Creation perspective of nucleocytoplasmic large DNA viruses

Jean O'Micks

The large-sized genomes and unique gene content of nucleocytoplasmic large DNA viruses (NCLDVs or giant viruses (short: giruses)) have become the focus of attention in recent research. Some evolutionists claim with much fanfare that these viruses form a new 'fourth domain' of life besides eukaryotes, bacteria, and archaea. They believe them to have degenerated from the eukaryotic nucleus, and that they harbour genes from eukaryotes, bacteria, and archaea through horizontal gene transfer. The presence of genes from cellular organisms suggests that giruses actually could be degenerate bacteria. Evolutionists delineate about 50 core genes with varying distribution in the genomes of seven NCLDV families, purportedly demonstrating their monophyly. Upon closer examination this is not warranted, since only 14 of these genes are unique to NCLDVs, whereas other genes are most probably found in bacteria. NCLDVs also contain a high ratio of ORFan genes, without homologs in other species, supporting their independent origin apart from other organisms. Thus, the fact that most girus genes might not have originated from one of today's three cellular domains helps further the spectrum of intelligent design.

During the past 15 years, a number of genetically diverse microorganisms with interesting characteristics have been discovered, with implications for the creation/evolution debate. They are called nucleocytoplasmic large DNA viruses, or NCLDVs,[1] since their life-cycle is attached mainly either to the cytoplasm or the nucleus of host cells. According to one taxonomical division, they have been grouped into the order Megavirales.[2] They are peculiar in that they have large genomes (ranging from 0.1–2.5 Mbp),[3,4] with up to 2,500 coding sequences; even surpassing those of bacterial or even eukaryotic species, and share certain characteristics with cells. Due to their size, their proteins are incapable of self-assembly, which denotes that they need complex proteins in order to self-assemble.

Their genomes can be made up of both DNA and RNA. Their genomes contain ORFs numbering in the hundreds, which also encode enzymes, such as ones which take part in sugar metabolism. Just like Russian nested dolls, some of these viruses themselves harbour viruses. For example, a variant of the mimivirus, called mamavirus, contains a small, 50-nm-size virus with a couple of dozen genes, called Sputnik.[5]

NCLDVs are classified into seven families, based on virion morphology and host range, and are listed and characterized in table 1. According to other classifications, giant viruses, include viruses which exceed 500 Kbp in genome size. Other DNA viruses with genomes in the size range of 100–280 Kbp are called large DNA viruses (ascoviruses, asfarviruses, baculoviruses, herpesviruses, iridoviruses and some bacteriophages), whereas the ones with very large genomes are called giant viruses, or 'giruses'. Newly discovered viruses such as the Pandoraviruses and Pithoviruses are being considered as new families,[3] although, according to some studies, Pandoraviruses are derived phycodnaviruses.[6] Some NCLDVs also contain introns and inteins, which is not characteristic of viruses.

Some evolutionists claim that NCLDVs predates the origin of the eukaryotic cell, and serve as precursors to the eukaryotic nucleus.[7] NCLDVs also have an important role in horizontal gene transfer (HGT) between species.[8] Because of these peculiar characteristics, scientists are designating these interesting viruses to a new, fourth domain of life besides eukaryotes, bacteria, and archaea, thus broadening the classical conception of viruses which were originally defined as subcellular infectious particles.

However, some NCLDV species, such as Mimivirus, contain genes only found in soil bacteria. Thus it might be that NCLDVs are not really viruses but rather degenerate bacteria which acquired viral genes, such as viral capsid proteins. For example, Mimivirus contains a number of genes which are characteristic of only cellular organisms, such as aminoacyl-tRNA synthase;[9] a vacuolar sorting-associated protein, a Cu/ZN superoxide dismutase, a UDP-N-acetylglucoseamine2-epimerase, a dTDP-4-dehydrorhamnose reductase, a dTDP-d-glucose 4-6 dehydratase, and an ExoV-like protein.[10] Fischer *et al.*[11] report 14 genes from *Cafeteria roenbergensis* virus, which resemble bacterial genes, and of which seven are involved in carbohydrate metabolism. If NCLDVs were really viruses, the presence of cellular genes truly would be an inexplicable anomaly. Other genes include topoisomerase IA, IB, and IIA, which are involved in unwinding DNA

Table 1. Characteristics of the seven families of NCLDVs

Family	No. of genera	Genome size range	Number of genes	Hosts	Replication origin
Ascoviridae	1	119–186 Kbp	99–110	Insects	Nucleus and cytoplasm
Asfarviridae	1	170–182 Kbp	151	Mammals, dinoflagellates	Cytoplasm
Iridoviridae	5	102–212 Kbp	130–328	Insects, fish, amphibians	Nucleus and cytoplasm
Mimiviridae	2	617 Kbp–1.3 Mbp	444–457	Amoeba, algae	Cytoplasm
Marseilleviridae	1	346–368 Kbp	95–463	Amoeba	Cytoplasm
Phycodnaviridae	5	154–407 Kbp	150–886	Algae	Nucleus and cytoplasm
Poxviridae	14	134–359 Kbp	544–1120	Mammals, birds, reptiles, insects	Cytoplasm

during replication. These genes are found in Mimivirus as well as *Pseudomonas*, *Agrobacterium*, and *Sinorhizobium* species.[12]

The designation of these organisms is problematic; they are called viruses in the scientific literature, and will technically be called such in this paper, but we maintain that these organisms are most likely degenerate bacteria.

NCLDV ORFan genes

According to Claverie and Ogata,[13] "the disturbing fact that most girus genes might not have originated from one of today's three cellular domains only helps revive the spectrum of intelligent design". It has been reported that in several NCLDV species a large portion of their several hundred genes have no known functional homologues.[14] These genes are called ORFan genes (genes without homologues in other lineages), and their distribution is restricted to closely related species. The vast majority of ORFans are exclusive to a single virus family only. Three-D protein structure analyses demonstrate that many ORFans encode expressed proteins, although they do not contain known protein folds. For example, 300 of the 911 Mimivirus proteins have no homologs with any other protein, and only 21 were assigned recognizable structures.[15] Ogata and Claverie[12] have demonstrated that these ORFs show the same position-dependent nucleotide statistics as the rest of the genome, suggesting that these ORFs are characteristic of the host virus, and not a result of HGT. This deals a particularly deadly blow to evolution (confirming Claverie's fears) since here we have tons of unique genes which are not a result of HGT. Accumulating evidence also shows that at least some viral genes are only less similar with their host counterpart genes. Indeed, less than 35 genes from the seven NCLDV families are a result of HGT, and less than 15 in the great majority of species.[8,16] All of this supports the idea that these organisms all have independent origins.

In Pandoraviruses, 93% of ORFs have no recognizable homologs; in fact, even now evolutionists do not have a clear idea as to what other virus Pandoraviruses are related to.[4] In general, the percent of ORFan genes in different NCLDV subgroups ranges from 2.8–75.2%, with an average of 30%[17], which is significantly higher than those in bacteria (9%). Marine virome studies show that 91% of marine viral genes are new.[18]

Distribution of ORFan genes across different NCLDV families

Boyer *at al.*[16] studied the percent of ORFan genes per NCLDV family, and found that the largest number of new genes comes from newly discovered viral families, such as Marseillevirus,[19] with up to 70% of its genes being ORFans. They also found, for example, that 2.6% of the genes in the PBCV-NY2A NCLDV genome are species-level ORFans, but 36.2% of them are ORFans at the genus level. This would indicate that for these NCLDV species, the genus is approximately equal to the baraminic boundary. In Mimivirus, only 298 of its 1,262 ORFS (24%) could be associated with functional attributes, compared to 70% in bacteria and archaea.[20] Evolutionists could claim that with the discovery of newer and newer genes and NCLDV species, the proportion of ORFan genes may decline; however, Yin and Fischer[21] found that the proportion of ORFan genes is stable, despite the increasing number of sequenced genomes, and does not depend on genome size. Table 2 shows the percentage of homologs per total proteins for NCLDVs in the COG database for each of the 49 NCLDVs in this study.

Table 2. Percentage of homologs per total proteins for NCLDVs in COG database. Proteins were blasted against protein sequences from the Uniprot website for 10 major taxonomic categories: archaea, bacteria, fungi, human, invertebrates, mammals, plants, rodents, vertebrates, and viruses. A maximum e-score cutoff of 1e-4 was applied to determine homology.

Species	Family	No. homologs	No. proteins	Homolog / protein %
Invertebrate iridescent virus 3	Iridoviridae	117	125	93.6
Acanthamoeba polyphaga mimivirus	Mimiviridae	876	979	89.48
African swine fever virus	Ascoviridae	142	160	88.75
Acanthamoeba castellanii mamavirus	Mimiviridae	872	988	88.26
Vaccinia virus	Poxviridae	183	223	82.06
Frog virus 3	Iridoviridae	75	99	75.76
Myxoma virus	Poxviridae	97	169	57.4
Yaba-like disease virus	Poxviridae	82	151	54.3
Wiseana iridescent virus	Iridoviridae	90	193	46.63
Squirrelpox virus	Poxviridae	63	141	44.68
Acanthamoeba polyphaga moumouvirus	Mimiviridae	377	891	42.31
Megavirus chiliensis	Mimiviridae	473	1120	42.23
Orf virus	Poxviridae	54	130	41.54
Canarypox virus	Poxviridae	131	328	39.94
Molluscum contagiosum virus subtype 1	Poxviridae	63	163	38.65
Invertebrate iridescent virus 6	Iridoviridae	174	467	37.26
Nile *crocodilepox* virus	Poxviridae	42	173	24.28
Singapore grouper iridovirus	Iridoviridae	38	161	23.6
Spodoptera frugiperda ascovirus 1a	Ascoviridae	17	122	13.93
Trichoplusia ni ascovirus 2c	Ascoviridae	22	163	13.5
Cafeteria roenbergensis virus BV-PW1	Mimiviridae	69	544	12.68
Paramecium bursaria Chlorella virus NYs1	Phycodnaviridae	45	374	12.03
Micromonas sp. RCC1109 virus MpV1	Phycodnaviridae	29	244	11.89
Ostreococcus tauri virus 1	Phycodnaviridae	24	230	10.43
Ostreococcus lucimarinus virus OlV1	Phycodnaviridae	26	250	10.4
Bathycoccus sp. RCC1105 virus BpV1	Phycodnaviridae	21	203	10.34
Heliothis virescens ascovirus 3e	Ascoviridae	18	179	10.06
Mythimna separata entomopoxvirus 'L'	Poxviridae	30	306	9.8
Infectious spleen and kidney necrosis virus	Iridoviridae	12	125	9.6
Lymphocystis disease virus - isolate China	Iridoviridae	22	238	9.24
Amsacta moorei entomopoxvirus 'L'	Poxviridae	27	293	9.22
Micromonas pusilla virus SP1	Phycodnaviridae	22	242	9.09
Ostreococcus virus OsV5	Phycodnaviridae	23	264	8.71
Organic Lake phycodnavirus 1	Phycodnaviridae	27	398	6.78
Phaeocystis globosa virus 14T	Phycodnaviridae	27	433	6.24
Phaeocystis globosa virus	Phycodnaviridae	27	434	6.22

Phaeocystis globosa virus 12T	Phycodnaviridae	27	439	6.15
Anomala cuprea entomopoxvirus	Poxviridae	16	263	6.08
Melanoplus sanguinipes entomopoxvirus	Poxviridae	16	267	5.99
Organic Lake phycodnavirus 2	Phycodnaviridae	19	326	5.83
Ectocarpus siliculosus virus 1	Phycodnaviridae	12	240	5
Lausannevirus	Marseillevirus	20	442	4.52
Acanthocystis turfacea Chlorella virus 1	Phycodnaviridae	34	860	3.95
Marseillevirus	Marseillevirus	16	428	3.74
Feldmannia species virus	Phycodnaviridae	5	150	3.33
Pithovirus sibericum	Unassigned	14	466	3
Emiliania huxleyi virus 86	Phycodnaviridae	14	472	2.97
Pandoravirus dulcis	Pandoraviridae	32	1487	2.15
Pandoravirus salinus	Pandoraviridae	31	2543	1.22

Therefore, in order to study this, we examined the percentage of ORFans in the genomes of 53 NCLDV genomes studied by Boyer. We plotted the average proportion of ORFans per NCLDV genus as a function of the number of genomes studied per family. The result can be seen in figure 1. We fitted a curve to the points on the graph and found that a power law best describes the relationship between the number of species within a family and the average proportion of ORFans to follow the following equation: $y = 59.251\ x^{-0.301}$, with a correlation coefficient of 0.79. Based on this model, as an example, with a baramin of 1,000 members, it can be expected that 7.4% of the genes within the baramin will be ORFan genes. This means that even with baramins with a high number of members, the number of ORFan genes tends to approximate an asymptotic value, meaning that there will always be a minimum number of family-specific ORFan genes constituting a significant portion of NCLDV genomes, which do not originate from other species.

The 20,086 protein sequences for 49 NCLDV species were downloaded from the COG website and compared to protein sequences from archaea, bacteria, and eukaryotes from the Swissprot database. For each of the 49 species we calculated the percent ORFan proteins they had in their genome. On average, 75.4% of their proteins (e-score 1^{-40}) did not have homologs with any other protein in the Swissprot database (being ORFans), similar to other results.[16]

In figure 2 we can see the average percentage of ORFs as a function of the negative logarithm of the cutoff e-value for orthologous hits between proteins. The curve follows the equation $y = -0.0026\ x^2 + 0.5283\ x + 58.784$, and has a correlation coefficient of 0.9918. As we can see, the curve steadily increases as the cutoff e-value becomes tighter (an ever decreasing e-value, which corresponds to an ever increasing neglog value). Even at a neglog e-value of 5, the average ORF content is 56%.

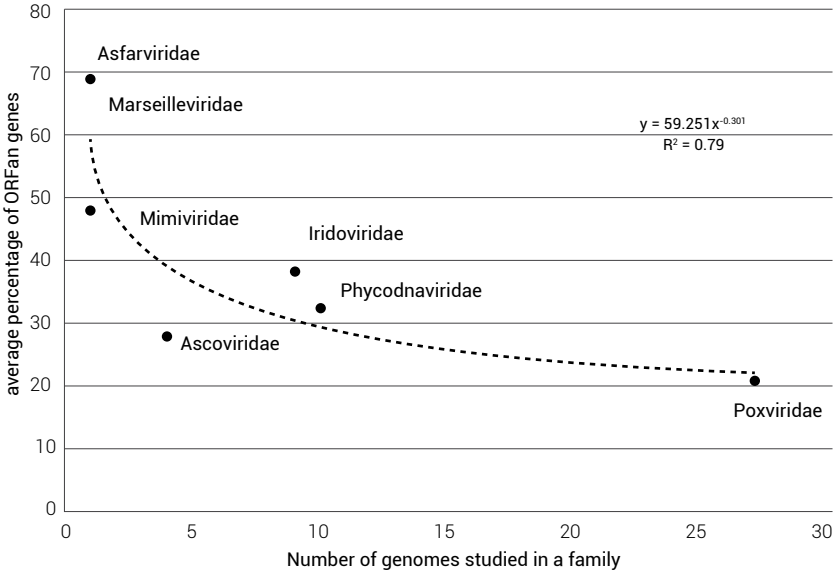

Figure 1. Average percentage of ORFan genes as a function of the number of genomes studied in a given NCLDV family. The points for the families Asfarviridae and Marseilleviridae overlap each other. The curve tapers off to the right, which shows that even with a high number of genomes in a given family, a substantial portion of the genes remain ORFans.

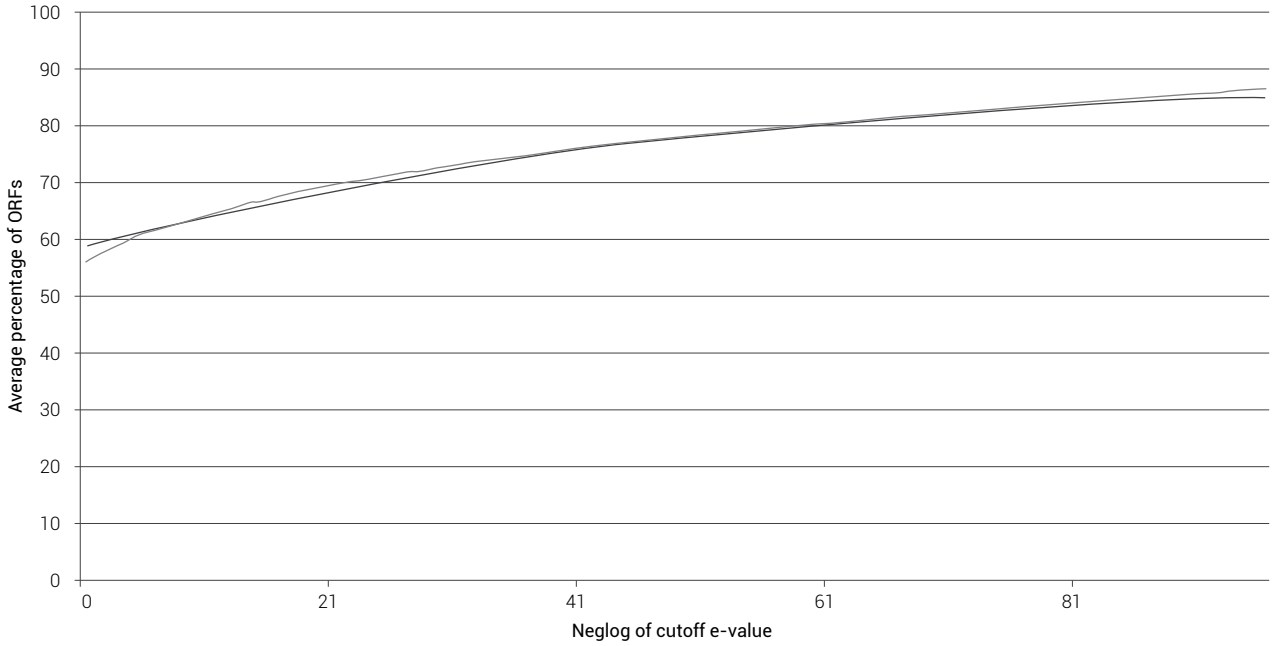

Figure 2. Average percent of ORFs in the 49 NCLDV species as a function of the negative logarithm of the cutoff e-value for orthologous proteins

A closer look at the common NCLDV core gene set

Yutin and Koonin[22] described the phylogenetic distribution of the 50 core NCLDV genes in Megavirales, and found that not all members of NCLDVs contain all of these genes. These genes are thought to be important for the basic replication machinery, and that they were present in a *hypothetical last common ancestor* of NCLDVs. These genes are involved in DNA replication, recombination and repair, transcription and RNA processing, nucleotide metabolism, virion structure, signal transduction, virus-host interactions, and also in other uncharacterized processes. However, other authors point out that NCLDVs are missing genes for translation systems, such as aminoacyl-tRNA synthetase, and translation factors, such as EF1-a, eIF-4a, eIF1, and SUA5.[23]

Yutin and Koonin[21] state that phylogenetic trees failed to show an NCLDV clade, and that deviations from simple vertical evolution probably occurred in almost all of the core genes. Indeed, only 14 of

Table 3. List of genes from Yutin and Koonin[22] which are missing from certain NCLDV subgroups

Gene/gene group	General functional group	Reason for exclusion
ATP-dependent ligase	DNA synthesis	polyphyletic
capping enzyme	mRNA synthesis	present in only one species of iridoviruses
DNA polymerase	viral replication	present in only some phycodnaviruses
dUTPase	nucleotide metabolism/ repair	present only in poxviruses, iridoviruses, and phycodnaviruses
FLAP nuclease	DNA synthesis	present in only poxviruses
polyA polymerase large catalytic subunit	mRNA synthesis	present in only one species of mimivirus
polyA polymerase small regulatory subunit	mRNA synthesis	present only in poxviruses
primase-helicase	viral replication	present in only some phycodnaviruses
ribonucleotide reductase (RR)	mRNA synthesis	present in only poxviruses and iridoviruses have different affinities
RNA polymerase (RNAP)	RNA synthesis	present in only majority of phycodnaviruses
thymidine kinase (TK)	dNTP synthesis	missing from some species across supposed NCLDV clade
thymidylate kinase (TMPK)	dNTP synthesis	missing from some species across supposed NCLDV clade

the 50 genes listed in table 1 of their paper were shown to be common to all NCLDVs. The 36 other genes were either thought to be polyphyletic, too divergent in sequence, missing from a number of subgroups, or acquired from other organisms by lateral transfer.[8,24] Table 3 lists groups of genes which are missing from some NCLDV subgroups along with their function. While it could be true that gene loss occurred in multiple NCLDV subgroups, it is also possible that these genes were never lost in the first place, but rather that NCLDVs are polyphyletic, forming different baramins within a single NCLDV apobaramin. Of course, the question can still be raised, if some of these important 50 genes needed for viral replication are missing from some species, then it must follow that they aren't necessary for viral replication in the first place. The authors also apply faulty logic in assuming that the monophyly of NCLDVs is the most appropriate null hypothesis, which they were unable to reject at a statistically significant level. In statistical hypothesis testing it is easier to reject a hypothesis than to prove it to be true.

Yutin *et al.*[22] studied the number of shared gene families as well as the Jaccard similarity (a measure of gene content similarity between two organisms) of gene complements in Iridoviridae, Marseilleviridae, Phycodnaviridae, Mimiviridae, and Poxviridae. The largest Jaccard similarity they found was 36% between *Acanthamoeba polyphaga* and *Megavirus chiliensis*, 17% between *Phaecocystis globosa* and Organic Lake phycodnavirus, 11% between Invertebrate iridescent virus and Lymphocystis disease virus, and 11% between *Amsacta moorei* entomopoxvirus and Vaccinia virus. For example, three viruses from the family Phycodnaviridae, PBCV-1, EhV, and EsV have only 14 genes in common (D5-type ATPase, DNA polymerase, A32-type ATPase, A18-type helicase, a capsid protein, a thiol-oxireductase, D6R-type helicase, a Ser/Thr protein kinase, a VLTF2-like transcription factor, a proliferating cell nuclear antigen, a ribonucleotide reductase large and small subunit, an A494R-like uncharacterized protein, and a group III thioredoxin/glutaredoxin),[25] whereas combined they have over 1,000 different genes, meaning that the Phycodnaviridae family itself can be broken down into separate baramins. Six strains of *Chlorella* viruses (NY-2A, AR158, MT325, FR483, PBCV-1, and ATCV-1) however have 80% of their genes in common, meaning that it is highly likely that they all belong to the same baramin. Indeed, common gene content may serve as a good marker for inclusion of NCLDVs into the same baramin. For example, in the case of the previously mentioned six *Chlorella* species, this must mean that a majority of genes resist genetic deletion, thus they must have some important function. Also, the genome of the white spot shrimp virus (WSSV),[26] from the family Nimaviridae, is dissimilar to any other virus, questioning its monophyletic relationship within the NCLDVs.[27] Of its 531 genes, only 45 have a higher than 20% similarity to any other known protein. It is the only eukaryotic virus genome to encode a collagen-like protein.

Furthermore, many evolutionists hold that despite increases or decreases in gene content or genome size, the size of the ancestral archaea or bacterial genome was not much different than their modern descendants.[28] For example, Iyer *et al.*[23] claim that due to the presence of SWI2/SNF2-like chromatin-remodelling ATPases of helicase SFII, the ancestral NCLDV chromosome was fairly large, in need of supercoil regulation. This obviously raises the question, if the ancestor of all NCLDVs is so similar to modern NCLDVs, then when did evolution happen?

Only 6.1% of Marseillevirus ORFs belong to the core NCLDV gene set.[16] Pandoraviruses are thought by evolutionists to have a distant relationship with the 7 families of NCLDVs, yet they have only 17 of the 50 core NCLDV proteins, which is all the more significant as the two viruses studied from this group (*Pandoravirus dulcis* and *salinus*) have 1,487 and 2,543 genes, respectively, the most of any giant virus.[21] Other giant virus families, such as Myoviridae, Nimaviridae, Herpesviridae, and Polydnaviridae, have large genome sizes, but their gene content precludes them from being classified as NCLDVs[2] due the evolutionary mispreconceptions that in order for all NCLDVs to be monophyletic they all have to have the same set of core genes.

Genome size variation in NCLDVs

As described in a previous work on bacterial genome decay,[29] NCLDVs also undergo a similar process involving gene loss. These species include poxviruses, African swine fever virus, and different species of chlorella viruses, in the range of 8–37 Kbp. For example, Mimivirus in *Acanthamoeba polyphaga* cultures can lose 17% of its genome, from 1.2 Mbp to 0.993 Mbp. This process also involves losing fibres from its surface.[30] These deletions covered 155 coding sequences, some of them duplicated genes (therefore unnecessary), and also included two uncharacterized genes from the set of core NCLDV genes, suggesting that these two genes are not absolutely necessary for function. A further 205 genes had gaps in them, being either deleted or turned into pseudogenes. This is remarkable, since in its original state the Mimivirus genome has no pseudogenes,[7] meaning that here pseudogenization was a completely downhill process. Some of these genes were involved in DNA replication and recombination, RNA processing, and translation. Some evolutionary theories have it that viruses with giant genomes acquired a lot of genes over evolutionary time from viruses with smaller genomes,[23]

yet here we have substantial downsizing of the Mimivirus in cell cultures, an evolutionary blink of an eye. This is proof that genome decay goes very fast, and hints at a recent origin, just as predicted by Terborg's baranome hypothesis,[31] which predicted the genomic breakup and decay of related organisms with a single pan-genome.

Interestingly, several NCLDV genera besides Mimiviridae have species which have large genome-size discrepancies—for example, the two *Pandoravirus* species, *dulcis* and *salinus*, with genome sizes of 1.9 and 2.5 Mbp, respectively.[4] The *Feldmannia* algal virus has two variants with different genome sizes, which are 158 and 178 Kbp, respectively.[32] Similar differences have been reported in two land species, *Arabidopsis thaliana* and *lyrata*.[33] In the current study, of 10 groups, we found several of them also showed a large within-group variation in genome size. For example, in the third group with species from Mimiviridae, genome size varied by 0.24 Mbp. In the fourth group, corresponding to ChPV species, the genome size ranged from 140 to more than 2.5 times its size, 360 Kbp, and in the fifth group (EPVs), the genome size ranged from 232 to 281 Kbp. According to Lefkowitz *et al*.,[34] "gene loss is a major mechanism responsible for genome diversity in the *Poxviridae*, and that acquisition of new genes has played essentially no role in determining the biology of individual species in the [orthopoxvirus genus]".

This means that large gains/losses in closely related NCLDV genomes (within a single group) are possible without upsetting species boundaries. Rapid large-scale genome size variation between two similar species is not what evolutionary theory predicts.

Patterns of genome decay involve loss of genes at the edges of chromosomes necessary for genetic variability, such as those which determine host-pathogen interaction, whereas more conserved housekeeping genes, such as those which are needed for replication, are located at the centre of the chromosome. Genes that have been acquired via HGT are also located at the end of the chromosome of the NCLDV.[35]

Summary and conclusion

A major question that needs to be addressed is, what kinds of organisms are NCLDVS exactly, and how did they originate? The Bible does not mention bacteria or viruses specifically, so therefore we are assuming that if God created different kinds of macroscopic organisms, then the same kind of logic can be applied to microorganisms. Therefore, the results of the analysis presented here are somewhat tentative. If NCLDVs are viruses, then the question would arise as to why God would create such pathogenic viruses in a good world. However, it is well known that there are both harmless 'passenger' viruses and harmless bacteria that do not harm their hosts.

The high number of ORFan genes in NCLDVs is significant because it means that these organisms harbour hundreds of genes which, if they are not homologous to known genes from the other three domains of life (eukaryotes, bacteria, and archaea), also must have originated independently from the main evolutionary tree of life. The high proportion of ORFan genes in NCLDV genomes has still held, despite the increase of newer genes in public databases and the discovery of newer species of NCLDVs. NCLDVs will always contain species or genera-specific genes. This, in turn, means that NCLDVs form their own apobaramin, separate from all other organisms. The high proportion of ORFans specific to NCLDV genera implies that this is the taxonomical limit to these virus species, and that an NCLDV genus corresponds to a biblical holobaramin. Despite their lack, or low content of core NCLDV genes, we suggest that the families Pandoraviridae, Myoviridae, Nimaviridae, Herpesviridae, and Polydnaviridae also be classified into the NCLDV/Megavirales apobaramin. This way the classification of these species is not forced unnecessarily into an evolutionary system.

Interestingly, there have been reports of discovering ancient samples of 'giant viruses', such as *Mollivirus sibericum* and *Pithovirus sibericum*. Some of these virus particles have retained their infectivity after being thawed out of permafrost after supposedly 30,000 years. The Pithovirus virions resemble those of Pandoravirus, and 16% of the Mollivirus genes have homologs in Pandoravirus.[36] As another example, they found traces of RNA of a coat protein ORF from tomato mosaic tobacovirus from a supposedly 140,000-year-old drill site in Greenland, which differed only by a few percent from extant strains.[37] This raises the obvious question as to how the RNA from this species could remain intact for so long, and it also fails to show any evidence of evolution of this organism over such a supposedly long timespan. Reminiscent of red blood cells isolated from dinosaur bones, active NCLDVs from the permafrost question the long-ages paradigm: can these 'virus' particles retain their infectivity for so long, especially ones the size of NCLDVs with their large genomes intact? Maybe they are not as ancient as evolutionary theory would have it.

There is evidence that the genomes of NCLDVs are also undergoing genome decay, which is the opposite process of gradual evolutionary buildup of genetic information. These genome reduction processes have been observed under laboratory conditions as opposed to unobserved evolutionary speculations. All of these interesting considerations regarding the genomic characteristics of NCLDVs greatly support the creation model and imply that the evolutionary model for the evolution of these species is highly questionable.

The speculation that the hypothetical last common ancestors of all NCLDVs contain 50 common genes has

been refuted, disproving that NCLDVs originate from a single ancestor. It is more likely that NCLDVs have no more than 9 genes in common, and also have independent origins. Furthermore, it would be interesting to analyze any genomic data from NCLDVs to see what kind of groups they cluster to, which could well be the object of future baraminology studies.

Materials and methods

The data for figure 1 came from Boyer *et al.*[17]. Protein sequences were downloaded from the Uniprot FTP website at: ftp.uniprot.org/pub/databases/uniprot/current_release/knowledgebase/taxonomic_divisions/ for 10 major taxonomic categories: Archaea, Bacteria, Fungi, Human, Invertebrates, Mammals, Plants, Rodents, Vertebrates and Viruses. These sequences represented protein sequences from all other domains of life, 549,215 in total. All 20,086 NCLDV proteins were blasted (blastp) against these protein sequences to see if any of them gave a hit with any other species. A maximum e-score cutoff of 1e-4 was applied.

Acknowledgments

I would like to hereby thank Peer Terborg for critical review of the manuscript as well as introducing the idea that NCLDVs are in reality degenerate bacteria. I would also like to thank Robin M. Wyle for critically reading the manuscript.

References

1. Boyer, M., Madoui, M.A., Gimenez, G., La Scola, B. and Raoult, D., Phylogenetic and phyletic studies of informational genes in genomes highlight existence of a 4th domain of life including giant viruses, *PLoS One* **5**(12): e15530, 2010.
2. Colson, P., de Lamballerie, X., Fournous, G. and Raoult, D., Reclassification of giant viruses composing a fourth domain of life in the new order Megavirales, *Intervirology* **55**(5):321–332, 2012.
3. Yutin, N., Wolf, Y.I. and Koonin, E.V., Origin of giant viruses from smaller DNA viruses not from a fourth domain of cellular life, *Virology* **466–467**: 38–52, 2014.
4. Philippe, N., Legendre, M., Doutre, G. *et al*., Pandoraviruses: amoeba viruses with genomes up to 2.5 Mb reaching that of parasitic eukaryotes, *Science* **341**(6143):281–286, 2013.
5. La Scola, B., Desnues, C., Pagnier, I. *et al.*, The virophage as a unique parasite of the giant mimivirus, *Nature* **455**(7209):100–104, 2008.
6. Yutin, N. and Koonin, E.V., Pandoraviruses are highly derived phycodnaviruses, *Biol. Direct.* **8**:25, 2013.
7. Durzyńska, J. and Goździcka-Józefiak, A., Viruses and cells intertwined since the dawn of evolution, *Virol. J.* **12**:16, 2015.
8. Filée, J., Pouget, N. and Chandler, M., Phylogenetic evidence for extensive lateral acquisition of cellular genes by Nucleocytoplasmic large DNA viruses, *BMC Evol. Biol.* **8**:320, 2008.
9. Raoult, D., Audic, S., Robert, C. *et al.*, The 1.2-megabase genome sequence of Mimivirus, *Science* **306**(5700):1344–1350, 2004.
10. Yoosuf, N., Pagnier, I., Fournous, G. *et al.*, Draft genome sequences of Terra1 and Terra2 viruses, new members of the family Mimiviridae isolated from soil, *Virology* **452–453**:125–132, 2014.
11. Fischer, M.G., Allen, M.J., Wilson, W.H. and Suttle, C.A., Giant virus with a remarkable complement of genes infects marine zooplankton, *Proc. Natl. Acad. Sci. USA* **107**(45):19508–19513, 2010.
12. Terborg, P., personal communication.
13. Claverie, J.M. and Ogata, H., Ten good reasons not to exclude giruses from the evolutionary picture, *Nat. Rev. Microbiol.* **7**(8):615, 2009.
14. Ludmir, E.B. and Enquist, L.W., Viral genomes are part of the phylogenetic tree of life, *Nat. Rev. Microbiol.* **7**(8):615, 2009.
15. Saini, H.K. and Fischer, D., Structural and functional insights into Mimivirus ORFans, *BMC Genomics* **8**:115, 2007.
16. Filée, J., Bapteste, E. Susko, E. and Krisch, H.M., A selective barrier to horizontal gene transfer in the T4-type bacteriophages that has preserved a core genome with the viral replication and structural genes, *Mol. Biol. Evol.* **23**(9): 1688–1696, 2006.
17. Boyer, M., Gimenez, G., Suzan-Monti, M. and Raoult, D., Classification and determination of possible origins of ORFans through analysis of nucleocytoplasmic large DNA viruses, *Intervirology* **53**(5):310–320, 2010.
18. Angly, F.E., Felts, B., Breitbart, M. *et al.*, The marine viromes of four oceanic regions, *PLoS Biol.* **4**(11):e368, 2006.
19. Aherfi, S., La Scola, B., Pagnier, I., Raoult, D. and Colson, P., The expanding family Marseilleviridae, *Virology* **466–467**:27–37, 2014.
20. Colson, P. and Raoult, D., Gene repertoire of amoeba-associated giant viruses, *Intervirology* **53**(5):330–343, 2010.
21. Yin, Y. and Fischer, D., On the origin of microbial ORFans: quantifying the strength of the evidence for viral lateral transfer, *BMC Evol. Biol.* **6**:63, 2006.
22. Yutin, N. and Koonin, E.V., Hidden evolutionary complexity of Nucleo-Cytoplasmic Large DNA viruses of eukaryotes, *Virol. J.* **9**:161, 2012.
23. Yutin, N., Wolf, Y.I. and Koonin, E.V., Origin of giant viruses from smaller DNA viruses not from a fourth domain of cellular life, *Virology* **466–467**: 38–52, 2014.
24. Iyer, L.M., Balaji, S., Koonin, E.V. and Aravind, L., Evolutionary genomics of nucleo-cytoplasmic large DNA viruses, *Virus Res.* **117**(1):156–184, 2006.
25. Allen, M.J., Schroeder, D.C., Holden, M.T.G. and Wilson, W.H., Evolutionary history of the occolithoviridae, *Mol. Biol. Evol.* **23**:86–92, 2006.
26. Van Hulten, M.C., Witteveldt, J., Peters, S. *et al.*, The white spot syndrome virus DNA genome sequence, *Virology* **286**(1):7–22, 2001.
27. Van Etten, J.L., Lane L.C. and Dunigan, D.D., DNA viruses: the really big ones (giruses), *Annu. Rev. Microbiol.* **64**:83–99, 2010.
28. Forterre, P., Giant viruses: conflicts in revisiting the virus concept, *Intervirology* **53**(5):362–378, 2010.
29. O'Micks, J., Bacterial genome decay from a baraminological viewpoint, *J. Creation* **29**(3):110–119, 2015.
30. Boyer, M., Azza, S., Barrassi, L. *et al*., Mimivirus shows dramatic genome reduction after intraamoebal culture, *Proc. Natl. Acad. Sci. USA* **108**(25): 10296–10301, 2011.
31. Terborg, P., Evidence for the design of life: part 2—Baranomes, *J. Creation* **22**(3): 68–76, 2008.
32. Ivey, R.G., Henry, E.C., Lee, A.M. *et al*., A *Feldmannia* algal virus has two genome size-classes, *Virology* **220**(2):267–273, 1996.
33. O'Micks, J., Small genome size of *Utricularia gibba* problematic for evolution but not creation, *J. Creation* **27**(3):4–6, 2013.
34. Lefkowitz, E.J., Wang, C. and Upton, C., Poxviruses: past, present and future, *Virus Res.* **117**(1):105–118, 2006.
35. Filée, J., Siguier, P. and Chandler, M., I am what I eat and I eat what I am: acquisition of bacterial genes by giant viruses, *Trends Genet.* **23**(1):10–15, 2007.
36. Legendre, M., Lartigue, A., Bertaux, L. *et al*., In-depth study of *Mollivirus sibericum*, a new 30,000-y-old giant virus infecting *Acanthamoeba*, *Proc. Natl. Acad. Sci. U.S.A* **112**(38):E5327–5335, 2015.
37. Castello, J.D., Rogers, S.O., Starmer, W.T., Catranis, C.M. and Ma, L., Detection of tomato mosaic tobamovirus RNA in ancient glacial ice, *Polar Biol.* **22**(3):207–212, 1999.

Jean O'Micks *has a Ph.D. in biology. He has been an active creationist for 11 years and takes a great interest in molecular biology. He has published a number of articles in* Journal of Creation.