3CAC: improving the classification of phages and plasmids in metagenomic assemblies using assembly graphs

Lianrong Pu and Ron Shamir

Blavatnik School of Computer Science, Tel Aviv University
lianrongpu@mail.tau.ac.il, rshamir@tau.ac.il

Abstract. Bacteriophages and plasmids usually coexist with their host bacteria in microbial communities and play important roles in microbial evolution. Accurately identifying sequence contigs as phages, plasmids, and bacterial chromosomes in mixed metagenomic assemblies is critical for further unravelling their functions. Many classification tools have been developed for identifying either phages or plasmids in metagenomic assemblies. However, only two classifiers, PPR-Meta and viralVerify, were proposed to simultaneously identify phages and plasmids in mixed metagenomic assemblies. Due to the very high fraction of chromosome contigs in the assemblies, both tools achieve high precision in the classification of chromosomes but perform poorly in classifying phages and plasmids. Short contigs in these assemblies are often wrongly classified or classified as uncertain.

Here we present 3CAC, a new three-class classifier that improves the precision of phage and plasmid classification. 3CAC starts with an initial three-class classification generated by existing classifiers and improves the classification of short contigs and contigs with low confidence classification by using proximity in the assembly graph. Evaluation on simulated metagenomes and on real human gut microbiome samples showed that 3CAC outperformed PPR-Meta and viralVerify in both precision and recall, and increased F1-score by 10-60 percentage points.

The 3CAC software is available on https://github.com/Shamir-Lab/3CAC.

Keywords: Metagenome · Three-class Classification · Assembly Graph · Phages · Plasmids.

1 Introduction

The metagenomes of microbial communities are mainly composed of bacterial chromosomes and the associated extrachromosomal mobile genetic elements (eMGEs), such as plasmids and bacteriophages (phages). These eMGEs carry genes related to antibiotic resistance [6, 36, 19], virulence factors [14, 30] and auxiliary metabolic pathways [11, 28, 12]. They can frequently move between species in the microbial community [32, 8] and enable their hosts to rapidly adapt to environmental changes [35, 33]. Despite their important roles in horizontal gene transfer events and in antibiotic resistance, our understanding of these eMGEs is still limited. Part of the difficulty is the challenge of identifying such elements efficiently from mixed metagenomic assemblies [3, 16, 1, 2, 24, 34, 38].

Multiple algorithms have been developed for identifying either phages or plasmids from metagenomic assemblies in recent years. VirSorter and VirSorter2 identify viral metagenomic fragments by searching for reference homologs and testing enrichment of virus-like proteins [29, 10]. These knowledge-based tools have high precision in virus classification but poor ability to identify novel viruses, due to reference database-associated bias. Other tools, such as DeepVirFinder [27], Seeker [4], and VIBRANT [12], use machine learning to learn k-mer signatures of viral sequences and perform better on novel virus classification, since they are more loosely linked to annotation databases. cBar is the first tool designed primarily for plasmid identification in metagenomes [40]. More recently, two supervised-learning approaches, PlasFlow [15] and PlasClass [23], were shown to classify plasmid fragments better from metagenomic assemblies. Although both phages and plasmids are commonly found in the metagenomes of microbial communities, all of these tools identify either only phages or only plasmids from metagenomic assemblies.

Currently, only two published tools, PPR-Meta [7] and viralVerify [2], can identify phages and plasmids simultaneously from metagenomic assemblies. However, due to the overwhelming abundance of chromosome fragments in the assemblies (usually $\geq 70\%$), both tools achieve high precision in chromosome classification but very low precision in classification of phages and plasmids [7, 2]. Moreover, classification of short contigs is challenging for all the existing classifiers, as they analyze each contig independently [2, 7, 15, 29, 26]. Here we present 3CAC (3-Class Adjacency based Classifier), an algorithm that employs existing two-class and three-class classifiers to generate an initial three-class classification with high precision, and then improves the classification of short contigs and of contigs classified with lower confidence by taking advantage of classification of their neighbors in the assembly graph. Evaluation on simulated and real metagenome datasets with short and long reads showed that 3CAC improved both precision and recall, and increased F1-score by at least 10 percentage points.

2 Methods

3CAC accepts as input a set of contigs and its associated assembly graph, uses the classification result of existing tools as a starting point, and repeatedly improves

the classification using the assembly graph. Its output is a classification of each contig in the input as phage, plasmid, chromosome, or uncertain. The details of the algorithm are described below.

2.1 Generating the initial classification

3CAC exploits existing two-class and three-class classifiers to generate an initial three-class classification as follows.

(1) Generating a three-class classification. The algorithm runs either viralVerify or PPR-Meta on the set of the input contigs and classifies each contigs as phage, plasmid, chromosome, or uncertain. viralVerify was designed to classify contigs as viral, non-viral or uncertain. Moreover, for non-viral contigs, viralVerify can further classify them as plasmid or non-plasmid using *-p* option. Here, we used *-p* option of viralVerify to classify each of the input contigs as viral, plasmid, chromosome, or uncertain. PPR-Meta calculates three scores representing the probabilities of a contig to be classified as a phage, plasmid, or chromosome. By default, PPR-Meta classifies a contig into the class with the highest score. If a specified score threshold is provided and no score passes the threshold, the sequence will be classified as uncertain. Here, we ran PPR-Meta with a score threshold of 0.7.

(2) Improving plasmid classification. To improve the precision of plasmid classification, PlasClass is run on contigs classified as plasmids in step (1). PlasClass outputs for each contig the probability that it originated from a plasmid. By default, PlasClass classifies a contig as plasmid if it has a probability > 0.5 and as chromosome otherwise. To assure high precision, here we identify contigs with probability ≥ 0.7 as plasmids. Contigs with probability ≤ 0.3 are moved to the chromosome class. The remaining contigs are reclassified as uncertain.

(3) Improving phage classification. Similarly, in order to improve the precision of phage classification, we run DeepVirFinder on all contigs classified as phages in step (1). DeepVirFinder generates a score and a p-value for each input contig. Contigs with higher scores or lower p-values are more likely to be viral sequences. Here, a contig is kept in the phage class if its p-value ≤ 0.03 and moved to the chromosome class if its p-value > 0.03 and its score ≤ 0.5. The remaining contigs are reclassified as uncertain.

We will denote the algorithm up to this step **Initial(vV)** and **Initial(PM)** if viralVerify or PPR-Meta were used in step (1), respectively.

2.2 Refining the classification using the assembly graph

In genomics and metagenomics, assembly graphs, such as de Bruijn graphs [18, 25] and string graphs [21, 31], are used as the core data structure to combine overlapped reads (or k-mers) into contigs. Initially, nodes are k-mers and edges represent $(k+1)$-long overlaps between them. Each longest path of nodes with indegree and outdegree=1 is then collapsed into a single node representing the corresponding sequence contig. In our description below, nodes in the graph are

the contigs, and the neighbors of a contig are its adjacent nodes. Existing classifiers take contigs as input and classify each of them independently based on its sequence. The overlap information between neighboring contigs in the assembly graphs was ignored by all the existing classifiers. However, recent studies showed that neighboring contigs in an assembly graph are more likely to come from the same taxonomic group [5, 20]. Based on this insight, here we exploit the assembly graph to improve the classification by the following two steps.

(1) Correction of classified contigs. A classified contig is called *incongruous* if it has ≥ 2 classified neighbors and all of them belong to same class, while this contig belongs to a different class. We reason that an incongruous contig was wrongly classified and its classification needs to be consistent with its classified neighbors. Therefore, 3CAC scans all the incongruous contigs in decreasing order of the number of their classified neighbors and corrects the classification of each incongruous contig to match its classified neighbors. Note that once an incongruous contig is corrected, this contig and all its neighbors will not be corrected anymore.

(2) Propagation from classified contigs to unclassified contigs. An unclassified contig is called *implied* if it has one or more classified neighbors and all of them belong to same class. 3CAC dynamically maintains a sorted list of implied contigs in decreasing order of the number of their classified neighbors. At each iteration, 3CAC classifies the first implied contig on the list according to its classified neighbors and then updates the sorted list. Note that only the unclassified neighbors of the first implied contig need to be updated at each iteration. We repeat this step until the list is empty.

Figure 1 shows the result of applying steps (1) and (2) in a small assembly graph, which is part of the graph generated by assembling simulated long reads (Sim4; see details in the Results section).

We will use the names **3CAC(vV)** and **3CAC(PM)** for the full 3CAC algorithms initialized with viralVerify and PPR-Meta solutions, respectively.

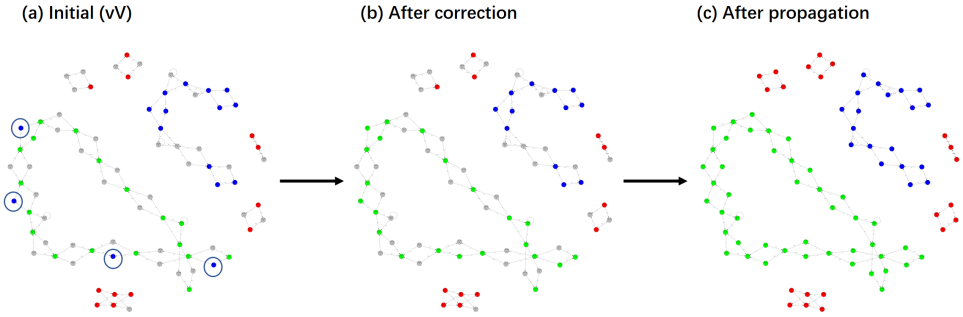

Fig. 1. An example of improving the classification using the assembly graph. Vertices with color red, blue, green, and grey represent contigs classified as phages, plasmids, chromosomes, and uncertain, respectively. (a) The result of Initial(vV). (b) After the correction step. The four contigs encircled in (a) were corrected. (c) After the propagation step.

3 Results

We tested 3CAC on both simulated and real metagenome assemblies and compared it to PPR-Meta and viralVerify.

3.1 Evaluation criteria

3CAC, viralVerify and PPR-Meta were evaluated based on precision, recall, and F1 score, calculated as follows.

- **Precision:** the fraction of correctly classified contigs among all classified contigs. Note that uncertain contigs were not included in the calculation.
- **Recall:** the fraction of correctly classified contigs among all contigs.
- **F1 score:** the harmonic mean of the precision and recall, which can be calculated as: $F1\ score = (2 * precision * recall)/(precision + recall)$.

Following [23, 7], the precision, recall, and F1 score here were calculated by counting the number of contigs and did not take into account their length. The precision and recall were also calculated separately for phage, plasmid and chromosome classification. For example, the precision of phage classification was calculated as the fraction of correctly classified phage contigs among all contigs classified as phages, and the recall of phage classification was calculated as the fraction of correctly classified phage contigs among all phage contigs.

3.2 Performance on simulated metagenome assemblies

We generated two short-read and two long-read metagenome assemblies as follows. Sequences of complete bacterial genomes were randomly selected from the NCBI database along with their associated plasmids. The abundance of bacterial genomes was modeled by the log-normal distribution and the copy numbers of plasmids were simulated by the geometric distribution as in [23]. The phage genomes and their abundance profiles were sampled from [26]. Two metagenomic datasets of different complexities were designed. For each of the datasets, 150bp-long short reads were simulated from the genome sequences using InSilicoSeq [9] and assembled by metaSPAdes [22]. Long reads were simulated from the genome sequences using NanoSim [39] and assembled by metaFlye [13]. The error rate of long reads was 9.8% and their average length was 14.9kb. For each assembly, contigs were matched to the reference genomes used in the simulation by minimap2 [17]. Contigs having matches to a reference genome with $\geq 90\%$ mapping identity along $\geq 80\%$ of the contig length were assigned to the class of that reference, and these assignments were used as the gold standard to test the classifiers. Table 1 presents a summary of the simulated metagenome assemblies.

Figure 2 shows the performance of PPR-Meta, viralVerify and the first phase of 3CAC on these simulated metagenome assemblies. Both PPR-Meta and viralVerify had high precision in chromosome classification, but their precision in phage and plasmid classification was usually low. Further analysis revealed that

Table 1. Properties of the simulated and the real metagenome datasets and of their assemblies. The number of genome references for the real human gut metagenomes is the number of all complete chromosome, plasmid and phage genomes in NCBI database.

	Read type	Number of reads	# of genome references			# of assembled contigs			Short contigs (< 1kb)
			chromosome	plasmid	phage	chromosome	plasmid	phage	
Sim1	MiSeq	61M	50	193	200	12,494	1,699	696	8,991
Sim2	MiSeq	100M	100	410	500	40,412	5,350	2,926	33,640
Sim3	Nanopore	0.5M	50	193	200	890	166	175	45
Sim4	Nanopore	1M	100	410	500	2,491	395	413	152
Sim2021-Miseq	MiSeq	100M	100	433	500	34,211	3,512	2,014	25,861
Sim2021-Nano	Nanopore	1M	100	433	500	2,413	345	342	143
Gut-HiSeq	HiSeq	53.8M	19,053	20,838	13,903	130,252	943	383	110,128
Gut-Pacbio	Pacbio	14.7M	19,053	20,838	13,903	4,671	64	8	723

both of the algorithms distinguished well between phages and plasmids. Their low precision in phage and plasmid classification was due to contamination from chromosome contigs (Supplementary Table A.1). Utilizing two-class classifiers, PlasClass and DeepVirFinder, the first phase of 3CAC improved markedly the precision in phage and plasmid classification, while it decreased a little bit the precision in chromosome classification (Figure 2, Supplementary Table A.2). In contrast, recall decreased in phage and plasmid classification, but increased in chromosome classification (Supplementary Figure B.1).

Figure 3 shows the results of initial phase of 3CAC on the short-read simulated metagenome assemblies for different contig lengths. Short contigs tended to have lower recall in the initial classification of 3CAC, while precision was not sensitive to the contig length. When the initial classification of 3CAC was generated based on PPR-Meta solution, recall decreased sharply for contigs with length $<$ 1kb. When viralVerify solution was used, recall was even lower for contig shorter than 1kb and improvement with size was roughly linear. We reasoned that these classifiers classified each of the input contigs independently, and so short contigs could not be classified reliably. However, Table 1 shows that more than half of the contigs assembled from short reads are shorter than 1kb. To assist in the classification of these short contigs, 3CAC was designed to take advantage of the longer contigs with confident classification and that are neighbors of these short contigs in the assembly graph. Figure 3 shows that 3CAC significantly increased recall for all contigs with almost no loss of precision. Remarkably, the recall for contigs shorter than 1kb increased from < 0.2 to ≥ 0.8. For contigs assembled from long reads, 3CAC not only improved the recall substantially but also slightly improved the precision (Figure 4).

The analysis above shows that the two phases of 3CAC algorithm improved the precision and recall for the three-class classification. Evaluation of PPR-Meta, viralVerify and 3CAC on these simulated metagenome assemblies showed that 3CAC performed the best in both precision and recall (Figure 5, Supplementary Table A.1, A.3). We also calculated the precision, recall and F1 scores for phage, plasmid, and chromosome classification separately (Supplementary Table A.4). 3CAC had the best F1 scores on all the datasets. Note that PPR-Meta here was run with default setting. Running PPR-Meta with 0.7 score thresh-

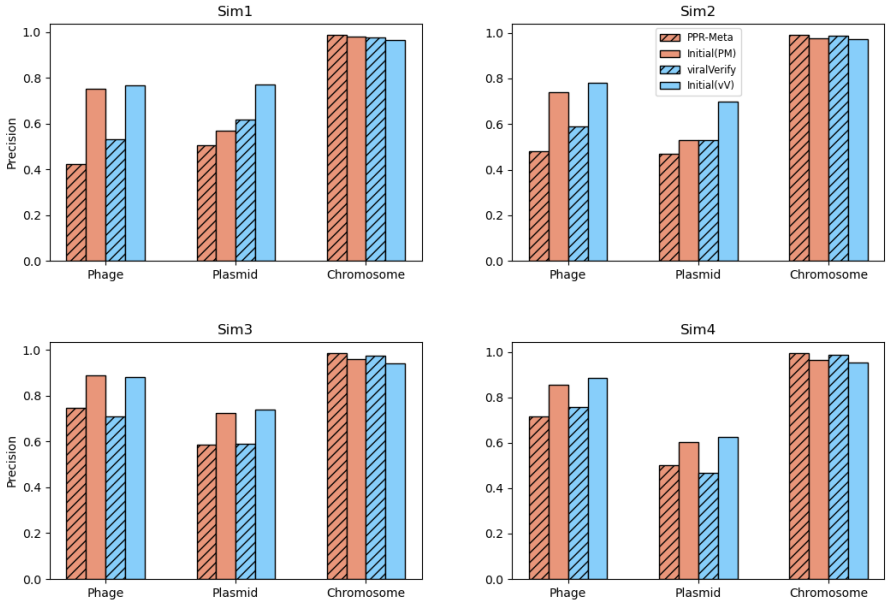

Fig. 2. Precision of the initial classification of 3CAC compared to PPR-Meta and viralVerify. Sim1 and Sim2 are assembled from short reads. Sim3 and Sim4 are assembled from long reads. See supplementary Figure B.1 for recall.

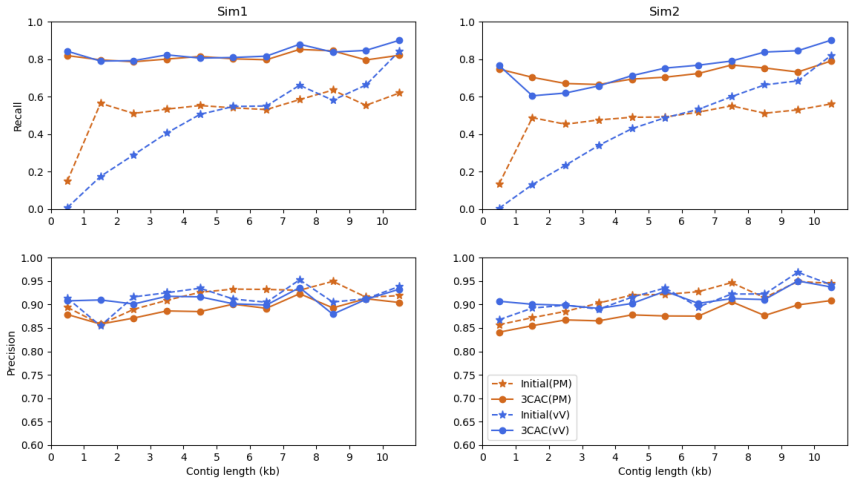

Fig. 3. Performance on contigs assembled from simulated short reads. Results are shown for contigs of lengths < 1 kb, 1-2 kb, ...,9-10 kb, $\geq$ 10kb .

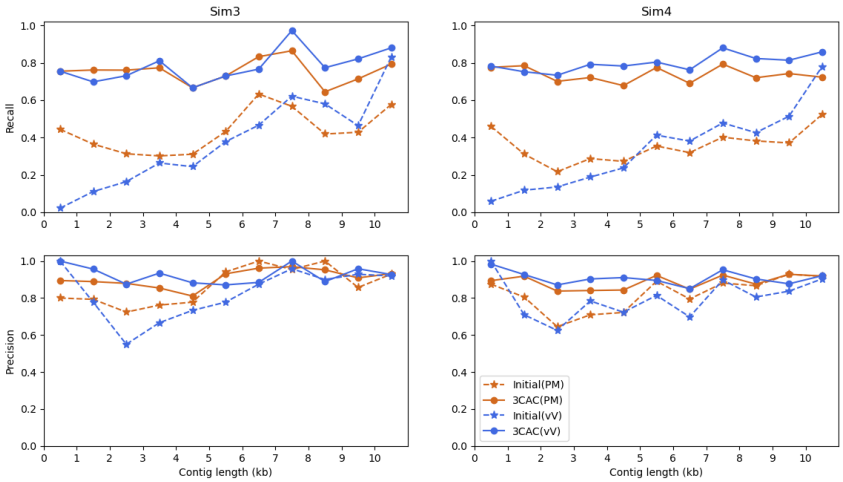

Fig. 4. Performance on contigs assembled from simulated long reads. Results are shown for contigs of lengths < 1 kb, 1-2 kb, ...,9-10 kb, ≥ 10kb .

old (as done in Initial(PM)) resulted in higher precision but lower recall and lower F1 score. Supplementary Table A.5 shows that 3CAC also outperformed PPR-Meta with 0.7 score threshold.

A potential bias in evaluating 3CAC may be created by the fact that the simulated datasets contain sequences that may have been used for training the various classifiers. To evaluate the ability of 3CAC to classify novel species, additional short-read and long-read metagenome simulations were designed by selecting genomes released on NCBI after January 2021 (Sim2021-Miseq and Sim2021-Nano in Table 1). Since all the classifiers that we used were developed prior to that date, all the tested genomes in these metagenome assemblies were not included in the training of the classifiers used by 3CAC. Evaluation on Sim2021-Miseq and Sim2021-Nano showed that 3CAC performed equally well in classification of novel species (Supplementary Figure B.2 and Table A.5).

Overall 3CAC(vV) achieved best precision, recall, and F1 score in all six simulations, and 3CAC(PM) was a close second (Supplementary Table A.5). PPR-Meta and viralVerify had 25-60 lower percentage points in F1 score on the short-read simulations, and 14-30 lower percentage points on the long-read simulations.

3.3 Performance on human gut microbiome samples

Five publicly available human gut microbiome samples with short-read sequencing datasets (NCBI accession numbers: ERR12976697, ERR1297651, ERR1297751, ERR1297845, ERR1297770) were selected and assembled together using metaSPAdes [22]. Another set of five human gut microbiome samples with long-read sequencing datasets (NCBI accessions: SRX2529348, SRX2529347, SRX2529346, SRX2529341, SRX2529340) were selected from [34] and assembled together us-

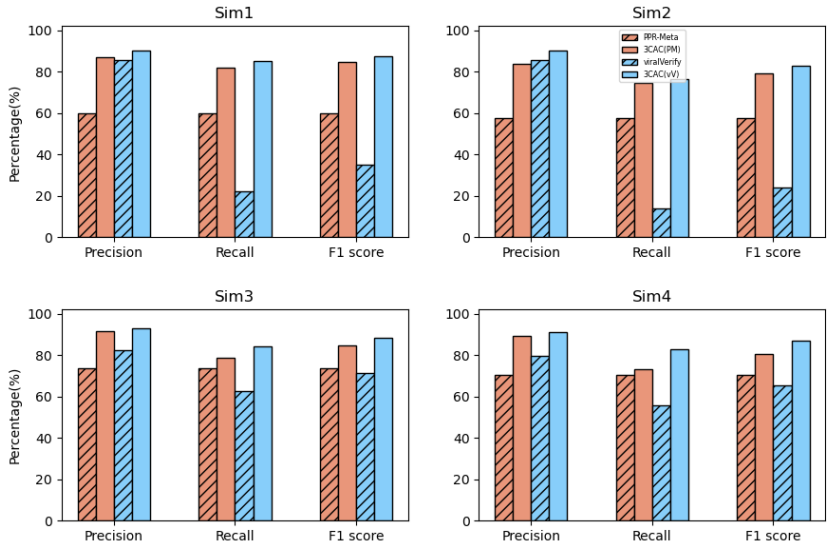

Fig. 5. Performance of three-class classifiers on the simulated metagenome assemblies. Sim1 and Sim2 are assembled from short reads. Sim3 and Sim4 are assembled from long reads.

ing metaFlye [13]. To identify the class of contigs in the real metagenome assemblies, we downloaded all complete phage, plasmid and chromosome genomes from NCBI database and mapped contigs to all the reference genomes using minimap2 [17]. A contig was considered matched to a reference sequence if it had $\geq$ 80% mapping identity along $\geq$ 80% of the contig length. Contigs that matched to reference genomes of two or more classes were excluded to avoid ambiguity. Overall, 131,578 out of 469,022 contigs in the short-read assembly and 4,743 out of 12,541 contigs in the long-read assembly had matches to a single class and were used as the gold standard to test the classifiers. Table 1 summarizes the properties of the datasets and the assemblies.

Figures 6(a) and 7(a) show the results of PPR-Meta, viralVerify and 3CAC on the short-read and long-read assemblies, respectively. On the long-read assembly, 3CAC(vV) and 3CAC(PM) had comparable performance. 3CAC was best in precision, recall and F1 score (Figure 7).

Interestingly, on the short-read assembly, 3CAC(PM) and PPR-Meta had higher F1 score than 3CAC(vV) (Figure 6 (a)). Further analysis revealed that this was due to a large number of isolated contigs in the short-read assembly graph. The second phase of 3CAC was only performed on contigs that have neighbours in the assembly graph. However, 59% of the contigs assembled from short reads were isolated and had no neighbors in the assembly graph, while the fraction on the long-read assembly was only 21%. Figures 6 (b) and 7 (b) show the results on the non-isolated contigs in the assembly graph. For both long-

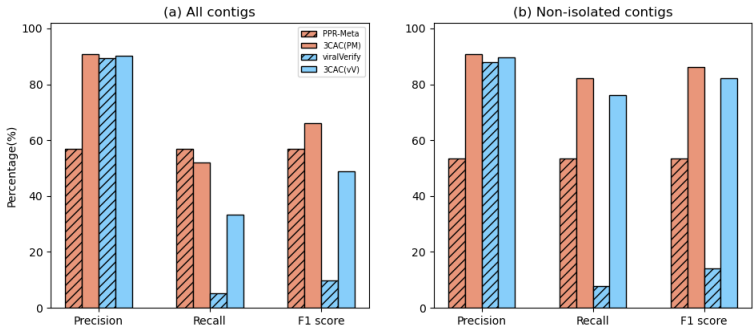

Fig. 6. Performance of three-class classifiers on contigs assembled from short-read sequencing of human gut microbiome samples. (a) performance on all contigs; (b) performance on non-isolated contigs in the assembly graph.

read and short-read assemblies, 3CAC(PM) and 3CAC(vV) had comparable performance and outperformed PPR-meta and viralVerify in precision, recall, and F1 score.

Supplementary Table A.4 shows the precision, recall and F1 score separately for phage, plasmid and chromosome classification in both short-read and long-read assemblies. Compared to PPR-Meta, 3CAC had higher precision in classification of phages and plasmids, at the cost of lower recall and tended to have better F1 scores. In chromosome classification, 3CAC performed the best in the long-read assembly while PPR-Meta performed slightly better in the short-read assembly.

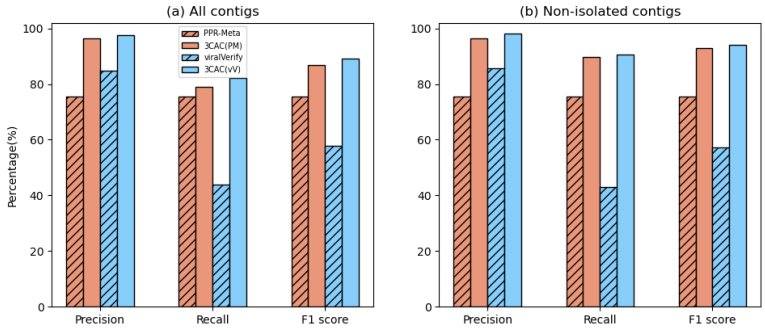

Fig. 7. Performance of three-class classifiers on contigs assembled from long-read sequencing of human gut microbiome samples. (a) performance on all contigs; (b) performance on non-isolated contigs in the assembly graph.

3.4 Software and Resource usage

3CAC uses classification results generated by existing classifiers as a starting point, and so the running time of its first phase depends on the classifiers used.

Supplementary Table A.6 shows that viralVerify was most time consuming, followed by PPR-Meta. PlasClass and DeepVirFinder were fast as they were only run on a subset of the input contigs. Time consumed by the second phase of 3CAC was tiny compared to PPR-Meta and viralVerify. Performance was measured on a 44-core, 2.2 GHz server with 792 GB of RAM. 3CAC is freely available on Github via https://github.com/Shamir-Lab/3CAC.

4 Discussion and Conclusion

In this study we introduced 3CAC, a new three-class classifier to classify contigs in assembly graphs into bacterial, plasmid, viral, and uncertain. 3CAC builds on initial three-class classifications generated by viralVerify or PPR-Meta and improves the classification of short and uncertain contigs by exploiting the structure of the assembly graph. Evaluation on real and simulated metagenomes assembled from both short and long reads showed that 3CAC significantly improved the recall with almost no loss of the precision.

In the correction step of 3CAC, the order of treating incongruous contigs with the same number of classified neighbors may affect the results. A similar situation occurs in the propagation step. We tried several random orders and the results were very stable ($< 1\%$ difference, Supplementary Table A.7).

Should a user interested specifically in a two-way classification prefer a dedicated two-way classifier over 3CAC? We compared PlasClass to 3CAC in plasmid classification, and DeepVirFinder to 3CAC in phage classification. The results on the six simulated datasets show that 3CAC had superior precision and F1 score in all cases (Supplementary Table A.8 and Table A.9).

Evaluation the performance of classifiers on real metagenome assemblies remains challenging due to the lack of gold standard. By mapping contigs to all the available reference genomes, we are able to identify the class of a fraction of the contigs. However, as shown in previous studies [7], some plasmid genomes are quite similar to their host bacterial chromosomes. Thus, many contigs from metagenome assemblies have matches to both plasmid and chromosome reference genomes, and it is hard to identify their classes. Additionally, many contigs with no matches to the reference database may represent novel species, but they were excluded from our evaluation. Keeping in mind these shortcomings of the gold standard for real metagenome assemblies, 3CAC outperformed existing three-class classifiers substantially.

3CAC has some limitations. The propagation step of 3CAC can greatly improve the recall, but it can only be performed on non-isolated contigs in the assembly graph. The recall of isolated contigs is still limited by the performance of existing classifiers. 3CAC relies on existing two-class and three-class classifiers. In the future, we plan to extend 3CAC to a stand-alone classification tool. The detection of prophages and other non-bacterial sequences that integrate into bacterial genome is challenging to all classifiers, including 3CAC. Finally, there is room for extending 3CAC to a four-class algorithm that would be able to classify also eukaryotic contigs in metagenome assemblies [37].

Acknowledgments

This study was supported in part by grant 2016694 from the United State - Israel Binational Science Foundation (BSF), Jerusalem, Israel and the United States National Science Foundation (NSF). L.P. was supported in part by a fellowship from the Edmond J. Safra Center for Bioinformatics at Tel-Aviv University. L.P. was also supported in part by postdoctoral fellowships from the Planning and Budgeting Committee (PBC) of the Council for Higher Education (CHE) in Israel.

References

1. Antipov, D., Raiko, M., Lapidus, A., Pevzner, P.A.: Plasmid detection and assembly in genomic and metagenomic data sets. Genome Research **29**(6), 961–968 (2019)
2. Antipov, D., Raiko, M., Lapidus, A., Pevzner, P.A.: Metaviral SPAdes: assembly of viruses from metagenomic data. Bioinformatics **36**(14), 4126–4129 (2020)
3. Arredondo-Alonso, S., Willems, R.J., Van Schaik, W., Schürch, A.C.: On the (im) possibility of reconstructing plasmids from whole-genome short-read sequencing data. Microbial Genomics **3**(10) (2017)
4. Auslander, N., Gussow, A.B., Benler, S., Wolf, Y.I., Koonin, E.V.: Seeker: alignment-free identification of bacteriophage genomes by deep learning. Nucleic Acids Research **48**(21), e121–e121 (2020)
5. Barnum, T.P., Figueroa, I.A., Carlström, C.I., Lucas, L.N., Engelbrektson, A.L., Coates, J.D.: Genome-resolved metagenomics identifies genetic mobility, metabolic interactions, and unexpected diversity in perchlorate-reducing communities. The ISME Journal **12**(6), 1568–1581 (2018)
6. Calero-Cáceres, W., Ye, M., Balcázar, J.L.: Bacteriophages as environmental reservoirs of antibiotic resistance. Trends in Microbiology **27**(7), 570–577 (2019)
7. Fang, Z., Tan, J., Wu, S., Li, M., Xu, C., Xie, Z., Zhu, H.: PPR-Meta: a tool for identifying phages and plasmids from metagenomic fragments using deep learning. GigaScience **8**(6), giz066 (2019)
8. Frost, L.S., Leplae, R., Summers, A.O., Toussaint, A.: Mobile genetic elements: the agents of open source evolution. Nature Reviews Microbiology **3**(9), 722–732 (2005)
9. Gourlé, H., Karlsson-Lindsjö, O., Hayer, J., Bongcam-Rudloff, E.: Simulating illumina metagenomic data with insilicoseq. Bioinformatics **35**(3), 521–522 (2019)
10. Guo, J., Bolduc, B., Zayed, A.A., Varsani, A., Dominguez-Huerta, G., Delmont, T.O., Pratama, A.A., Gazitúa, M.C., Vik, D., Sullivan, M.B., et al.: Virsorter2: a multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses. Microbiome **9**(1), 1–13 (2021)
11. Hurwitz, B.L., U'Ren, J.M.: Viral metabolic reprogramming in marine ecosystems. Current Opinion in Microbiology **31**, 161–168 (2016)
12. Kieft, K., Zhou, Z., Anantharaman, K.: Vibrant: automated recovery, annotation and curation of microbial viruses, and evaluation of viral community function from genomic sequences. Microbiome **8**(1), 1–23 (2020)
13. Kolmogorov, M., Bickhart, D.M., Behsaz, B., Gurevich, A., Rayko, M., Shin, S.B., Kuhn, K., Yuan, J., Polevikov, E., Smith, T.P., et al.: metaFlye: scalable long-read metagenome assembly using repeat graphs. Nature Methods **17**(11), 1103–1110 (2020)

14. Kraushaar, B., Hammerl, J., Kienol, M., Heinig, M., Sperling, N., Dinh Thanh, M., Reetz, J., Jackel, C., Fetsch, A., Hertwig, S.: Acquisition of virulence factors in livestock-associated mrsa: lysogenic conversion of cc398 strains by virulence gene-containing phages. Sci Rep 7: 2004 (2017)
15. Krawczyk, P.S., Lipinski, L., Dziembowski, A.: Plasflow: predicting plasmid sequences in metagenomic data using genome signatures. Nucleic Acids Research **46**(6), e35–e35 (2018)
16. Krishnamurthy, S.R., Wang, D.: Origins and challenges of viral dark matter. Virus Research **239**, 136–142 (2017)
17. Li, H.: Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics **34**(18), 3094–3100 (2018)
18. Lin, Y., Yuan, J., Kolmogorov, M., Shen, M.W., Chaisson, M., Pevzner, P.A.: Assembly of long error-prone reads using de Bruijn graphs. Proceedings of the National Academy of Sciences **113**(52), E8396–E8405 (2016)
19. Lopatkin, A., Meredith, H., Srimani, J., Pfeiffer, C., Durrett, R., You, L.: Persistence and reversal of plasmid-mediated antibiotic resistance. Nature Communications 8: 1689 (2017)
20. Mallawaarachchi, V., Wickramarachchi, A., Lin, Y.: Graphbin: refined binning of metagenomic contigs using assembly graphs. Bioinformatics **36**(11), 3307–3313 (2020)
21. Myers, E.W.: The fragment assembly string graph. Bioinformatics **21**(suppl_2), ii79–ii85 (2005)
22. Nurk, S., Meleshko, D., Korobeynikov, A., Pevzner, P.: metaSPAdes: a new versatile de novo metagenomics assembler. arXiv preprint arXiv:1604.03071 (2016)
23. Pellow, D., Mizrahi, I., Shamir, R.: Plasclass improves plasmid sequence classification. PLoS Computational Biology **16**(4), e1007781 (2020)
24. Pellow, D., Zorea, A., Probst, M., Furman, O., Segal, A., Mizrahi, I., Shamir, R.: Scapp: An algorithm for improved plasmid assembly in metagenomes. Microbiome **9**(1), 1–12 (2021)
25. Pevzner, P.A., Tang, H., Waterman, M.S.: An Eulerian path approach to DNA fragment assembly. Proceedings of the National Academy of Sciences **98**(17), 9748–9753 (2001)
26. Ren, J., Ahlgren, N.A., Lu, Y.Y., Fuhrman, J.A., Sun, F.: Virfinder: a novel k-mer based tool for identifying viral sequences from assembled metagenomic data. Microbiome **5**(1), 1–20 (2017)
27. Ren, J., Song, K., Deng, C., Ahlgren, N.A., Fuhrman, J.A., Li, Y., Xie, X., Poplin, R., Sun, F.: Identifying viruses from metagenomic data using deep learning. Quantitative Biology pp. 1–14 (2020)
28. Rosenwasser, S., Ziv, C., Van Creveld, S.G., Vardi, A.: Virocell metabolism: metabolic innovations during host–virus interactions in the ocean. Trends in Microbiology **24**(10), 821–832 (2016)
29. Roux, S., Enault, F., Hurwitz, B.L., Sullivan, M.B.: Virsorter: mining viral signal from microbial genomic data. PeerJ **3**, e985 (2015)
30. Sarowska, J., Futoma-Koloch, B., Jama-Kmiecik, A., Frej-Madrzak, M., Ksiazczyk, M., Bugla-Ploskonska, G., Choroszy-Krol, I.: Virulence factors, prevalence and potential transmission of extraintestinal pathogenic escherichia coli isolated from different sources: recent reports. Gut Pathogens **11**(1), 1–16 (2019)
31. Simpson, J.T., Durbin, R.: Efficient de novo assembly of large genomes using compressed data structures. Genome Research **22**(3), 549–556 (2012)

32. Sitaraman, R.: Prokaryotic horizontal gene transfer within the human holobiont: ecological-evolutionary inferences, implications and possibilities. Microbiome **6**(1), 1–14 (2018)
33. Smalla, K., Jechalke, S., Top, E.M.: Plasmid detection, characterization, and ecology. Microbiology Spectrum **3**(1), 3–1 (2015)
34. Suzuki, Y., Nishijima, S., Furuta, Y., Yoshimura, J., Suda, W., Oshima, K., Hattori, M., Morishita, S.: Long-read metagenomic exploration of extrachromosomal mobile genetic elements in the human gut. Microbiome **7**(1), 1–16 (2019)
35. Thomas, C.M., Nielsen, K.M.: Mechanisms of, and barriers to, horizontal gene transfer between bacteria. Nature Reviews Microbiology **3**(9), 711–721 (2005)
36. Wein, T., Hülter, N., Mizrahi, I., Dagan, T.: Emergence of plasmid stability under non-selective conditions maintains antibiotic resistance. Nature Communications 10: 2595 (2019)
37. West, P.T., Probst, A.J., Grigoriev, I.V., Thomas, B.C., Banfield, J.F.: Genome-reconstruction for eukaryotes from complex natural microbial communities. Genome Research **28**(4), 569–580 (2018)
38. Yahara, K., Suzuki, M., Hirabayashi, A., Suda, W., Hattori, M., Suzuki, Y., Okazaki, Y.: Long-read metagenomics using promethion uncovers oral bacteriophages and their interaction with host bacteria. Nature Communications **12**(1), 1–12 (2021)
39. Yang, C., Chu, J., Warren, R.L., Birol, I.: Nanosim: nanopore sequence read simulator based on statistical characterization. GigaScience **6**(4), gix010 (2017)
40. Zhou, F., Xu, Y.: cBar: a computer program to distinguish plasmid-derived from chromosome-derived sequence fragments in metagenomics data. Bioinformatics **26**(16), 2051–2052 (2010)

Supplementary Material

A Supplementary tables

Table A.1. Performance of viralVerify and PPR-Meta on simulated metagenome assemblies. PPR-Meta was run with 0.7 score threshold to assure high precision.

	True Category	# contigs	viralVerify classification				PPR-Meta classification			
			phage	plasmid	chromosome	uncertain	phage	plasmid	chromosome	uncertain
Sim1	phage	696	279	6	1	410	572	1	1	122
	plasmid	1,699	23	380	58	1,238	80	452	43	1,124
	chromosome	12,494	224	231	2,624	9,415	696	442	3,749	7,607
Sim2	phage	2,926	737	7	3	2,179	1,919	14	2	991
	plasmid	5,350	53	652	60	4,585	158	1,351	71	3,770
	chromosome	40,412	462	572	5,448	33,930	1,926	1,520	9,568	27,398
Sim3	phage	175	163	2	0	10	164	0	0	11
	plasmid	166	9	124	12	21	1	88	5	72
	chromosome	890	58	84	485	263	55	62	366	407
Sim4	phage	413	376	4	0	33	362	1	0	50
	plasmid	395	6	301	16	72	2	197	4	192
	chromosome	2,491	114	337	1,161	879	142	196	889	1,264
Sim2021-Miseq	phage	2,014	591	2	2	1,419	1,455	30	24	505
	plasmid	3,512	33	768	55	2,656	149	988	59	2,316
	chromosome	34,211	502	498	5,718	27,493	2,104	1,041	9,547	21,519
Sim2021-Nano	phage	342	290	3	1	48	306	0	1	35
	plasmid	345	5	274	5	61	4	196	2	143
	chromosome	2,413	153	215	1,158	887	152	119	1,045	1,097

Table A.2. Performance of Initial(vV) and Initial(PM) on simulated metagenome assemblies.

	True Category	# contigs	Initial(vV)				Initial(PM)			
			phage	plasmid	chromosome	uncertain	phage	plasmid	chromosome	uncertain
Sim1	phage	696	249	3	8	436	442	0	8	246
	plasmid	1,699	7	310	88	1,294	10	372	75	1,242
	chromosome	12,494	69	90	2,797	9,538	137	284	3,936	8,137
Sim2	phage	2,926	653	4	35	2,234	1,136	9	45	1,736
	plasmid	5,350	18	499	132	4,701	33	964	213	4,140
	chromosome	40,412	165	213	5,803	34,231	363	842	10,092	29,115
Sim3	phage	175	139	1	11	24	145	0	9	21
	plasmid	166	3	108	22	33	0	82	9	75
	chromosome	890	16	37	540	297	18	31	407	434
Sim4	phage	413	321	2	34	71	318	0	23	87
	plasmid	395	2	262	36	103	1	175	14	213
	chromosome	2,491	38	153	1,280	1012	50	114	970	1,349
Sim2021-Miseq	phage	2,014	503	2	45	1,464	1,068	17	65	864
	plasmid	3,512	17	581	97	2,817	19	689	126	2,678
	chromosome	34,211	172	182	6,092	27,765	272	613	9,983	23,343
Sim2021-Nano	phage	342	261	3	15	63	291	0	5	46
	plasmid	345	3	236	21	85	3	167	17	158
	chromosome	2,413	36	96	1,292	989	38	69	1,129	1,177

Table A.3. Performance of 3CAC on simulated metagenome assemblies.

	True Category	# contigs	3CAC(vV)				3CAC(PM)			
			phage	plasmid	chromosome	uncertain	phage	plasmid	chromosome	uncertain
Sim1	phage	696	642	7	21	26	646	7	17	26
	plasmid	1,699	18	1,135	404	142	32	1,103	411	153
	chromosome	12,494	393	538	10,901	662	434	902	10,481	677
Sim2	phage	2,926	2,728	9	77	112	2,714	26	80	106
	plasmid	5,350	50	2,815	1,440	1,045	98	2,871	1,695	686
	chromosome	40,412	915	1,542	31,712	6,243	1,278	3,760	30,817	4,557
Sim3	phage	175	145	1	11	18	148	1	8	18
	plasmid	166	3	124	25	14	0	110	18	38
	chromosome	890	12	28	769	81	25	36	711	118
Sim4	phage	413	336	2	35	40	328	0	27	58
	plasmid	395	2	306	41	46	4	210	43	138
	chromosome	2,491	35	148	2,087	221	71	143	1,877	400
Sim2021-Miseq	phage	2,014	1,575	3	128	308	1,662	24	153	175
	plasmid	3,512	55	2,596	472	389	66	2,402	730	314
	chromosome	34,211	674	1,260	30,076	2,201	1,008	2,131	29,364	1,708
Sim2021-Nano	phage	342	297	5	15	25	297	0	5	40
	plasmid	345	3	272	21	49	11	218	25	91
	chromosome	2,413	33	94	2,066	220	45	94	1,961	313

Table A.4. Classification of phages, plasmids, and chromosomes on simulated and real metagenome assemblies. PPR-Meta and PPR-Meta(0.7) represent running PPR-Meta on default setting and with a score threshold of 0.7, respectively.

	Tool	Phage			Plasmid			Chromosome		
		precision(%)	recall(%)	F1 score(%)	precision(%)	recall(%)	F1 score(%)	precision(%)	recall(%)	F1 score(%)
Sim1	viralVerify	53.04	40.09	45.66	61.59	22.37	32.82	97.8	21.0	34.58
	3CAC(vV)	60.97	92.24	73.41	67.56	66.80	67.18	96.25	87.25	91.53
	PPR-Meta(0.7)	42.43	82.18	55.97	50.5	26.6	34.85	98.84	30.01	46.04
	3CAC(PM)	58.09	92.82	71.46	54.82	64.92	59.44	96.08	83.89	89.57
	PPR-Meta	19.06	94.54	31.72	25.97	57.74	35.83	95.31	58.43	72.45
Sim2	viralVerify	58.87	25.19	35.28	52.97	12.19	19.81	98.86	13.48	23.73
	3CAC(vV)	73.87	93.23	82.43	64.48	52.62	57.95	95.43	78.47	86.13
	PPR-Meta (0.7)	47.94	65.58	55.39	46.83	25.25	32.81	99.24	23.68	38.23
	3CAC(PM)	66.36	92.75	77.37	43.13	53.66	47.82	94.55	76.26	84.43
	PPR-Meta	21.83	90.46	35.17	24.31	61.7	34.88	95.86	54.53	69.52
Sim3	viralVerify	70.87	93.14	80.49	59.05	74.7	65.96	97.59	54.49	69.94
	3CAC(vV)	90.63	82.86	86.57	81.05	74.70	77.74	95.53	86.40	90.74
	PPR-Meta (0.7)	74.55	93.71	83.04	58.67	53.01	55.70	98.65	41.12	58.05
	3CAC(PM)	85.55	84.57	85.06	74.83	66.27	70.29	96.47	79.89	87.40
	PPR-Meta	60.63	99.43	75.33	42.04	84.34	56.11	97.38	66.85	79.28
Sim4	viralVerify	75.81	91.04	82.73	46.88	76.2	58.05	98.64	46.61	63.30
	3CAC(vV)	90.08	81.36	85.50	67.70	77.47	72.26	96.49	83.90	89.76
	PPR-Meta (0.7)	71.54	87.65	78.78	50.0	49.87	49.94	99.55	35.69	52.54
	3CAC(PM)	81.39	79.42	80.39	59.49	53.16	56.15	96.40	75.35	84.59
	PPR-Meta	60.3	98.55	74.82	33.17	84.05	47.57	97.47	63.51	76.91
Sim2021-Miseq	viralVerify	52.49	29.34	37.64	60.57	21.87	32.13	99.01	16.71	28.60
	3CAC(vV)	68.36	78.20	72.95	67.27	73.92	70.44	98.04	87.91	92.70
	PPR-Meta (0.7)	39.24	72.24	50.86	47.98	28.13	35.47	99.14	27.91	43.55
	3CAC(PM)	60.75	82.52	69.98	52.71	68.39	59.54	97.08	85.83	91.11
	PPR-Meta	17.16	85.55	28.59	23.43	63.30	34.20	96.62	57.08	71.76
Sim2021-Nano	viralVerify	64.73	84.80	73.42	55.69	79.42	65.47	99.48	47.99	64.75
	3CAC(vV)	89.19	86.84	88.00	73.32	78.84	75.98	98.29	85.62	91.52
	PPR-Meta (0.7)	66.23	89.47	76.12	62.22	56.81	59.39	99.71	43.31	60.39
	3CAC(PM)	84.14	86.84	85.47	69.87	63.19	66.36	98.49	81.27	89.06
	PPR-Meta	36.42	98.28	53.15	29.39	87.85	44.04	99.46	72.45	83.84
Gut-Hiseq	viralVerify	8.65	8.88	8.76	7.13	3.77	4.93	99.78	5.15	9.79
	3CAC(vV)	11.97	40.21	18.44	2.66	10.52	4.25	99.49	33.62	50.26
	PPR-Meta (0.7)	2.24	55.09	4.30	3.50	19.38	5.93	99.81	26.78	42.23
	3CAC(PM)	7.02	58.22	12.53	3.75	16.39	6.11	99.56	52.15	68.45
	PPR-Meta	1.13	80.94	2.23	1.51	49.94	2.94	99.62	56.87	72.41
Gut-Pacbio	viralVerify	3.80	37.50	6.90	6.54	32.81	10.91	100	44.02	61.13
	3CAC(vV)	27.59	100	43.24	17.65	23.44	20.13	99.74	82.89	90.54
	PPR-Meta (0.7)	8.25	100	15.24	19.55	40.63	26.40	99.63	39.94	57.02
	3CAC(PM)	10.00	87.50	17.95	24.14	32.81	27.81	99.65	79.49	88.43
	PPR-Meta	2.47	100	4.82	4.95	67.19	9.22	99.55	75.68	85.99

Table A.5. Performance of three-class classifiers on simulated and real metagenome assemblies. PPR-Meta and PPR-Meta(0.7) represent running PPR-Meta on default setting and with a score threshold of 0.7, respectively.

Dataset	Evaluation Criteria	**viralVerify**	**3CAC(vV)**	**PPR-Meta(0.7)**	**3CAC(PM)**	**PPR-Meta**
Sim1	Precision	85.81%	**90.18%**	79.08%	87.15%	60.04%
	Recall	22.05%	**85.15%**	32.06%	82.14%	60.04%
	F1 score	35.08%	**87.59%**	45.62%	84.57%	60.04%
Sim2	Precision	85.53%	**90.23%**	77.67%	83.99%	57.47%
	Recall	14.04%	**76.52%**	26.37%	74.77%	57.47%
	F1 score	24.12%	**82.81%**	39.37%	79.11%	57.47%
Sim3	Precision	82.39%	**92.84%**	83.4%	91.67%	73.84%
	Recall	62.71%	**84.32%**	50.2%	78.72%	73.84%
	F1 score	71.22%	**88.38%**	62.68%	84.70%	73.84%
Sim4	Precision	79.4%	**91.34%**	80.76%	89.35%	70.35%
	Recall	55.71%	**82.81%**	43.89%	73.20%	70.35%
	F1 score	65.48%	**86.87%**	56.87%	80.47%	70.35%
Sim2021-Miseq	Precision	86.63%	**92.96%**	77.87%	89.05%	59.07%
	Recall	17.81%	**86.18%**	30.17%	84.12%	59.07%
	F1 score	29.55%	**89.45%**	43.49%	86.51%	59.07%
Sim2021-Nano	Precision	81.84%	**93.91%**	84.77%	93.22%	74.91%
	Recall	55.55%	**85.00%**	49.90%	79.87%	74.91%
	F1 score	66.18%	**89.23%**	62.82%	86.03%	74.91%
Gut-Hiseq	Precision	89.24%	90.12%	71.41%	**90.64%**	56.90%
	Recall	5.15%	33.48%	26.81%	51.92%	**56.90%**
	F1 score	9.74%	48.83%	38.98%	**66.02%**	56.90%
Gut-Pacbio	Precision	84.70%	**97.47%**	90.35%	96.35%	75.61%
	Recall	43.86%	**82.12%**	40.05%	78.87%	75.61%
	F1 score	57.79%	**89.14%**	55.50%	86.74%	75.61%

Table A.6. Running time in minutes of the classifiers. Row 3CAC shows the running time for the second phase of 3CAC algorithm.

	Sim1	**Sim2**	**Sim3**	**Sim4**	**Sim2021 -Miseq**	**Sim2021 -Nano**	**Gut-HiSeq**	**Gut-Pacbio**
viralVerify	141	201	165	218	314	324	238	292
PPR-Meta	22	32	23	33	69	65	67	50
PlasClass	0.43	0.61	0.72	0.72	0.6	0.68	0.7	0.22
DeepVirFinder	2.5	5.5	2	3.5	6.3	3.4	32	1
3CAC	1.5	3.5	0.5	0.5	2.1	0.5	4	55

Table A.7. Performance of 3CAC with different contig scanning orders. In the correction step, each set of incongruous contigs with the same number of classified neighbors was implemented by a queue data structure. The same was done for implied contigs in the propagation step. These queues determine the contig scanning order. To test the effect of the order, we modified 3CAC to use lists instead of queues, and inserted any new contig to a random location in the corresponding list. The table shows the results for the original 3CAC (O) and for two random orders R1 and R2.

Dataset	Evaluation Criteria	**3CAC(vV)**			**3CAC(PM)**		
		O	**R1**	**R2**	**O**	**R1**	**R2**
Sim1	Precision	90.18%	89.70%	89.49%	87.15%	86.71%	86.65%
	Recall	85.15%	84.96%	84.79%	82.14%	82.05%	82.11%
	F1 score	87.59%	87.27%	87.08%	84.57%	84.31%	84.32%
Sim2	Precision	90.23%	90.19%	89.76%	83.99%	83.72%	83.71%
	Recall	76.52%	76.70%	76.32%	74.77%	74.76%	74.73%
	F1 score	82.81%	82.90%	82.50%	79.11%	78.99%	78.96%
Sim3	Precision	92.84%	92.68%	92.67%	91.67%	90.52%	91.52%
	Recall	84.32%	84.32%	84.24%	78.72%	78.31%	78.88%
	F1 score	88.38%	88.30%	88.26%	84.70%	83.97%	84.73%
Sim4	Precision	91.34%	91.09%	90.98%	89.35%	89.65%	89.32%
	Recall	82.81%	82.72%	82.60%	73.20%	73.81%	73.54%
	F1 score	86.87%	86.70%	86.59%	80.47%	80.96%	80.67%
Sim2021-Miseq	Precision	92.96%	92.92%	92.55%	89.05%	88.94%	88.73%
	Recall	86.18%	86.44%	86.09%	84.12%	84.24%	84.05%
	F1 score	89.45%	89.56%	89.21%	86.51%	86.53%	86.32%
Sim2021-Nano	Precision	93.91%	93.91%	93.67%	93.22%	92.66%	92.43%
	Recall	85.00%	85.13%	84.90%	79.87%	79.45%	79.55%
	F1 score	89.23%	89.31%	89.07%	86.03%	85.55%	85.51%
Gut-Hiseq	Precision	90.12%	89.86%	89.73%	90.64%	90.50%	90.35%
	Recall	33.48%	33.42%	33.37%	51.92%	52.00%	51.93%
	F1 score	48.83%	48.72%	48.65%	66.02%	66.05%	65.95%
Gut-Pacbio	Precision	97.47%	97.53%	97.58%	96.35%	96.10%	96.33%
	Recall	82.12%	82.24%	82.35%	78.87%	78.83%	79.00%
	F1 score	89.14%	89.23%	89.32%	86.74%	86.61%	86.81%

Table A.8. Comparison of PlasClass and 3CAC on plasmid classification in simulated metagenome assemblies. PlasClass(0.5) and PlasClass(0.7) classified a contig as plasmid if its score ≥ 0.7 and > 0.5, respectively. Here, precision was calculated as the fraction of correctly classified plasmid contigs among all contigs classified as plasmids, and the recall was calculated as the fraction of correctly classified plasmid contigs among all plasmid contigs.

Dataset	Evaluation Criteria	**PlasClass(0.5)**	**PlasClass(0.7)**	**3CAC(vV)**	**3CAC(PM)**
Sim1	Precision	20.32%	21.48%	**67.56%**	54.82%
	Recall	**68.16%**	57.27%	66.80%	64.92%
	F1 score	31.30%	31.25%	**67.18%**	59.44%
Sim2	Precision	17.49%	19.04%	**64.48%**	43.13%
	Recall	**68.41%**	57.42%	52.62%	53.66%
	F1 score	27.86%	28.60%	**57.95%**	47.82%
Sim3	Precision	31.20%	42.41%	**81.05%**	74.83%
	Recall	**87.95%**	82.53%	74.70%	66.27%
	F1 score	46.06%	56.03%	**77.74%**	70.29%
Sim4	Precision	27.32%	35.00%	**67.11%**	59.49%
	Recall	**89.37%**	77.97%	77.47%	53.16%
	F1 score	41.85%	48.31%	**71.92%**	56.15%
Sim2021-Miseq	Precision	15.67%	16.56%	**67.27%**	52.71%
	Recall	71.50%	58.66%	**73.92%**	68.39%
	F1 score	25.71%	25.82%	**70.44%**	59.54%
Sim2021-Nano	Precision	26.98%	36.40%	**73.32%**	69.87%
	Recall	**86.96%**	81.45%	78.84%	63.19%
	F1 score	41.18%	50.31%	**75.98%**	66.36%

Table A.9. Comparison of DeepVirFinder and 3CAC on phage classification in simulated metagenome assemblies. DeepVirFinder(0.05) and DeepVirFinder(0.03) classified a contig as phage if its pvalue ≤ 0.05 and ≤ 0.03, respectively. Here, precision was calculated as the fraction of correctly classified phage contigs among all contigs classified as phages, and the recall was calculated as the fraction of correctly classified phage contigs among all phage contigs. 76 contigs longer than 2Mb were excluded from the comparison since DeepVirFinder did not terminate successfully on these long contigs.

Dataset	Evaluation Criteria	**DeepVirFinder(0.05)**	**DeepVirFinder(0.03)**	**3CAC(vV)**	**3CAC(PM)**
Sim1	Precision	29.46%	39.10%	**60.97%**	58.09%
	Recall	83.76%	77.59%	92.24%	**92.82%**
	F1 score	43.59%	52.00%	**73.41%**	71.46%
Sim2	Precision	30.99%	39.52%	**73.87%**	66.36%
	Recall	70.95%	63.81%	**93.23%**	92.75%
	F1 score	43.14%	48.81%	**82.43%**	77.37%
Sim3	Precision	49.09%	60.98%	**90.63%**	85.55%
	Recall	**92.00%**	85.71%	82.86%	84.57%
	F1 score	64.02%	71.26%	**86.57%**	85.06%
Sim4	Precision	46.22%	58.63%	**90.08%**	81.39%
	Recall	**90.31%**	83.05%	81.36%	79.42%
	F1 score	61.15%	68.74%	**85.50%**	80.39%
Sim2021-Miseq	Precision	28.52%	37.34%	**68.36%**	60.75%
	Recall	74.48%	67.03%	78.20%	**82.52%**
	F1 score	47.97%	41.24%	**72.95%**	69.98%
Sim2021-Nano	Precision	37.63%	55.70%	**89.19%**	84.14%
	Recall	**94.74%**	90.06%	86.84%	86.84%
	F1 score	53.87%	68.83%	**88.00%**	85.47%

B Supplementary figures

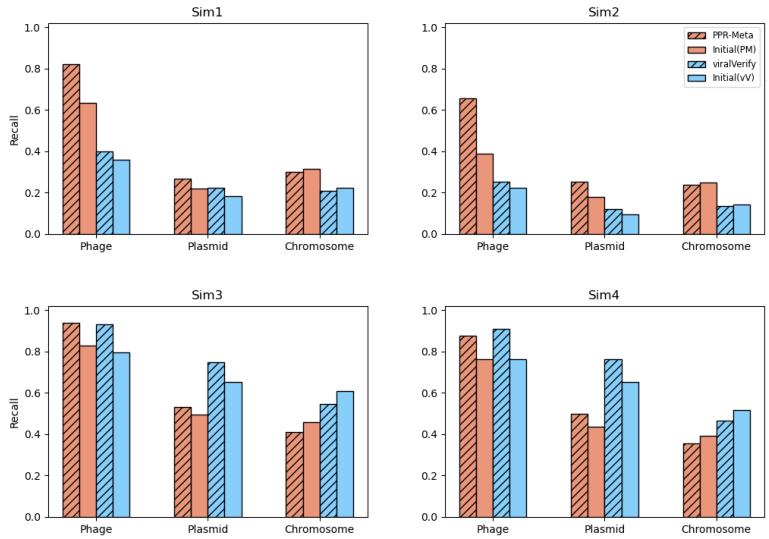

Fig. B.1. Recall of the initial classification of 3CAC compared to PPR-Meta and viralVerify.

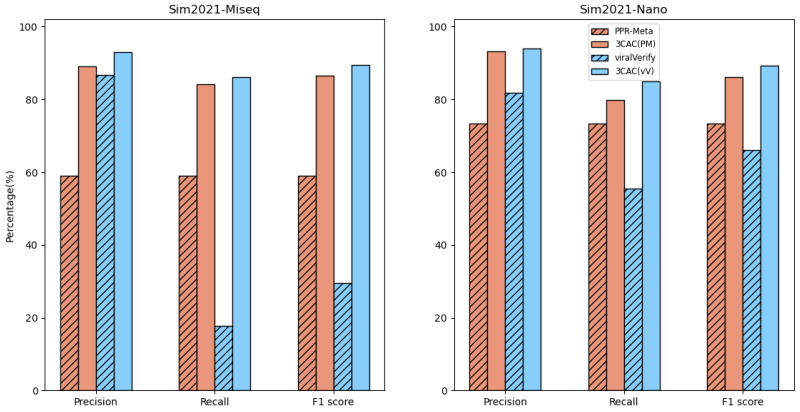

Fig. B.2. Performance of three-class classifiers on metagenomes simulated from newly released genomes.