

DOI: [https://doi.org/10.71317/kjard.2.9\(s\).2026.647](https://doi.org/10.71317/kjard.2.9(s).2026.647)

Kashmir Journal of Academic Research and Development

Journal homepage: <https://rjsaonline.org/index.php/KJARD>

Plasmid-Resolved Metagenomics for Tracking Horizontal Transfer of Antimicrobial Resistance Determinants Across Microbial Communities

Arooj Khan

IMBB/CRIMM (Institute of Molecular Biology and Biotechnology/Centre for Research in Molecular Medicine), University of Lahore (UOL)

Email: arooj6098@gmail.com**Maham Waheed**

IMBB, The University of Lahore

Email: waheedmaham352@gmail.com**Muhammad Medrar Hussain**

Abasyn University, Peshawar

Email: medrarbiotech@gmail.com**ARTICLE INFO****ABSTRACT****Received:**

August 08, 2026

Revised:

August 25, 2026

Accepted:

September 07, 2026

Available Online:

September 20, 2026

Keywords:

Plasmid-resolved metagenomics; antimicrobial resistance; horizontal gene transfer; long-read sequencing; host-plasmid association; One Health; bioinformatics; mobile genetic elements

Corresponding Author:**Arooj Khan**Email: arooj6098@gmail.com

The spread of antimicrobial resistance (AMR) throughout microbial communities is a serious concern for both human and environmental health, since plasmids act as the main carriers for the horizontal gene transfer (HGT) of antibiotic resistance genes (ARGs). Yet, conventional short-read metagenomics has in the past been unable to determine the extrachromosomal context of ARGs, thereby hiding the mechanisms involved in the spread of resistance. Plasmid-resolved metagenomics has now become a groundbreaking approach that brings together third-generation long-read sequencing (using PacBio HiFi and Oxford Nanopore R10.4), improved assembly algorithms, machine learning classifiers, and proximity ligation assays (Hi-C) in order to fully reconstruct plasmids directly from complex microbial communities. This article brings together the technological frameworks, bioinformatic advances, and ecological findings that are shaping this quickly developing area. We carry out a critical assessment of the long-read assemblers such as hifiasm-meta, metaMDBG, metaFlye, HiCanu, and myloasm, showing that they are capable of recovering complete circular plasmids in a variety of environmental samples. We look at the computational methods used for identifying and separating plasmids, pointing out geNomad's deep neural network design and PlasMAAG's assembly-alignment graph method based on contrastive learning, both of which allow for accurate plasmid reconstruction even in complicated metagenomes. We also consider how Hi-C-based proximity ligation helps to establish associations between hosts and plasmids, finding that plasmids carrying AMR have a much wider range of phylogenetic host species than those that do not carry resistance, and thus form well-connected transmission networks across the various One Health interfaces such as wastewater treatment systems, aquaculture sediments, and microplastic plastispheres. Although there have been major progressions, difficulties still remain in resolving plasmid microdiversity, in standardising benchmarking protocols, and in experimentally verifying computational predictions. The future integration of this approach with single-cell microfluidics and stable isotope probing will enable the direct measurement of plasmid transfer kinetics, which in turn will allow for proactive monitoring and control of the spread of AMR mediated by plasmids in interconnected ecosystems.

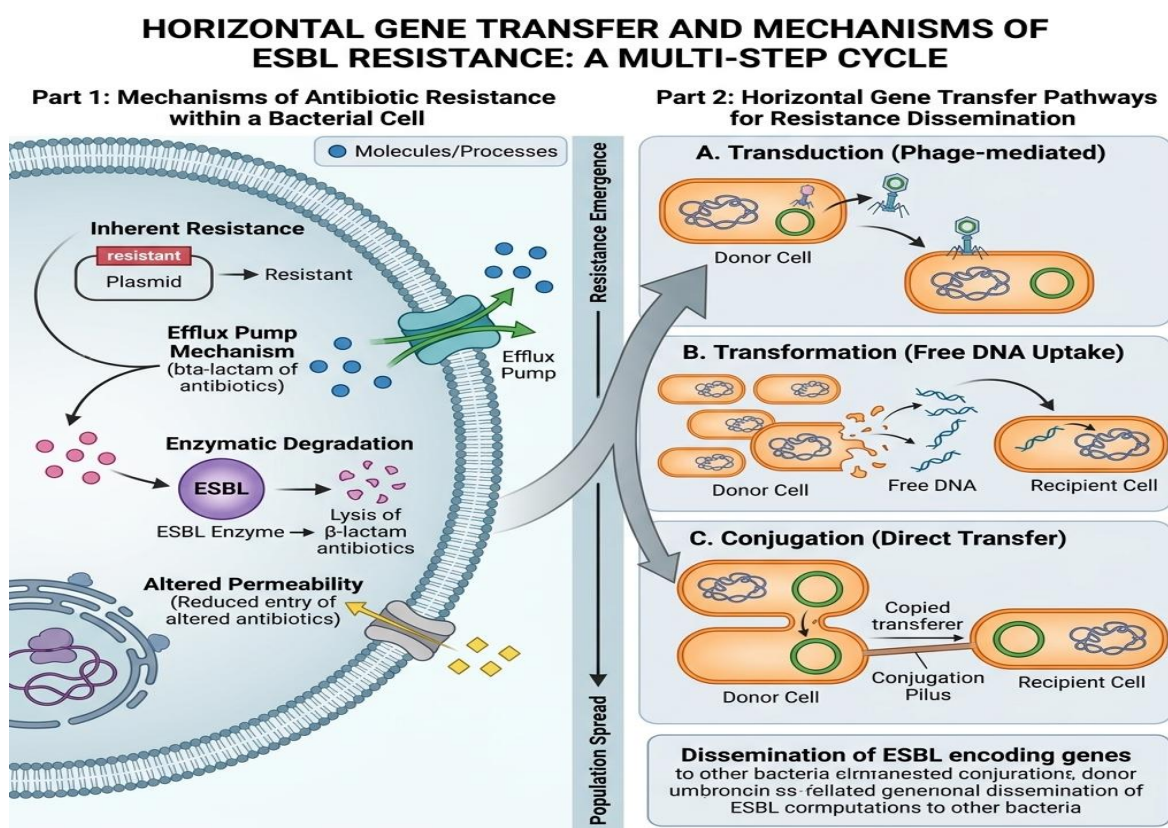
Introduction

The quick spread of antimicrobial resistance (AMR) throughout microbial communities represents a serious threat to clinical medicine, to veterinary health, and to environmental sustainability. Although mutations at particular points in chromosomal genes lead to treatment failure, most of the transmission of antibiotic resistance genes (ARGs) among different bacterial groups is due to horizontal gene transfer (HGT) (Castañeda-Barba et al., 2024). Of the main mobile genetic elements—these comprising transposons, insertion sequences, integrons, and integrative conjugative elements—plasmids act as the main independent vectors for the spread of multi-drug resistance. They enable the physical transfer of complex, multi-gene resistance cassettes over wide phylogenetic distances and thus create genetic links that connect environmental matrices, agricultural reservoirs, and human clinical pathogens as part of the One Health approach (Rodríguez-Beltrán et al., 2021).

In the past, studies of the ecology and evolutionary paths of plasmid-mediated antimicrobial resistance (AMR) have depended on sequencing of cultured isolates or on conventional short-read shotgun metagenomics. Yet short-read shotgun metagenomics generally results in sequence data being collapsed into highly fragmented contigs, and thus fails to associate specific ARGs with their underlying genomic or extrachromosomal contexts (Brito, 2021). Since plasmids have a mosaic structure and often share repetitive elements, transposons, and insertion sequences with their chromosomal backbones, short-read assembly programmes usually produce fragmented assembly graphs. As a result, researchers have frequently been unable to determine whether an ARG is located on a broad-host-range conjugative plasmid or on a stable chromosomal site, which in turn blurs the exact mechanisms responsible for the spread of resistance (Ghaly & Gillings, 2022).

Plasmid-resolved metagenomics has become a highly effective method which allows the complex extrachromosomal structures to be separated out directly from microbiomes that cannot be cultivated. Through the use of third-generation long-read sequencing, graph-aware assembly programs, proximity ligation assays, and specific machine learning classification tools, plasmid-resolved metagenomics makes it possible to fully reconstruct the extrachromosomal components from complex microbial communities (Wang et al., 2023). The present review offers a thorough summary of the technological approaches, bioinformatic advancements, host-linking strategies, and ecological findings that define plasmid-resolved metagenomics, emphasizing its important function in monitoring the horizontal spread of antimicrobial resistance genes in both natural and man-made ecosystems (Pellow et al., 2021).

Figure 1: Intracellular Mechanisms of ESBL-Mediated Antibiotic Resistance and Subsequent Horizontal Gene Transfer Dissemination



Sequencing Paradigms and Assembly Architectures for Metagenomic Plasmid Resolution

Limitations of Short-Read Shotgun Sequencing

For more than twenty years, second-generation short-read sequencing technologies, especially those made by Illumina, have been the mainstay of metagenomic research because of their low cost per base and their high accuracy for individual reads (greater than 99.9%). Although they have these advantages, short reads (usually between 150 and 300 base pairs) have inherent structural drawbacks when it comes to dealing with extrachromosomal DNA. About 97% of the bacterial genome assemblies found in public repositories come from short-read data, and this greatly hampers the structural accuracy of the annotations for mobile elements (Hu et al., 2021). These short-read assemblers create genome assemblies by using de Bruijn graphs that are based on overlapping k-mers. But when the length of the k-mers is less than that of the repetitive insertion sequences (for example, IS26, ISCR1) and the transposable elements adjacent to ARGs, the assembly graph splits and collapses, making it impossible to fully resolve the sequence (Li et al., 2015).

The computational collapse results in serious analytical bottlenecks in the field of mobile element genomics. Due to structural mosaicism and modular recombination, the assembly graphs become entangled and thus the full-length plasmid backbones cannot be reconstructed. Moreover, natural microbial communities contain plasmid variants that differ from one another by single-nucleotide variants or minor indels (Poot, 2024). De Bruijn graphs usually collapse these variant structures down to a single chimeric consensus sequence or fail to carry out the assembly at all. As a result, the short-read assemblies produce short contigs containing ARGs which lack the necessary replicon, relaxase, or host marker sequences, so it is impossible to tell whether an expansion is due to clonal chromosomal spread or to active horizontal plasmid transfer (Antipov et al., 2016).

Third-Generation Long-Read and Hybrid Sequencing Architectures

The development of third-generation long-read sequencing technology, which is mainly the development of PacBio HiFi reads and the development of ONT R10.4 technology, has eliminated the above obstacles (Wenger et al., 2019). PacBio HiFi reads provide read lengths of 15-25 kilobases (kb) with base accuracies greater than Q40 (>99.99%) via circular consensus sequencing. The ONT R10.4 chemistry provides continuous reads of tens to hundreds of kilobases with mean accuracies greater than 99% and can traverse ultra-long repeat regions (Feng et al., 2022).

The long reads can easily cover complex transposable elements, multi-drug resistance regions, full-length operons, and structural repeats within one read. Evaluations on systematic benchmarks have indicated that long-read metagenomics provides 44 times to 64 times more complete metagenome-assembled genomes (cMAGs) per gigabase of sequence data compared to short-read methods (Kim et al., 2022). PacBio HiFi technology can offer great cost-effectiveness when used for the direct recovery of closed, circular genomes of microorganisms from feces and environmental samples. Although the early bridge was provided by hybrid approaches that utilize short reads with high coverage to correct errors of long reads, it is now preferred to employ single-platform HiFi and ONT R10.4 approaches due to their simplicity of library preparation and lower composition biases (Kolmogorov et al., 2020).

Comparative Evaluation of Long-Read Metagenomic Assemblers

Assembling metagenome assemblies via long reads poses additional challenges to conventional algorithms than those required in assembly from single isolate. The metagenome assembly has a high number of species diversity, huge variations in coverage depth for abundant and rare taxa, and strains micro-diversity. Long-read specific assemblers use overlap-layout-consensus or string graph formulation to assemble complete circular plasmids along with the host genome (Ruan & Li, 2020).

The hifiasm-meta is an extension of the haplotype-resolved assembler hifiasm used specifically in PacBio HiFi metagenomes. It uses coverage-aware graph pruning and containment-read protection to phase the closely related strain variants that have an average nucleotide identity of $\geq 99.9\%$. From empirical testing, hifiasm-meta shows extremely low misassembly and chimera rate (0.001 real chimeric reads per Mbp) (Nurk et al., 2017).

MetaMDBG uses a minimizer-based de Bruijn graph for quick processing of large amounts of data in the long-reads. Although it is fast at creating a large number of circular contigs through deep sequencing, from comparisons, it can have high levels of contig clipping and false circularization events especially in complex environments such as surface ocean metagenome (Benoit et al., 2024).

The MetaFlye uses repeat graph algorithm whose parameterizations depend on variation of coverage in the community. It is good at resolving structural repeats and circular plasmids despite the uneven abundance distribution. It, however, may sometimes collapse intra-species strain heterogeneity into the species level consensus sequence (Kolmogorov et al., 2019).

The HiCanu modifies the classic OLC pipeline to accommodate high accuracy long reads. Although HiCanu has high structural fidelity and solves long inverted repeats problems, the high computation power requirements limit its scalability in the large metagenomic analysis (Nurk et al., 2020).

The newly developed myloasm assembler can work efficiently with both PacBio HiFi and ONT R10.4 reads. The improved error rate model in ONT R10.4 chemistry has allowed it to achieve the contiguity and circular plasmid assembly metrics equivalent to the traditional HiFi assemblies (Shaw et al., 2026).

Table 1: Performance Benchmarks and Algorithmic Characteristics of Metagenomic Assemblers for Plasmid Recovery

Assembler	Target Sequencing Platform	Graph Architecture / Core Algorithm	Strengths in Metagenomic Plasmid Recovery	Error Profiles and Operational Limitations
hifiasm-meta	PacBio HiFi	Haplotype-phased string graph with containment preservation	Exceptional strain phasing ($\geq 99.9\%$ ANI); lowest chimera and misassembly rate (0.001 per Mbp); highly accurate plasmid circularization	High memory requirements; optimized strictly for high-accuracy ($> Q30$) long-read inputs
metaMDBG	PacBio HiFi / ONT R10.4	Low-memory minimizer de Bruijn graph (mDBG)	Ultra-fast execution time; yields high total assembled output and large volume of circular contigs from deep datasets	Increased rate of contig clipping and false circularization in complex metagenomes (up to 46 errors per 100 Mb)
metaFlye	ONT (R9.4/R10.4) & PacBio	Disentangled repeat graph with adaptive coverage thresholds	Highly robust across extreme coverage gradients; excellent recovery of diverse mobile element sizes	Tends to collapse intra-species strain heterogeneity into species-level consensus sequences
HiCanu	PacBio HiFi / Oxford Nanopore	Modified Overlap-Layout-Consensus (OLC) with homopolymer suppression	High structural fidelity; rigorous resolution of long repetitive plasmid backbones and inverted repeats	Extremely high computational resource requirements; prone to graph fragmentation in low-coverage regions
myloasm	ONT R10.4 / PacBio HiFi	Multi-k-mer overlap graph with precise error modeling	Recovers complete circular genomes from ONT R10.4 data matching or exceeding HiFi performance; high sensitivity	Emerging tool; requires broader validation across diverse soil and sediment matrices

Computational Approaches for Plasmid Identification, Reconstruction, and Binning

Plasmid Detection and Machine Learning Classification

After an assembly separating plasmid contigs from chromosomal fragments is a big bioinformatics problem. Standard metagenomic binning tools such as CheckM2 and MetaBAT2 use tetranucleotide frequency (TNF) and differences in coverage between samples (Krawczyk et al. 2018). Because plasmids normally have a nucleotide composition changing copy numbers or varying amounts over time normal binners often put plasmid contigs into host chromosomal MAGs or leave them as unbinned contigs (Kang et al., 2019).

Plasmid informatics tools are usually divided into detection algorithms, which decide if each contig is a plasmid or chromosomal and plasmid reconstruction software, which puts together contigs that come from the same real plasmid. Machine-learning classifiers now use more than sequence matching to reference databases. They also use natural language processing, k-mer embeddings, deep neural networks and random forests (Krawczyk et al. 2018).

The geNomad framework gives a method for finding mobile genetic elements. GeNomad mixes a neural network that looks at sequence signatures with a conditional random field that checks gene structure, marker genes such as relaxases, replication

initiation proteins, type IV secretion systems and gene density patterns. Tests show that tools like geNomad, Plasmer, PlasmidEC and gplas2 find Gram-negative Enterobacteriaceae plasmids accurately (Li et al., 2026). The accuracy can drop when the samples come from scarce Gram-positive groups, like Enterococcus species or new environmental isolates because the databases are not well represented (Camargo et al., 2024).

Graph-Based Metagenomic Plasmid Binning and Assembly-Alignment Networks

Because big conjugative plasmids often break into parts across different contigs in short-read or mixed assemblies putting these parts together into full plasmid units needs special binning methods. Early programs like SCAPP and MOB-suite used circular path finding inside one-sample assembly graphs or distance-based grouping using fixed database reference genomes.. One-sample assembly graphs often break apart because of a lot of small sequence differences not enough coverage or tangled structures (Pradier et al. 2021).

To fix these problems better computer methods use information from samples and graph neural networks. One example is PlasMAAG (Plasmid and Organism Metagenomic Binning using Assembly-Alignment Graphs). PlasMAAG joins each samples assembly graphs with alignment graphs between samples to make an Assembly-Alignment Graph (AAG). Contigs are moved into an embedding space using fastnode2vec and communities of contigs are found, cleaned and improved using a Variational Autoencoder (VAE) that uses loss (Wang et al., 2026). At the end PlasMAAG takes geNomad annotation scores across all contigs in a community of looking at single contigs, which makes the classification better. Tested on data PlasMAAG found 50% to 121% more nearly complete plasmids on made-up tests and 33% more plasmid parts, in actual hospital sewage samples compared to methods that look at one sample at a time (Piera L ndez et al. 2026).

Reference Databases and Curation Infrastructure

Large conjugative plasmids are often split into pieces, called contigs when we use short-read or hybrid assembly methods. To bring these pieces into full plasmid units we need special binning algorithms. Early tools like SCAPP and MOB-suite looked for paths in single-sample assembly graphs or grouped contigs based on distance from fixed reference genomes.. Single-sample assembly graphs are often broken because of high microdiversity in the sequences, shallow coverage or tangled structures (Pradier et al., 2021).

To annotate metagenomes accurately we need fresh well-checked reference collections. The PLSDb database is a source of complete bacterial and archaeal plasmid sequences. In its 2025 release PLSDb added 72,360 plasmid entries. PLSDb uses seqkit2 to remove sequences maps BioSample metadata to connect plasmids with host diseases and ecological settings classifies plasmid mobility as conjugative, mobilizable or non-mobilizable marks biosynthetic gene clusters and includes ARG detection (Schmartz et al., 2022).

When profiling resistance we compare contigs with specialized resistance databases, especially the Comprehensive Antibiotic Resistance Database or CARD. CARD uses Protein Homology and Protein Variant Models to label resistance mechanisms. These include enzymes that break down antibiotics, such, as extended-spectrum β -lactamases and carbapenemases efflux pump systems and mutations that change target sites. This approach gives a map of resistomes carried on plasmids (Alcock et al. 2023).

Table 2: Bioinformatic Tools and Computational Frameworks for Metagenomic Plasmid Analysis

Tool / Resource	Tool Category	Primary Algorithm / Framework	Underlying /	Key Analytical Capabilities	Reference Dependencies & Limitations
geNomad	MGE Identification & Classification	Deep neural networks combined with conditional random fields (CRFs)	with random	Accurately classifies contigs into plasmid, viral, or chromosomal origins based on gene markers and sequence structure	Precision drops slightly on novel environmental backbones underrepresented in reference marker sets
PlasMAAG	Multi-Sample Plasmid Binning	Assembly-Alignment Graphs (AAG), fastnode2vec, VAE with contrastive learning		Reconstructs complete plasmids across multiple samples; geNomad scores over contig communities	Requires multi-sample datasets and assembly graphs; higher memory usage during graph embedding

MOB-suite	Plasmid Reconstruction & Typing	Pairwise Mash distance clustering, replicon, and relaxase typing	Reconstructs plasmid contigs; classifies plasmids into precise MOB families and mating-pair formation groups	Dependent on reference databases; performance declines when analyzing novel or uncultured plasmid backbones
gplas2	Plasmid Reconstruction	De Bruijn assembly graph extraction with neural network scoring	Clusters short-read assembly graph nodes into discrete plasmid components based on coverage and sequence content	Sensitive to graph fragmentation in low-coverage or highly mosaic metagenomic samples
PLSDB (v2025)	Reference Database	Curated repository of complete bacterial and archaeal plasmids	Hosts 72,360 complete entries; incorporates BioSample ecological metadata, mobility typing, and ARG mapping	Restricted to complete plasmid records; does not index fragmented draft contigs from unpolished metagenomes
CARD (v4.0+)	Resistance Database	Ontology-driven curation (ARO) with BLASTp and Resistance Gene Identifier (RGI)	High-precision annotation of ARGs, efflux systems, and resistance-modifying agents	Requires high sequence accuracy to distinguish specific single-nucleotide resistance variants

Unraveling Host-Plasmid Associations and In Situ Transfer Dynamics

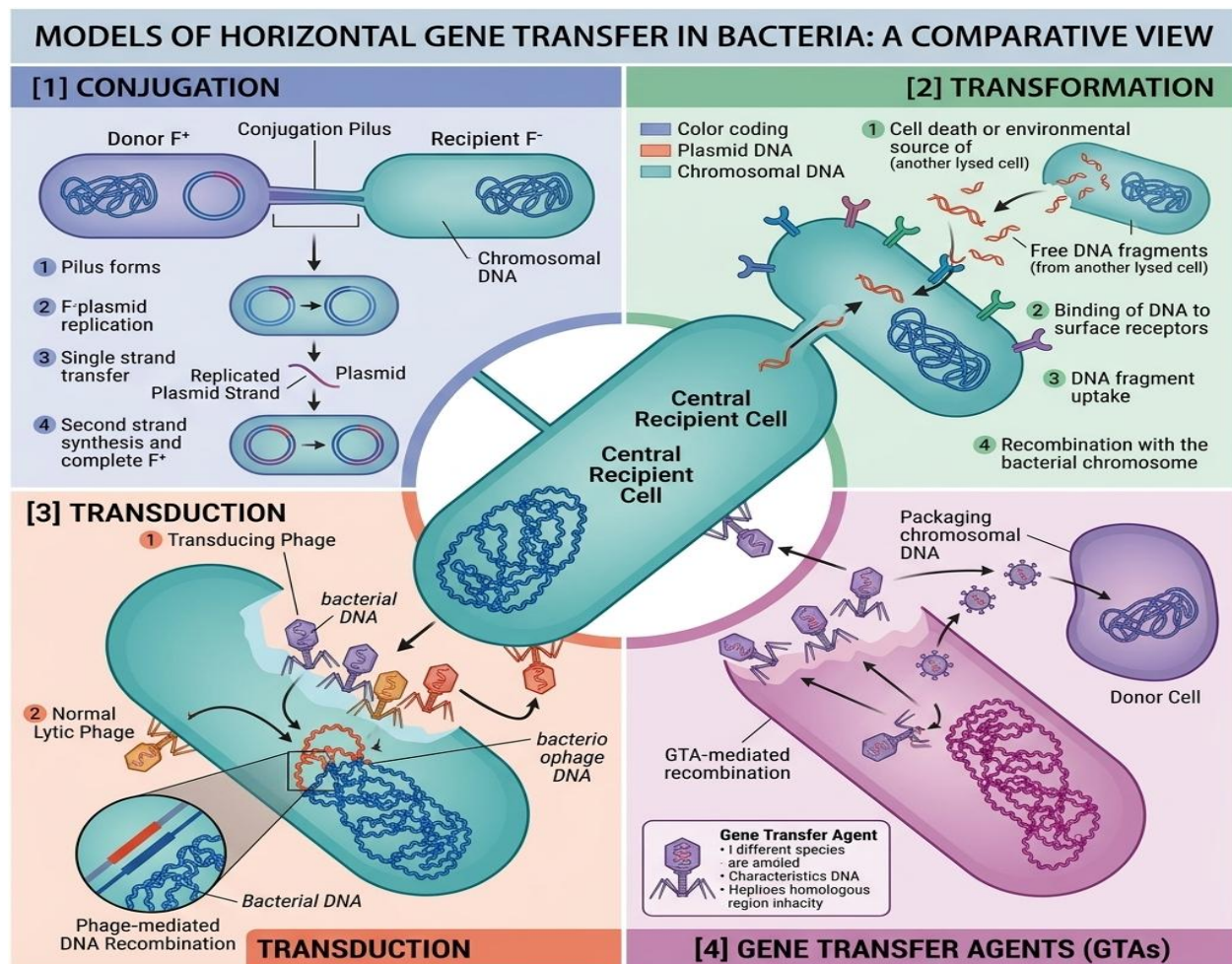
Proximity Ligation Metagenomics (Hi-C) for Physical Host-Plasmid Linkage

While read sequencing shows what is inside plasmids linking a specific plasmid to its bacterial host in a mix of different microbes is still hard. Methods that try to find the host using CRISPR spacers, prophage signals or similar DNA patterns are based on chance. Often miss plasmids that can work with many different bacteria, especially those that are not easy to grow in the lab (Beitel et al., 2014).

Proximity ligation metagenomics, also known as Hi-C helps find the connection between a plasmid and its host by looking at how DNA's placed inside the cell before it breaks apart. In a Hi-C process whole microbes in a sample are treated with formaldehyde. This makes the DNA that's close to each other stick together inside the cell. The DNA is cut with enzymes made blunt with biotin connected under weak conditions and then broken into small pieces. Pieces that have biotin are. Then read (Stalder et al., 2019).

When used on samples that have different microbes Hi-C creates maps that show how plasmid DNA is connected to the DNA of the host bacteria. This way no need to grow the bacteria in a lab. It directly shows which plasmids belong to which bacteria in environments, like wastewater, human poop and soil (Bickhart et al., 2019).

Figure 2: Comparative Overview of Bacterial Horizontal Gene Transfer Pathways and Gene Transfer Agent (GTA)-Mediated Recombination



Network Topologies of Plasmid-Borne AMR Dissemination

Integrating plasmid-resolved metagenomics with Hi-C proximity capture allows researchers to build host-plasmid interaction networks. In these networks bacterial metagenome-assembled genomes (MAGs) and putative plasmid bins act as nodes. Normalized Hi-C crosslinking contacts form the edges that show connections between hosts and plasmids (Risely et al. 2024).

When researchers look at the structure of these networks, they find differences between plasmids that carry antibiotic resistance genes (ARGs) and those that do not. Plasmids with ARGs have a higher degree centrality meaning they interact with different bacterial host MAGs compared to plasmids without resistance markers in the same environment. Also, AMR-bearing plasmids are found in a range of bacterial hosts. They can link across taxonomic groups, such as between Gammaproteobacteria, Alphaproteobacteria and Bacilli (Given et al., 2022).

In evolutionary theory keeping large conjugative plasmids usually comes with a cost. The host cell pays a metabolic price when no antibiotics are present. In environments shaped by human activity the benefit of having antibiotic resistance genes, on the plasmid is strong enough to outweigh that cost. This powerful selection helps resistance plasmids stay in place across many different bacterial hosts. As a result, host-plasmid networks become highly connected and very stable forming transmission webs (Robertson & Nash 2018).

Plasmid Dynamics in Environmental and One Health Interfaces

Plasmid-mediated HGT happens at rates in environmental hotspots. These tiny places have lots of cells, thick biofilm layers, plenty of food. Are constantly exposed to low levels of things that select for resistance, such as antibiotics, heavy metals or biocides (Ahmad et al. 2023). Plasmid-mediated HGT is therefore very active in these settings.

Municipal wastewater treatment plants are the mixing spots where clinical pathogens meet environmental saprophytes. Hi-C network studies of wastewater show webs of interaction, where broad-host-range plasmids spread extended-spectrum β -lactamases such as bla_{CTX-M} and carbapenemases like bla_{KPC} bla_{NDM} across many bacterial hosts. Plasmid-mediated HGT and plasmid exchange are processes in this environment (Carattoli et al. 2014).

In aquaculture sediments many stresses act together heat, redox pressure, low oxygen and antibiotics. Plasmid-resolved metagenomic studies show that higher temperatures and redox changes boost gene transfer a lot. Genome-context analysis shows that antibiotic resistance genes sit close to class 1 integrons (intI1) insertion sequences and conjugative transfer operons speeding the growth of the resistome when climate stress hits. Horizontal gene transfer and plasmid movement are accelerated in this setting (Zhou et al. 2023).

Microplastic particles that build up in water create biofilm communities called the plastisphere. Metagenomic comparison shows that plastisphere biofilms hold conjugative plasmids with broad-host-range backbones such as IncFIB and IncFII. These conjugative plasmids link bacterial genera *Pseudomonas*, *Vibrio*, *Sphingomonas*, *Rhizobium* allowing quick transfer of multi-drug resistance genes, inside the biofilm across both marine and freshwater habitats. Plasmid-mediated HGT. Plasmid spread continue to shape resistance in these systems (Harrison et al. 2018).

Table 3: Plasmid Dynamics, Replicon Distributions, and Host Profiles Across One Health Interfaces

Environmental Interface / Hotspot	Dominant Plasmid Replicons & Mobility Profiles	Primary Host Taxa & Phylogenetic Breadth	Environmental Drivers & Selective Pressures	Dissemination Mechanisms & Resistance Impact
Municipal Wastewater Systems	Conjugative IncF, IncA/C, IncL/M; high proportion of mobilizable elements	Broad phylogenetic range; bridges Gammaproteobacteria, Bacilli, and Bacteroidota	Sub-inhibitory antibiotics, heavy metals, quaternary ammonium compounds	Bipartite network hub expansion; wide horizontal spread of bla _{CTX-M} , bla _{NDM} , and aminoglycoside resistance
Aquaculture Sediments	Integron-associated conjugative plasmids; high frequency of IS elements	Sediment microbiomes, <i>Vibrio</i> , <i>Aeromonas</i> , and <i>Pseudomonas</i> spp.	Thermal stress, hypoxia, redox compression, oxytetracycline loading	Climate-driven acceleration of conjugation frequencies; co-selection of heavy metal and antibiotic resistance
Microplastic Plastispheres	Conjugative IncFIB, IncFII, and novel uncharacterized rep_cluster backbones	Core plastisphere genera: <i>Pseudomonas</i> , <i>Vibrio</i> , <i>Sphingomonas</i> , <i>Paracoccus</i>	High local cell density, persistent surface biofilms, hydrophobic organic pollutants	Biofilm-mediated surface conjugation; rapid cross-species exchange of multi-drug resistance cassettes
Clinical-Environmental Boundary	Broad-host-range incompatibility groups (IncP-1, IncW, IncQ)	Cross-phylum transfer between environmental saprophytes and clinical Enterobacteriaceae	Antibiotic therapy selection, hospital effluent discharges	Transfer of carbapenemases and colistin resistance (mcr) from environmental reservoirs to clinical pathogens

Methodological Challenges, Standardization, and Future Perspectives

Despite progress in bioinformatics and sequencing plasmid-resolved metagenomics still faces many technical roadblocks that must be solved carefully. Plasmid populations in nature behave like quasispecies constantly exchanging modules, recombining and rearranging their structures (Olawade et al. 2025). When we use long reads from PacBio HiFi or ONT R10.4 the assembly graphs can become tangled if several closely related plasmid backbones exist in the same microbial community. Better

algorithms for haplotype phasing and for untangling reference-free assembly graphs are essential to resolve intra-sample plasmid microdiversity (Darby et al., 2023).

A second operational challenge involves plasmid copy number variations and binning bias. Plasmid copy numbers vary widely inside cells from single-copy broad-host-range conjugative plasmids to high-copy-number small mobilizable plasmids that exceed one hundred copies per cell (Shahzadi et al. 2026). These differences distort the sequence-coverage signals that metagenomic binning algorithms use, often causing high-copy plasmids to be misassigned or broken apart. Adding quantitative coverage-normalization steps to the pipeline front-ends is necessary to keep binning (Kang et al., 2015).

Standardization and benchmarking remain priorities. Validating assemblers and binning pipelines requires realistic benchmark datasets where host-plasmid connections and sequence identities are known for sure. Past benchmarks relied mainly on data or on mixtures of isolates dominated by *E. Coli*. Creating mock communities that contain structurally diverse, low-abundance, Gram-positive and environmental plasmids is critical, for fair algorithm testing (Chklovski et al. 2023).

Finally, metagenomic predictions of host range and transferability stay only hypotheses until we test them in the lab. Combining metagenomics with high-throughput single-cell microfluidics, fluorescent reporters and stable-isotope probing (SIP) will allow us to measure plasmid transfer kinetics directly and in real time under controlled environmental gradients (Gallegos et al. 2020).

Conclusion

Plasmid-resolved metagenomics has completely changed how we follow the spread of resistance between different types of microbes. By using read sequencing better ways to put the genetic pieces together machine learning tools and methods that look at how close different parts are to each other scientists can now build full extra parts of DNA directly from microbes that we have never grown in the lab. This helps them see the parts and which microbes they belong to something that was hard before with short-read methods. Some tools like hifiasm-meta, metaMDBG, metaFlye and myloasm have done a job in finding the round shapes of plasmids in complicated environments. Other tools such as geNomad and PlasMAAG help. Group plasmids more accurately than ever before. Importantly a method called Hi-C has shown that plasmids with resistance are found in different types of hosts than plasmids without resistance. These plasmids connect in networks across areas like wastewater plants, places where fish are raised and tiny plastic parts in water. These findings show how important it is to watch for plasmid-related resistance in hospitals, farms and nature. There are still problems like figuring out small differences in plasmids making sure tests are the same everywhere and checking computer results with real experiments. In the future combining plasmid-resolved metagenomics with tools that look at cells ways to track how organisms use food and tests that see how plasmids move in real time could help measure how fast plasmids spread. This would help plan ways to stop problems before they get worse. As the tools to read DNA get better and the computer programs get more developed plasmid-resolved metagenomics will keep being a part of understanding and dealing with the worlds problem with antimicrobial resistance, in the connected world of people, animals and the environment.

References

- Ahmad, N., Joji, R. M., & Shahid, M. (2023). Evolution and implementation of One Health to control the dissemination of antibiotic-resistant bacteria and resistance genes: A review. *Frontiers in Cellular and Infection Microbiology*, 12, 1065796.
- Alcock, B. P., Huynh, W., Chalil, R., Smith, K. W., Raphenya, A. R., Wlodarski, M. A., Edalatmand, A., Petkau, A., Syed, S. A., Tsang, K. K., Baker, S. J. C., Dave, M., McCarthy, M. C., Mukiri, K. M., Nasir, J. A., Golbon, B., Imtiaz, H., Jiang, X., Kaur, K., ... Gray, K. L. (2023). CARD 2023: Expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance Database. *Nucleic Acids Research*, 51(D1), D690–D699. <https://doi.org/10.1093/nar/gkac920>
- Antipov, D., Hartwick, N., Shen, M., Raiko, M., Lapidus, A., & Pevzner, P. A. (2016). plasmidSPAdes: Assembling plasmids from whole genome sequencing data. *Bioinformatics*, 32(22), 3380–3387.
- Beitel, C. W., Froenicke, L., Lang, J. M., Korf, I. F., Micheltore, R. W., Eisen, J. A., & Darling, A. E. (2014). Strain- and plasmid-level deconvolution of a synthetic metagenome by sequencing proximity ligation products. *PeerJ*, 2, e415.
- Benoit, G., Raguideau, S., James, R., Phillippy, A. M., Chikhi, R., & Quince, C. (2024). High-quality metagenome assembly from long accurate reads with metaMDBG. *Nature Biotechnology*, 42, 1180–1188.

- Bickhart, D. M., Watson, M., Koren, S., Panke-Buisse, K., Cersosimo, L. M., Press, M. O., et al. (2019). Assignment of virus and antimicrobial resistance genes to microbial hosts in a complex microbial community by combined long-read assembly and proximity ligation. *Genome Biology*, 20, 153.
- Brito, I. L. (2021). Examining horizontal gene transfer in microbial communities. *Nature Reviews Microbiology*, 19, 442–453.
- Camargo, A. P., Roux, S., Schulz, F., Babinski, M., Xu, Y., Hu, B., Chain, P. S. G., Nayfach, S., & Kyrpides, N. C. (2024). Identification of mobile genetic elements with geNomad. *Nature Biotechnology*, 42, 1303–1312.
- Carattoli, A., Zankari, E., García-Fernández, A., Larsen, M. V., Lund, O., Villa, L., Aarestrup, F. M., & Hasman, H. (2014). In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrobial Agents and Chemotherapy*, 58(7), 3895–3903.
- Castañeda-Barba, S., Top, E. M., & Stalder, T. (2024). Plasmids, a molecular cornerstone of antimicrobial resistance in the One Health era. *Nature Reviews Microbiology*, 22(1), 18–32.
- Chklovski, A., Parks, D. H., Woodcroft, B. J., & Tyson, G. W. (2023). CheckM2: A rapid, scalable and accurate tool for assessing microbial genome quality using machine learning. *Nature Methods*, 20(8), 1203–1212.
- Darby, E. M., Trampari, E., Siasat, P., Gaya, M. S., Alav, I., Webber, M. A., & Blair, J. M. A. (2023). Molecular mechanisms of antibiotic resistance revisited. *Nature Reviews Microbiology*, 21(5), 280–295. <https://doi.org/10.1038/s41579-022-00820-y>
- Feng, X., Cheng, H., Portik, D., & Li, H. (2022). Metagenome assembly of high-fidelity long reads with hifiasm-meta. *Nature Methods*, 19, 671–674.
- Gallegos, J. E., Rogers, M. F., Cialek, C. A., & Peccoud, J. (2020). Rapid, robust plasmid verification by de novo assembly of short sequencing reads. *Nucleic Acids Research*, 48(18), e106. <https://doi.org/10.1093/nar/gkaa727>
- Ghaly, T. M., & Gillings, M. R. (2022). New perspectives on mobile genetic elements: A paradigm shift for managing the antibiotic resistance crisis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 377(1842), 20200462.
- Given, C., Penttinen, R., & Jalasvuori, M. (2022). Plasmid viability depends on the ecological setting of hosts within a multiplasmid community. *Microbiology Spectrum*, 10(2), e00133-22.
- Harrison, J. P., Hoellein, T. J., Sapp, M., Tagg, A. S., Ju-Nam, Y., & Ojeda, J. J. (2018). Microplastic-associated biofilms: a comparison of freshwater and marine environments. *Freshwater microplastics*, 58, 181-201.
- Hu, T., Chitnis, N., Monos, D., & Dinh, A. (2021). Next-generation sequencing technologies: An overview. *Human immunology*, 82(11), 801-811.
- Kang, D. D., Froula, J., Egan, R., & Wang, Z. (2015). MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ*, 3, e1165.
- Kang, D. D., Li, F., Kirton, E., Thomas, A., Egan, R., An, H., & Wang, Z. (2019). MetaBAT 2: An adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ*, 7, e7359.
- Kim, C. Y., Ma, J., & Lee, I. (2022). HiFi metagenomic sequencing enables assembly of accurate and complete genomes from human gut microbiota. *Nature Communications*, 13(1), 2915. <https://doi.org/10.1038/s41467-022-34149-0>
- Kolmogorov, M., Bickhart, D. M., Behsaz, B., Gurevich, A., Rayko, M., Shin, S. B., Kuhn, K., Yuan, J., Polevikov, E., Smith, T. P. L., & Pevzner, P. A. (2020). metaFlye: Scalable long-read metagenome assembly using repeat graphs. *Nature Methods*, 17, 1103–1110.
- Kolmogorov, M., Yuan, J., Lin, Y., & Pevzner, P. A. (2019). Assembly of long, error-prone reads using repeat graphs. *Nature Biotechnology*, 37, 540–546.
- Krawczyk, P. S., Lipinski, L., & Dziembowski, A. (2018). PlasFlow: predicting plasmid sequences in metagenomic data using genome signatures. *Nucleic acids research*, 46(6), e35-e35.
- Krawczyk, P. S., Lipinski, L., & Dziembowski, A. (2018). PlasFlow: Predicting plasmid sequences in metagenomic data using genome signatures. *Nucleic Acids Research*, 46(6), e35.

- Li, D., Liu, C.-M., Luo, R., Sadakane, K., & Lam, T.-W. (2015). MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*, 31(10), 1674–1676.
- Li, R., Dong, W., Yang, Z., Wang, M., Xiong, J., Ma, Y., ... & Li, J. (2026). MicroWorldOmics: All-in-one Desktop Solution for Microbiome Profiling, Virome Analysis, and Unexplored “Dark Matter” Discovery. *Genomics, Proteomics & Bioinformatics*, qzago59.
- Nurk, S., Meleshko, D., Korobeynikov, A., & Pevzner, P. A. (2017). metaSPAdes: A new versatile metagenomic assembler. *Genome Research*, 27(5), 824–834. <https://doi.org/10.1101/gr.213959.116>
- Nurk, S., Walenz, B. P., Rhie, A., Vollger, M. R., Logsdon, G. A., Grothe, R., Miga, K. H., Eichler, E. E., Phillippy, A. M., & Koren, S. (2020). HiCanu: Accurate assembly of segmental duplications, satellites, and allelic variants from high-fidelity long reads. *Genome Research*, 30(9), 1291–1305. <https://doi.org/10.1101/gr.263566.120>
- Olawade, D. B., Kade, A., Egbon, E., Usman, S. O., Fapohunda, O., Ijiwade, J., & Ogbonna, C. E. (2025). Bioinformatics and artificial intelligence in genomic data analysis: current advances and future directions. *Molecular Genetics and Genomics*, 300(1), 111.
- Pellow, D., Zorea, A., Probst, M., Furman, O., Segal, A., Mizrahi, I., & Shamir, R. (2021). SCAPP: An algorithm for improved plasmid assembly in metagenomes. *Microbiome*, 9, 144.
- Piera Líndez, P., Danielsen, L. S., Kovačić, I., Pielies Avelli, M., Nesme, J., Jensen, L. J., Andersen, J. N., Sørensen, S. J., & Rasmussen, S. (2026). Accurate plasmid reconstruction from metagenomics data using assembly–alignment graphs and contrastive learning. *Nature Biotechnology*. Advance online publication. <https://doi.org/10.1038/s41587-026-02612-w>
- Poot, M. (2024). Methods of detection and mechanisms of origin of complex structural genome variations. *Cancer Cytogenetics and Cytogenomics: Methods and Protocols*, 39-65.
- Pradier, L., Tissot, T., Fiston-Lavier, A.-S., & Bedhomme, S. (2021). PlasForest: A homology-based random forest classifier for plasmid detection in genomic datasets. *BMC Bioinformatics*, 22, 349.
- Risely, A., Newbury, A., Stalder, T., Simmons, B. I., Top, E. M., Buckling, A., & Sanders, D. (2024). Host-plasmid network structure in wastewater is linked to antimicrobial resistance genes. *Nature Communications*, 15, 555.
- Robertson, J., & Nash, J. H. E. (2018). MOB-suite: Software tools for clustering, reconstruction and typing of plasmids from draft assemblies. *Microbial Genomics*, 4(8), e000206.
- Rodríguez-Beltrán, J., DelaFuente, J., León-Sampedro, R., MacLean, R. C., & San Millán, Á. (2021). Beyond horizontal gene transfer: The role of plasmids in bacterial evolution. *Nature Reviews Microbiology*, 19, 347–359.
- Ruan, J., & Li, H. (2020). Fast and accurate long-read assembly with wtdbg2. *Nature Methods*, 17, 155–158.
- Schmartz, G. P., Hartung, A., Hirsch, P., Kern, F., Fehlmann, T., Müller, R., & Keller, A. (2022). PLSDB: Advancing a comprehensive database of bacterial plasmids. *Nucleic Acids Research*, 50(D1), D273–D278.
- Shahzadi, I., Xue, W., Ubaid Ullah, H., Maddamsetti, R., You, L., & Wang, T. (2026). Integrating theory and machine learning to reveal determinants of plasmid copy number. *Nature Communications*.
- Shaw, J., Marin, M. G., & Li, H. (2026). High-resolution metagenome assembly for modern long reads with myloasm. *Nature Biotechnology*.
- Stalder, T., Press, M. O., Sullivan, S., Liachko, I., & Top, E. M. (2019). Linking the resistome and plasmidome to the microbiome. *The ISME Journal*, 13, 2437–2446.
- Wang, X. W., Wang, T., & Liu, Y. Y. (2026). Artificial intelligence for microbiology and microbiome research. *Cell Systems*, 17(2).
- Wang, X., Zhang, H., Long, X., Xu, X., Ren, H., Mao, D., Alvarez, P. J. J., & Luo, Y. (2023). Global increase of antibiotic resistance genes in conjugative plasmids. *Microbiology Spectrum*, 11(2), e04478-22.
- Wenger, A. M., Peluso, P., Rowell, W. J., Chang, P.-C., Hall, R. J., Concepcion, G. T., Ebler, J., Fungtammasan, A., Kolesnikov, A., Olson, N. D., Töpfer, A., Alonge, M., Mahmoud, M., Qian, Y., Chin, C.-S., Phillippy, A. M., Schatz, M. C., Myers, G.,

DePristo, M. A., ... Koren, S. (2019). Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome. *Nature Biotechnology*, 37(10), 1155–1162. <https://doi.org/10.1038/s41587-019-0217-9>

Zhou, Q., Zhang, J., Fang, Q., Zhang, M., Wang, X., Zhang, D., & Pan, X. (2023). Microplastic biodegradability dependent responses of plastisphere antibiotic resistance to simulated freshwater-seawater shift in onshore marine aquaculture zones. *Environmental Pollution*, 330, 121798. <https://doi.org/10.1016/j.envpol.2023.121798>



2026 by the authors; Journal of Kashmir Journal of Academic Research and Development. This is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).