genus Beta or other genera? We can clarify this by looking at the genera Gamma and Mu (we realize that this analysis is rather rough and does not claim to draw any solid conclusions). We also note that Beta-1 is the only genus of human papillomaviruses containing types HPV8, HPV47, HPV99, which have transmitted patterns of length 4, and type HPV8 has a complex set of such patterns, including not one, but four members. This type, along with HPV5, HPV20 (Beta-1), HPV17 (Beta-2) and especially HPV28 (Alpha-2), also unique among all human papillomaviruses. It has a pattern length of 6 and is associated with a high risk of developing squamous cell skin cancer [13].

Gamma papillomaviruses

The Gamma papillomavirus genus is highly diverse, but most healthy adults chronically shed Gamma virions from apparently healthy skin surfaces. Recent metagenomic studies have nearly doubled the number of known Gamma HPV types [14]. While the Beta papillomavirus genus is related to epidermodysplasia verruciforma, patients with the WHIM syndrome (warts, hypogammaglobulinemia, infections, myelokathexis) have been found to be uniquely susceptible to Gamma HPV- associated skin warts.

Neural replicator analysis of Gamma papillomaviruses shows that they share some properties with Beta papillomaviruses, but also differ from them. Like Beta papillomaviruses, their types can form replicators with a maximum length of up to 5. More importantly, the only kind of non-periodic transmission pattern is the same as that of Beta papillomaviruses. The number of species of Gamma papillomaviruses is large and, as can be seen from Fig. 6, the proportion of Gamma papillomaviruses that do not generate neural replicators (black boxes) exceeds 60%, while for Beta papillomaviruses this figure is about 11%. Thus, we can assume that Alpha-14 papillomaviruses species are more similar to Beta, and not to Gamma human papillomaviruses.

Mu papillomaviruses

Mu papillomaviruses are among the HPV types associated with cutaneous disease. The HPV1 type is responsible for about 30% cases of common warts [15]. The results obtained for Mu human papillomaviruses show, that they are similar to those for Beta and Gamma papillomaviruses (see Fig. 7 left): the type HPV63 has the same non-periodic transmitted pattern as for the Beta and Gamma genera.

Nu papillomaviruses

The most interesting result of NRA was obtained for Nu (HPV41) human papillomavirus. Initially, this virus was isolated from a facial wart, but subsequently its DNA was found in some skin carcinomas and

precancerous keratoses [16]. The genomic sequence of this virus is most distantly related to all other types of human papillomaviruses, and HPV41 virus has been identified as the first type of new Nu genus. But NRA analysis shows that it is ideal for Alpha-2 specie because it has a maximal replicator size of 7 as well as the same periodic transmitted pattern (Fig. 7 right). The clinical manifestations of Nu virus infection are similar to those of the Alpha-2 species (although it also causes malignant skin lesions), so this result is not inconsistent with the characteristics of this genus. What is also interesting is that NRA may provide some additional information about the problem of virus transfer to another host, as well as the taxonomy of viruses. As the Van Doorslayer paper says [17]:

“Because of the absence of cross-species infections, it is unlikely that horizontal gene transfer played any role in the evolution of the Papillomaviridae. In fact, a study specifically looking at the influence of horizontal gene transfer identified only a single potential cross-species transmission event. This event involved ancestors of a porcupine (EdPV1) and human (HPV41) papillomavirus [18]. These two viruses are the only members of a divergent genus (Nu papillomavirus); it will be of interest to see how the inclusion of more viruses in this genus will affect the conclusion of cross-species infection [17]”.

In this situation, it was very interesting to use NRA to study the porcupine EdPV1 virus. It turned out that indeed it has a replicator of maximum size 6 (7 for HPV41), but the pattern transmitted by this replicator (3-periodic) differs significantly not only from the 2-period pattern of HPV41, but also from any pattern transmitted by the replicators of all papillomaviruses (see Fig. 7 right bottom). Thus, from the point of view of the NRA, the porcupine sigma virus EdPV1 cannot be combined with the HPV41 virus into one genus, nor can it be attached to any other genera of human papillomaviruses.

Materials

Below the data used in the study is presented: it contains the species name, type, NCBI and GenBank accession number and the length of the virus ssDNA are shown.

Alpha papillomaviruses

Alpha-1	HPV32	NC_001586.1	7961 bp	Alpha-7	HPV18	LC636309	7857 bp
	HPV42	LR862086	7920 bp		HPV39	LR862071	7833 bp
Alpha-2	HPV3	X74462.1	7820 bp		HPV45	EF202167.1	7849 bp
	HPV10	X74465	7919 bp		HPV59	LR862080.1	7898 bp
	HPV28	U31783.1	7959 bp		HPV68	GQ472851.1	7830 bp
	HPV29	U31784.1	7916 bp		HPV70	U21941.1	7905 bp
	HPV77	Y15175	7887 bp		HPV97	EF436229.1	7843 bp
	HPV78	AB793779	7805 bp	Alpha-8	HPV7	MK463913	8037 bp
	HPV94	GU117628	7872 bp		HPV40	X74478	7909 bp
	HPV117	GQ246950.1	7895 bp		HPV43	LR861953	8007 bp
	HPV125	FN547152.2	7809 bp		HPV91	AF419318.1	7966 bp
	HPV160	AB745694	7779 bp	Alpha-9	HPV16	NC_001526.4	7906 bp
Alpha-3	HPV61	U31793.1	7989 bp		HPV31	LR862053	7878 bp
	HPV62	AY395706.1	8092 bp		HPV33	M12732.1	7909 bp
	HPV72	X94164.1	7988 bp		HPV35	M74117.1	7851 bp
	HPV81	AJ620209.1	8070 bp		HPV52	LC373207.1	7906 bp
	HPV83	AF151983	8104 bp		HPV58	LC376008	7824 bp
	HPV84	AF293960	7948 bp		HPV67	D21208	7801 bp
	HPV86	AF349909	7983 bp	Alpha-10	HPV6	AF092932	8012 bp
	HPV87	KU298941.1	7992 bp		HPV11	HE574705	7933 bp
	HPV89	KU298945.1	8074 bp		HPV13	X62843	7880 bp
	HPV102	DQ090083.1	8078bp		HPV44	LR862067	7836 bp
	HPV114	GQ244463.1	8069 bp		HPV74	LR862050	7902 bp
Alpha-4	HPV2	MN605988.1	7859 bp	Alpha-11	HPV34	KF436817	7788 bp
	HPV27	AB211993.1	7831 bp		HPV73	LR862011	7716 bp
	HPV57	MK463925	7848 bp	Alpha-13	HPV54	HPU37488	7759 bp
Alpha-5	HPV26	NC_001583.1	7855 bp	Alpha-14	HPV71	NC_039089	8017 bp
	HPV51	KF436884	7815 bp		HPV90	NC_004104	8033 bp
	HPV69	KF436864.1	7705 bp		HPV196	DQ080082	8035 bp

	HPV82	AB027021.1	7821 bp
Alpha-6	HPV30	LR862000	7843 bp
	HPV53	NC_001593.1	7856 bp
	HPV56	LR862083	7866 bp
	HPV66	LC511686.1	7818 bp

Beta papillomaviruses

Beta-1	HPV5	JN211194	7746 bp		HPV23	U31781.1	7324 bp
	HPV8	M12737.1	7654 bp		HPV37	U31786.1	7421 bp
	HPV12	X74466.1	7673 bp		HPV38	JN211196	7397 bp
	HPV14	X74467.1	7439 bp		HPV80	Y15176.1	7427 bp
	HPV19	X74470.1	7685 bp		HPV100	FM955839.1	7380 bp
	HPV20	U31778.1	7757 bp		HPV104	FV955840	7386 bp
	HPV21	U31779.1	7779 bp		HPV107	EF42222.1	7562 bp
	HPV24	U31782.1	7452 bp		HPV110	EU410348.1	7423 bp
	HPV25	X74471.1	7713 bp		HPV111	EU410349.1	7384 bp
	HPV36	U31785.1	7722 bp		HPV113	FM955842.1	7412 bp
	HPV47	M32305.1	7726 bp		HPV120	FN598907	7304 bp
	HPV93	AY382778	7450 bp		HPV122	GQ845444.1	7397 bp
	HPV98	FM955837.2	7466 bp		HPV145	HM999997	7375 bp
	HPV99	FM955838	7698 bp		HPV151	FN77756	7386 bp
	HPV105	FM955841.1	7667 bp		HPV159	HE963025	7443 bp
	HPV118	GQ246951.1	7597 bp		HPV174	HF930491.1	7359 bp
	HPV124	GQ845446.1	7489 bp	Beta-3	HPV49	NC_001591.1	7560 bp
	HPV143	HM999995	7715bp		HPV75	Y15173.1	7537 bp
	HPV152	JF304768	7480 bp		HPV76	Y15174	7549 bp
Beta-2	HPV9	NC_001596.1	7434 bp		HPV115	FJ947080.1	7476 bp
	HPV15	X74468.1	7412 bp	Beta-4	HPV92	NC_004500.1	7461 bp
	HPV17	JN211195	7426 bp	Beta-5	HPV96	NC_005134.2	7438 bp
	HPV22	U31780.1	7368 bp		HPV150	FN677755.1	7336 bp

Gamma papillomaviruses

Gamma-1	HPV4	NC_001457.1	7353 bp	Gamma-11	HPV126	NC_016157.1	7326 bp
	HPV65	X70829.1	7308 bp		HPV136	NC_017994.1	7319 bp
	HPV95	AJ620210.1	7337 bp		HPV140	NC_017996.1	7341 bp
	HPV158	KT698168.1	7192 bp		HPV141	HM999993	7384 bp
	HPV173	KF006400.1	7297 bp		HPV154	NC_021483.1	7286 bp
	HPV205	KT698167.1	7298 bp		HPV169	JX413105.1	7252 bp
Gamma-2	HPV48	NC_001690.1	7100 bp		HPV171	KF006398.1	7261 bp
	HPV200	KP692114.1	7137 bp		HPV202	KP692116.1	7344 bp
Gamma-3	HPV50	NC_001691.1	7184 bp	Gamma-12	HPV127	NC_014469.1	7181 bp
Gamma-4	HPV60	NC_001693.1	7313 bp		HPV132	NC_014955.1	7125 bp
Gamma-5	HPV88	NC_010329.1	7326 bp		HPV148	GU129016.1	7164 bp
Gamma-6	HPV101	LR861930	7259 bp		HPV157	KT698166.1	7154 bp
	HPV103	NC_008188.1	7263 bp		HPV165	JX444072.1	7129 bp
	HP108	NC_012213.1	7149 bp		HPV199	KJ913662.1	7184 bp
Gamma-7	HPV109	NC_012485.1	7346 bp	Gamma-13	HPV128	NC_014952.1	7259 bp
	HPV123	GQ845445.1	7329 bp		HPV153	JN171845	7240 bp
	HPV134	NC_014956.1	7309 bp	Gamma-14	HPV131	NC_014954.1	7182 bp
	HPV138	HM999990.1	7353 bp	Gamma-15	HPV135	NC_017993.1	7293 bp
	HPV139	HM999991.1	7360 bp		HPV146	HM999998	7265 bp
	HPV149	GU117629.1	7333		HPV179	NC_022095.1	7228

			bp				bp
	HPV155	JF906559.1	7352 bp	Gamma-16	HPV137	NC_017995.1	7236 bp
	HPV170	JX413110.1	7417 bp	Gamma-17	HPV144	NC_017997.1	7271 bp
Gamma-8	HPV112	NC_012486.1	7227 bp	Gamma-18	HPV175	NC_038524.1	7226 bp
	HPV119	GQ845441.1	7251 bp	Gamma-19	HPV161	NC_038522.1	7238 bp
	HPV147	HM999996.1	7224 bp		HPV162	JX413108.1	7214 bp
	HPV164	JX413106.1	7233 bp		HPV166	NC_019023.1	7212 bp
	HPV168	KC862317.1	7204 bp	Gamma-20	HPV163	NC_028125.1	7233 bp
Gamma-9	HPV116	NC_013035.1	7184 bp	Gamma-21	HPV167	NC_022892.1	7228 bp
	HPV129	NC_014953.1	7219 bp	Gamma-22	HPV172	NC_038523.1	7203 bp
Gamma-10	HPV121	NC_014185.1	7342 bp	Gamma-23	HPV156	NC_033781.1	7329 bp
	HPV130	GU117630.1	7388 bp	Gamma-24	HPV178	NC_023891.1	7314 bp
	HPV133	GU117633.1	7358 bp		HPV197	KM085343	7278 bp
	HPV142	HM999994.1	7374 bp	Gamma-25	HPV184	NC_038914.1	7324 bp
	HPV180	KC108722.1	7356 bp	Gamma-27	HPV207	MK645900.1	7247 bp

Mu, Nu and Sigma papillomaviruses

Mu-1	**HPV1**	**NC_001356.1**	**7815 bp**
Mu-2	HPV63	NC_001458.1	7348 bp
Mu-3	HPV204	NC_038525.1	7227 bp
Nu	HPV41	NC_001354.1	7614 bp
Sigma	EdPV-1	NC_006951.1	7428 bp

Discussion

While recognizing the limitations of the NRA, we nevertheless believe that its application to human papillomaviruses may provide further insight into the problem of virus taxonomy, or at least provide some additional features such as the form of RT, as well as the form of transmission patterns that may be added to the characteristics of viruses, especially when performing polythetic classification. It has been shown here that this approach can be used to isolate the Alpha human papillomavirus species responsible for the formation of warts and to hypothesize the similarity of the genus Nu with this subgenus of the genus Alpha, as well as the dissimilarity between Nu of the human papillomavirus and the Sigma porcupine virus. The advantages of NRA are associated with the ability to process complete genomes without the need to know their structure, as well as without alignment [19] of nucleotide sequences of different types of viruses. We hope that NRA can indeed be a useful element in the studies of the virus genome, and the considered case of human papillomavirus is more interesting than the case of viroids [1]. Perhaps the main advantage of NRA will not be in the ordering of virus types, but rather in the classification of genera. Indeed, such a feature of transmitted patterns as their periodicity can be considered suitable for monothetic classification, but the existence of types that do not have replicators makes this difficult. On the other hand, the absence of replicators and the proportion of such “dead” non-replicating networks in the genus can be considered as a feature for the polythetic classification [20] of genera (for example, as in the case of differentiation of Beta and Gamma papillomaviruses in this work). The hard question for the NRA's approach is "Why does it work?" also requires future research. The general understanding is that neural networks allow for non-linear data transformation, which proves to be very useful in many applications, including data categorization, classification, and pattern recognition. In addition, being complex systems, they have emergent properties and exhibit emergent patterns—for example, the fine-grained patterns transmitted by replicators. In any case, the world of known viral genomes is so large that there is a large field for future research on the application of NRA and the “artificial pathogen” (neural replicator) model of genomic sequences and for the assessment of their usefulness.

Declarations

Acknowledgments

I am very grateful to Professor Marc H. V. van Regenmortel for his interest to our attempts to apply neural replicators to genome analysis and also to Professor Eörs Szathmáry for his interest to our model of neural self-reproducible networks.

Compliance with ethical standards

Conflict of interest

The author declares that he has no conflict of interest.

Ethical approval

This research did not contain studies involving human participants or animals by the author.

References

1. Ezhov AA (2020) Can artificial neural replicators be useful for studying RNA replicators? Arch Virol 165(11):2513–2529
2. Hopfield JJ (1982) Neural networks and physical systems with emergent collective computational abilities. PNAS 79:2554–2558
3. Ezhov AA, Vvedensky VL, Khromov AG, Knizhnikova LA (1991) Self-reproducible neural networks with synchronously changing neuronal threshold. In: Holden AV, Kryukov VI (eds) Neurocomputers and attention II: connectionism and neurocomputers. Manchester University Press, pp 523–534
4. Ezhov AA, Khromov AG, Knizhnikova LA, Vvedensky VL (1991) Self-reproducible networks: classification, antagonistic rules and generalization. Neural Networks World 1:52–57
5. Flores 2012] Flores R, Ruiz-Ruiz S, Serra P (2012) Viroids and hepatitis delta virus. Semin Liver Dis 32(3):201–210
6. Van Regenmortel MHV, Fauquet CM, Bishop DHL, Calisher CH, Carsten EB, Estes MK, Lemon SM, Maniloff J, Mayo MA, McGeoch DJ, Pringle CR, Wickner RB (2002) Virus Taxonomy. Seventh Report of the International Committee for the Taxonomy of Viruses" Academic Press, New-York, San Diego
7. Gheit T (2019) Mucosal and Cutaneous Human Papillomavirus Infections and Cancer Biology. Front Oncol 9:355
8. Bzhalava 2015] Bzhalava D, Eklund C, Dillner J (2015) International standardization and classification of human papillomavirus types. Virology 476:341–344
9. Bruni 2021] Bruni L, Albero G, Serrano B, Mena M, Collado JJ, Gómez D, Muñoz J, Bosch FX, de Sanjosé S (2021) ICO/IARC Information Centre on HPV and Cancer (HPV Information Centre). Human Papillomavirus and Related Diseases in the World. Summary Report
10. Van Doorslaer et al (2018) ICTV Virus Taxonomy Profile: Papillomaviridae. J Gen Virol 99:989–990
11. La Rosa G (2016) Papillomavirus. In: Rose JB and Jiménez-Cisneros B, (eds) Water and Sanitation for the 21st Century: Health and Microbiological Aspects of Excreta and Wastewater Management (Global Water Pathogen Project). (Meschke JS, and Girones R (eds), Part 3: Specific Excreted Pathogens: Environmental and Epidemiology Aspects - Section 1: Viruses), Michigan State University, E. Lansing, MI, UNESCO
12. Maria 2020] Flores-Miramontes MG, Olszewski D, Artaza-Irigaray C, Willemsen A, Bravo IG, Vallejo-Ruiz V, Leal-Herrera YA, Piña-Sánchez P, Molina-Pineda A, Cantón-Romero JC, Martínez-Silva MG, Jave-Suárez LF, Aguilar-Lemarroy A (2020) Detection of Alpha, Beta, Gamma, and Unclassified Human Papillomaviruses in Cervical Cancer Samples From Mexican Women. Front Cell Infect Microbiol 10:234

13. Tampa 2020] Tampa M, Mitran CI, Mitran MI, Nicolae I, Dumitru A, Matei C, Manolescu L, Popa GL, Caruntu C, Georgescu SR (2020) The Role of Beta HPV Types and HPV-Associated Inflammatory Processes in Cutaneous Squamous Cell Carcinoma. Hindawi J Immunol Res 2020 1:1–10
14. Pastrana 2018] Pastrana DV, Peretti A, Welch NL, Borgogna C, Olivero C, Badolato R et al (2018) Metagenomic discovery of 83 new human papillomavirus types in patients with immunodeficiency. mSphere 3:e00645
15. Doorbar 2016] Doorbar J, Egawa N, Griffin H, Kranjec C, Murakami I (2016) Human papillomavirus molecular biology and disease association. Rev Med Virol 25:2–23
16. Hirt 1991] Hirt L, Hirsch-Behnam A, de Villiers EM (1991) Nucleotide sequence of human papillomavirus (HPV) type 41: an unusual HPV type without a typical E2 binding site consensus sequence. Virus Res 18:179–189
17. Van Pocrupine 2] K (2013) Evolution of the Papillomaviridae. Virology 445:11–20
18. Pocrupine 1] Shah SD, Doorbar J, Goldstein RA (2010) Analysis of Host–Parasite Incongruence in Papillomavirus Evolution Using Importance Sampling. Mol Biol Evol 27(6):1301–1314
19. Shen-Gunther J, Cai H, Zhang H, Wang Y (2019) Abundance of HPV L1 Intra-Genotype Variants With Capsid Epitopic Modifications Found Within Low- and High-Grade Pap Smears With Potential Implications for Vaccinology. Front Genet 10:489
20. Van Regenmortel MHV (2019) Solving the species problem in viral taxonomy: recommendations on non–Latinized binomial species names and on abandoning attempts to assign metagenomic viral sequences to species taxa. Arch Virol 164:2223–2229

Figures

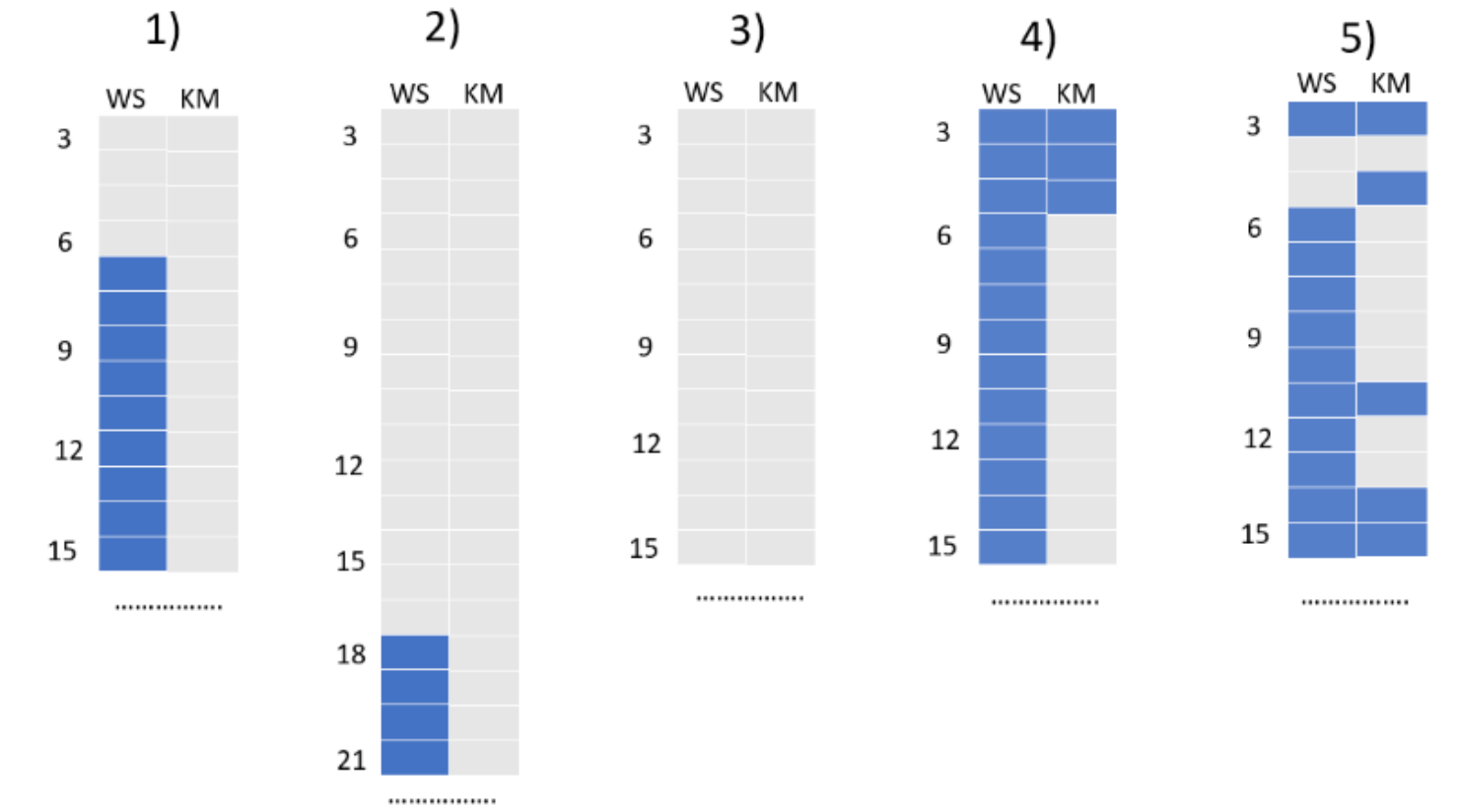

Figure 1

The examples of Replicator Tables of viruses with short genome sequencies are characterized by different forms of RT where both WS and also KM coded sequencies generate self-reproducible neural networks – replicators (gray boxes). The absence of replicators is illustrated by blue boxes. 1) *Hepatitis delta virus* genomic RNA, GenBank: D01075.1M, 1682 bp, 2) *Saccharomyces 23S RNA narnavirus* NC_004050 2891 bp, 3) *Fusarium poae narnavirus 1* NCBI Reference Sequence: NC_030865, 2297 bp, 4) *Ophiostoma mitovirus 5* NC_004053.1 2474 bp 5) *Binucleate Rhizoctonia mitovirus K1 isolate* NC_027921.1 2794 bp

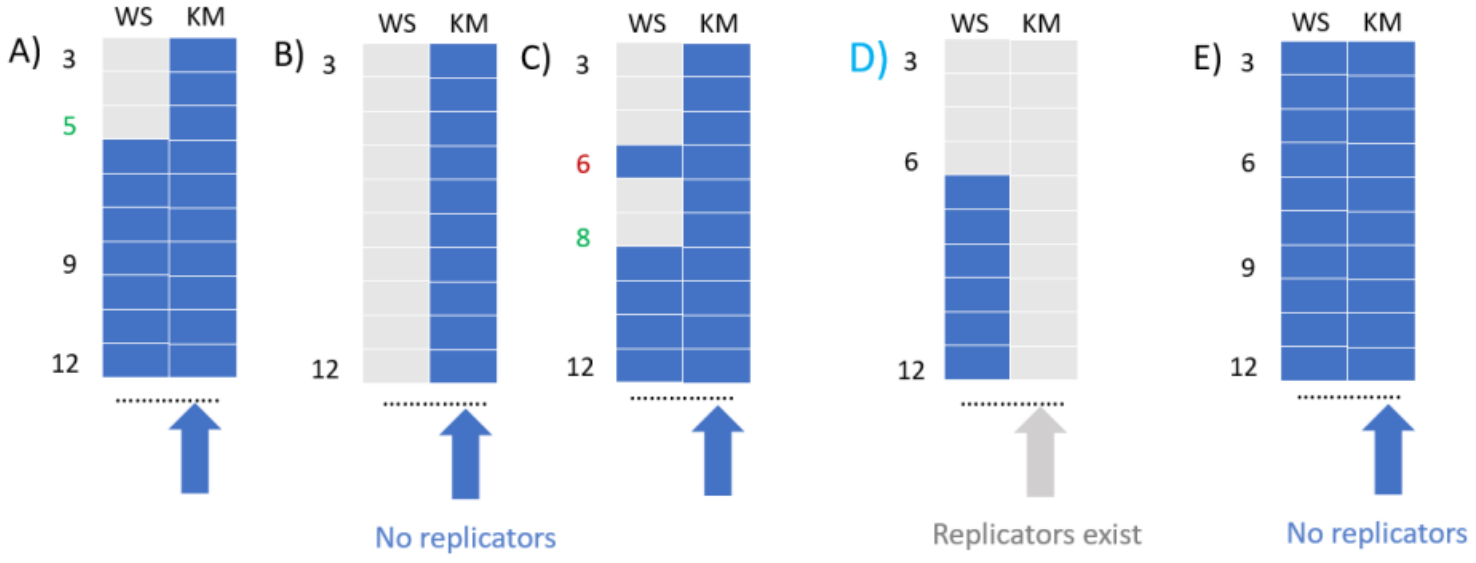

Figure 2

Replicator tables of hepatitis viruses. With the exception of delta viruses (D), all other viruses (A,B,C,E) are characterized by the absence of replicators built using KM-encoded genomic sequences (only blue boxes in the right columns). A) *Hepatitis A isolate p16 virus* genomic RNA, GenBank: KP879217.1, 7476 bp, B) *Hepatitis B virus isolate MT,* GenBank: KC492739.1, 3215 bp, C) *Hepatitis C virus genotype 1,* NCBI Reference Sequence: NC_004102, 9646 bp, D) *Hepatitis delta virus* genomic RNA, GenBank: D01075.1M, 1682 bp, E) *Hepatitis E virus,* NC_001434, 7176 bp.

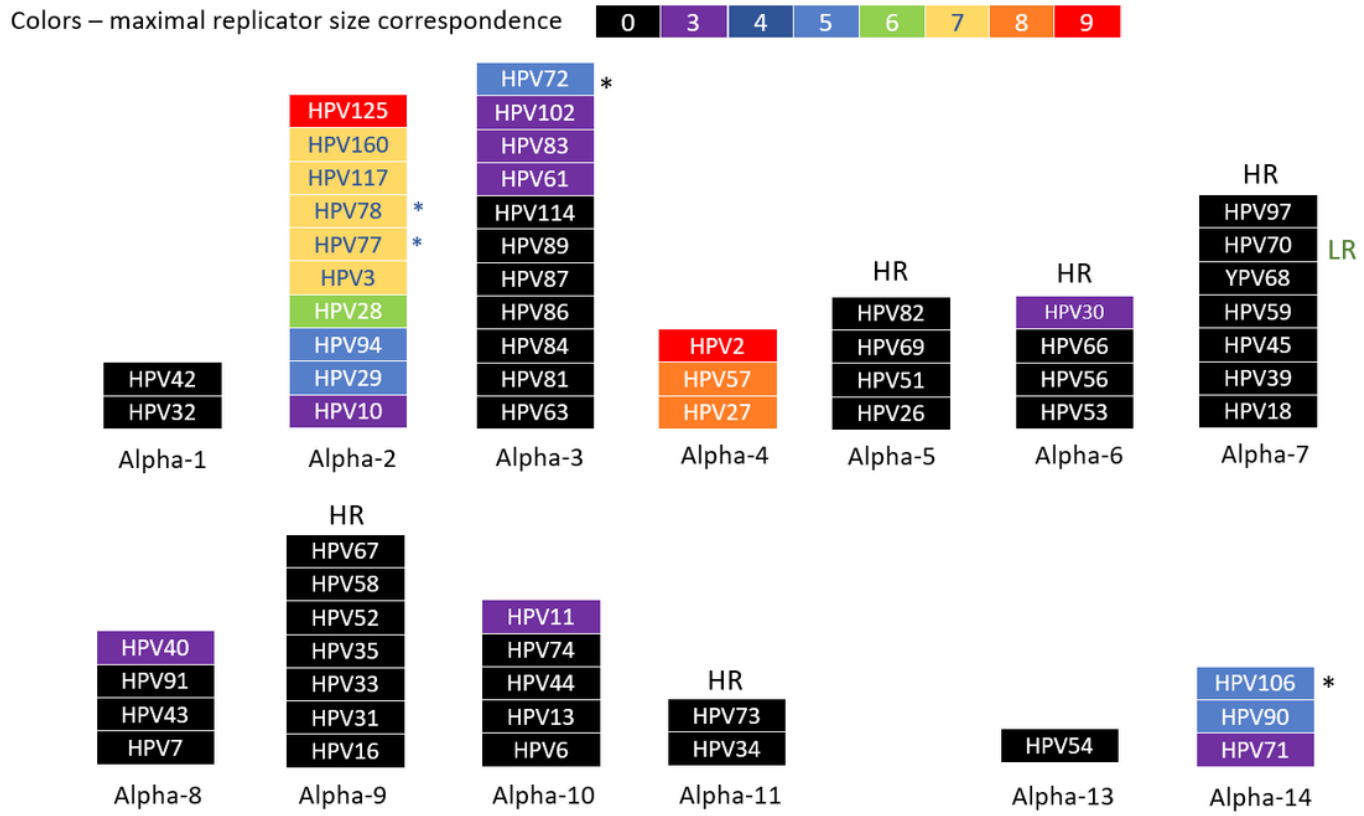

Figure 3

The maximal size of replicators for different species and types of the genus Alpha of human papillomaviruses. The sets of replicator patterns of the types HPV77, HPV78, HPV72 and HPV106 are non-monotonic (replicator of one of the size lower than the maximal size does not exist – corresponding colored boxes are marked by the asterisk)

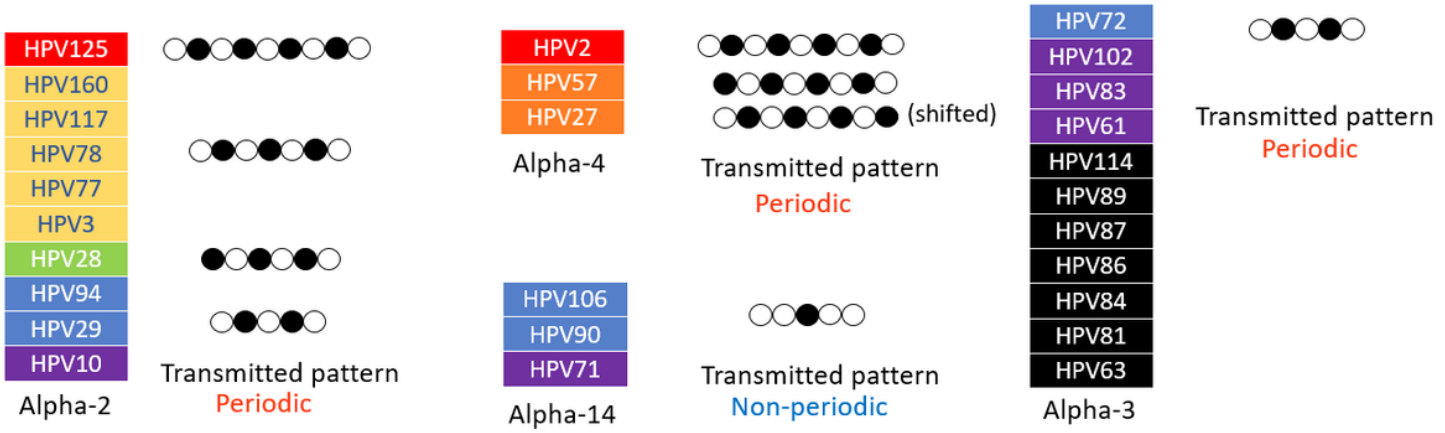

Figure 4

The single patterns transmitted with replicator networks of maximal size are presented. Neuron states (pattern binary components) are represented by black (state equals to −1) or white (+1) circles. Transmitted patterns corresponding to the Alpha-2,3,4 species are periodic with period equals to 2 (note, that patterns of HPV27 and HPV57 are complementary – shifted by one position). On contrary, the pattern transmitted by replicators corresponding to types HPV90 and HPV106 belonging to Alpha-14 is not periodic.

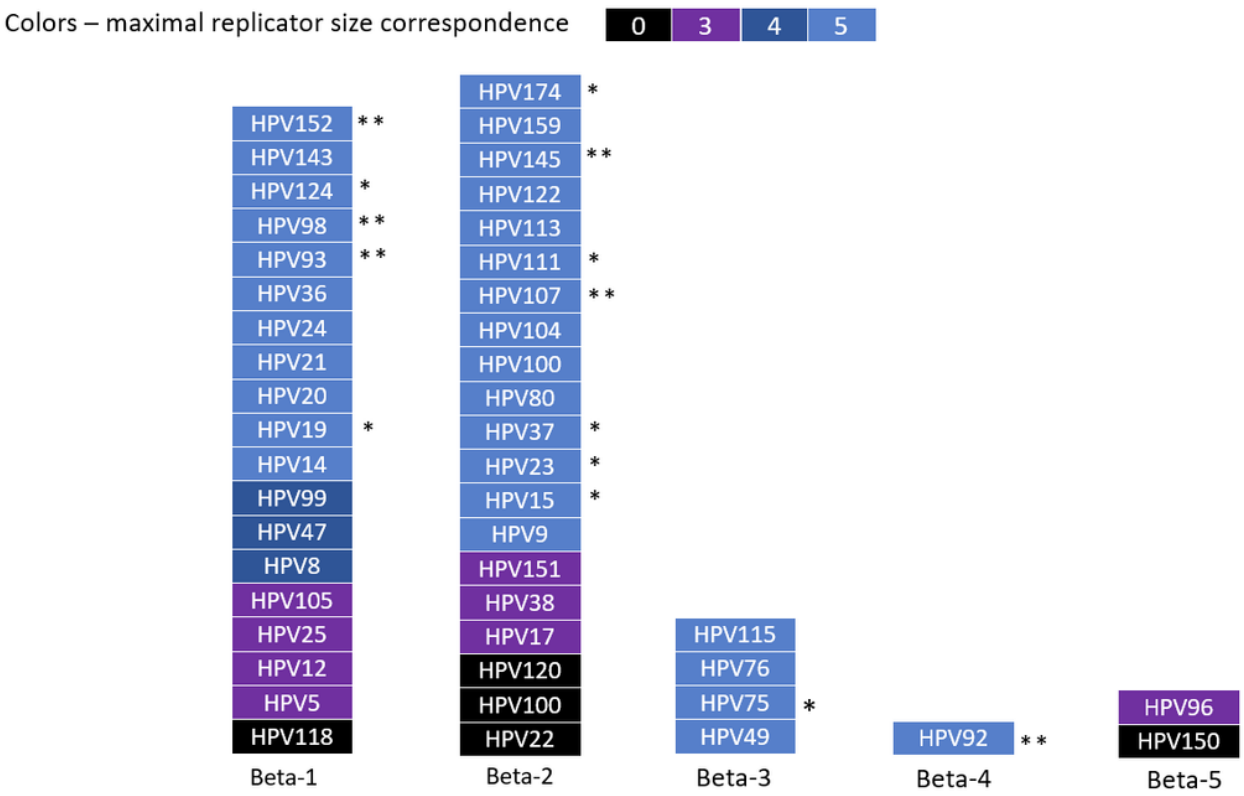

Figure 5

The maximum size of replicators for different species and types of the genus Beta human papillomavirus. The sets of many replicator patterns are non-monotonic: one (one asterisk) or two (two asterisks) replicators of length less than the maximum do not exist. Pay attention to a small part of virus types for which there are no replicators (black boxes).

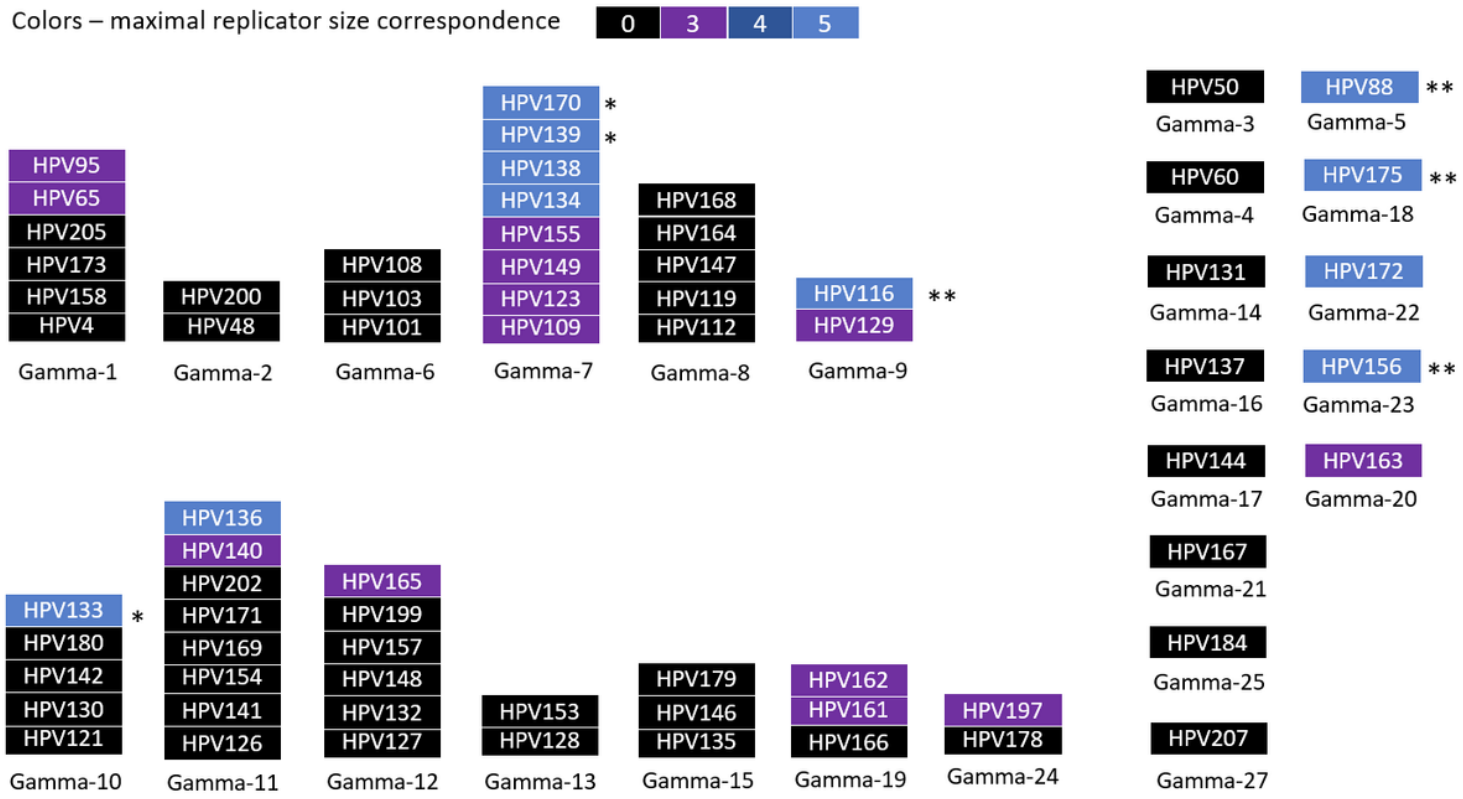

Figure 6

The maximal size of replicators for different species and types of the Gamma genus of human papillomaviruses. The sets of some replicator patterns are non-monotonic, with one (one asterisk) or two (two asterisks) less than maximum replicators missing. Pay attention to most of the types of viruses for which there are no replicators (black boxes).

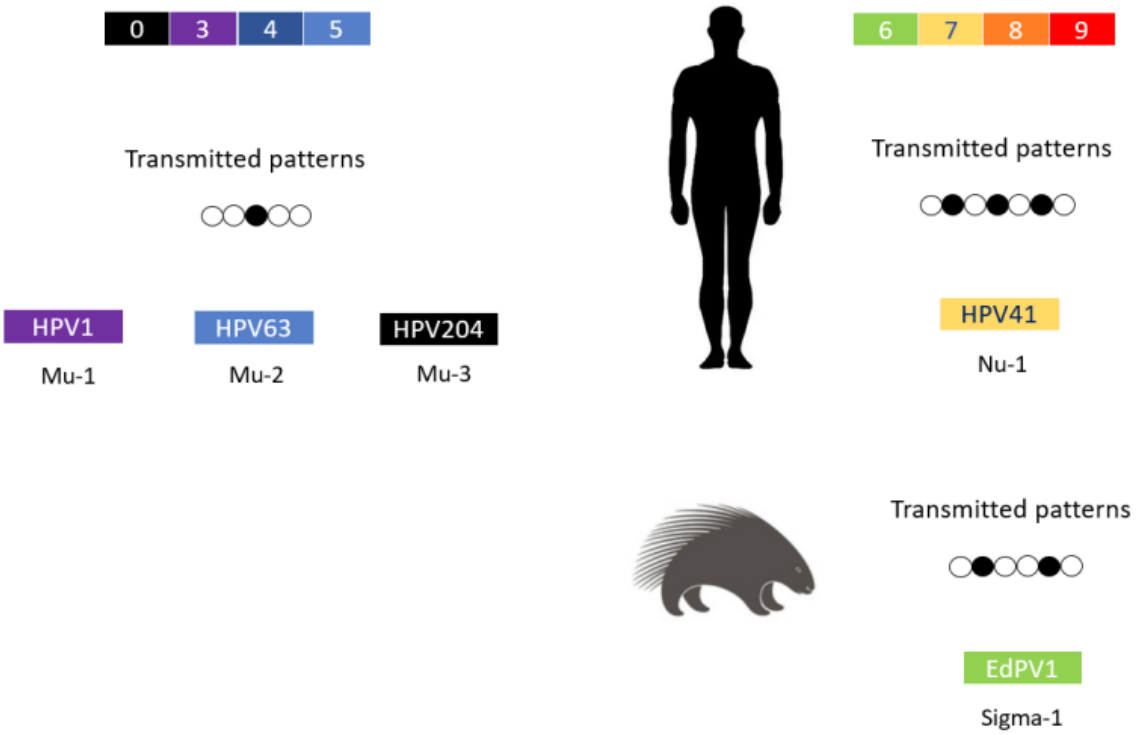

Figure 7

Maximum size of replicators for different species and genus types Mu (top left) and Nu (top right) of human papillomavirus and porcupine papillomavirus EdPV1 (bottom right). The patterns transmitted by replicators are presented. The pattern of the EdPv1 virus differs both from the 2-period pattern of human papillomavirus Nu (and Alpha genus) and from the aperiodic pattern of the Mu, Beta and Gamma genera.