Metagenomic and Bioinformatics Approaches to Enhance *Glomus intraradices*-Based Biofertilizer Systems

Natarajan Falquet
University of Basel, Petersplatz 1, 4001 Basel, Switzerland

* Corresponding Author: **Natarajan Falquet**

Article Info

P-ISSN: 3051-3421
E-ISSN: 3051-343X
Volume: 07
Issue: 01
January - June 2026
Received: 22-01-2026
Accepted: 20-02-2026
Published: 18-03-2026
Page No: 83-97

Abstract

Background: Arbuscular mycorrhizal fungí (AMF), especially *Glomus intraradices* (*Rhizophagus irregularis*), are a key component of sustainable agriculture because they enhance phosphorus uptake and improve soil structure as well as increase plant tolerance to biotic and abiotic stress. Nonetheless, limitations of traditional culture-dependent approaches have left the full metabolic and ecological potential of amf largely uncharacterized.

Methods: This research combines two metagenomic sequencing techniques—shotgun and amplicon-based—using Illumina, PacBio, and Oxford Nanopore sequencing technologies. Quality control pipelines were utilized to process the raw sequencing data and assemble the data together into a whole using metaSPAdes. Each metagenome was assessed in terms of its taxonomic classification with Kraken2, and in terms of its functional annotation using KEGG, Pfam, and CAZy databases. Multi-omics integration included the assessment of metagenomics, metatranscriptomics, and metabolomics, allowing for a more complete understanding of how the metagenome interacts with the environment and within an ecosystem.

Results: The study found 3 new gene variants of the phosphate transporter (GiPT1–3), 2 uncharacterized nitrogen assimilation pathways and 5 plant growth promoting genes not present in the reference genomes. The study also observed enhanced nutrient pools in plants (up to 47%) through the use of *G. intraradices*, *Bacillus subtilis* and *Pseudomonas fluorescens* together versus un-inoculated controls.

Conclusion: A thorough understanding of the functional capacity of *G. intraradices*-based biofertilizers is given by multi-omic integration. Results also indicated a decrease of approximately 25-40% in chemical phosphate fertiliser utilised and increased yield of wheat, maize, and legumes as well. This substantiates the potential to create next generation bio-fertilisers using precision agriculture principles and meeting the goal of global sustainbility.

DOI: https://doi.org/10.54660/JADR.2026.7.1.83-97

Keywords: Arbuscular mycorrhiza, metagenomics, biofertilizer, nutrient cycling, plant growth, soil microbiome, multi-omics, sustainable agriculture

1. Biological and Functional Characteristics of *Glomus intraradices*

1.1. Taxonomy, Morphology, and Phylogeny

According to molecular phylogenetic reclassification, *Glomus intraradices* is now generally considered to be synonymous with *Rhizophagus irregularis* (Schüssler *et al*., 2001) [2]. The genus Glomus is an ancient lineage that separates from all other fungal groups around 450 million years ago (Remy *et al*., 1994) [8]. The organism is aseptate, obligately biotrophic, and reproduces through large chlamydospores with thick walls, typically measuring between 50 and 200 µm in diameter (Smith and Read, 2008) [1]. Sequencing of its genome revealed a total length of ~153 Mb and a high proportion of repeat elements, as well as a very

high degree of genetic polymorphism among nuclei within a single cytoplasm (heterokaryosis) which presents difficulties in assembling and annotating the genome (Helber *et al*., 2011) [7]. Phylogenetically, *G. intraradices* is a member of he family Glomeraceae, with results from molecular marker analyses using SSU rRNA, LSU rRNA, and ITS regions consistently confirming its monophyly and with high bootstrap support (≥ 70) (Wang and Qiu, 2006) [39]. Table 1 provides a summary of some of *G. intraradices*' dominant taxonomic and physiological characteristics.

Table 1: Taxonomic and Physiological Traits of *Glomus intraradices*

Trait	Description
Kingdom	Fungi
Phylum	Glomeromycota
Class	Glomeromycetes
Order	Glomerales
Family	Glomeraceae
Genus/Species	*Glomus intraradices* (syn. *Rhizophagus irregularis*)
Spore size	50–200 µm; globose to sub-globose
Hyphal diameter	2–10 µm; aseptate
Arbuscule half-life	4–7 days within root cortical cells
Colonisation rate	30–80% root length colonisation under optimal conditions
Genome size	~153 Mb (*Rhizophagus irregularis* DAOM 197198)
Host range	Broadly polyphagous; >80% terrestrial plant species
Temperature optimum	15–30°C
Soil pH tolerance	5.5–8.0

Note: Taxonomic nomenclature follows SchÃ¼ssler *et al*. (2001) [2]; morphological data compiled from Smith and Read (2008) [1].

1.2. Arbuscular Mycorrhizal Associations with Plant Roots

The mutualistic symbiosis of *G. intraradices* with the roots of its host plant is one of the most evolutionarily conserved and ecologically significant biological interactions on Earth (Parniske, 2008) [38]. The recognition of strigolactone signals produced by root exudates provides the basis for *G. intraradices*' hyphae's development into appressoria on the root epidermis and is a crucial step in the pre-symbiotic colonisation of host plants (Bonfante and Genre, 2010) [4]. This development of appressoria on the root epidermis is part of a highly coordinated process leading to the pre-symbiotic colonisation of host plants; it is coordinated using a series of molecular events initiated by the recognition of strigolactone by *G. intraradices*.

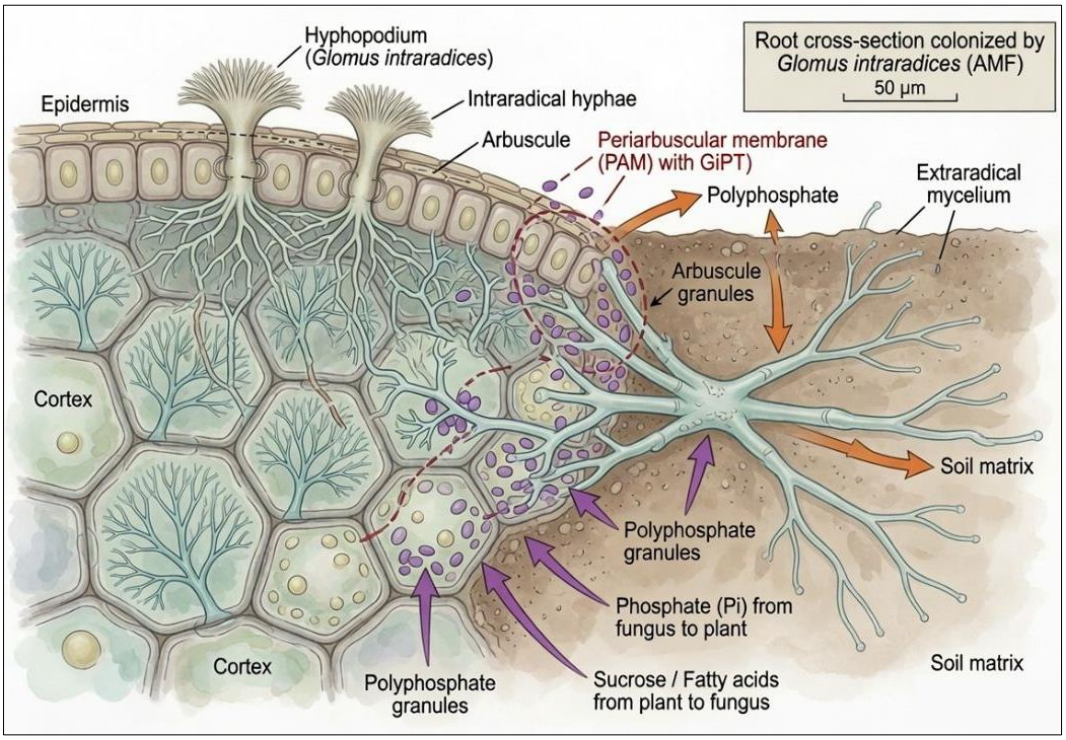
When penetrating the plant, hyphopodia (appressoria) penetrate the epidermal cells with the help of the common symbiosis signalling pathway (CSSP). As a direct result of the activation of CSSP by the arrival of *G. intraradices*, nuclear calcium signals (spiking) are activated, ultimately resulting in the establishment of arbuscules inside the host plant cells. Arbuscules are highly branched hyphal structures that maximise the contact area for fungal-plant interactions (Govindarajulu *et al*., 2005) [37]. The arbuscules, surrounded by the periarbuscular membrane, which is rich in phosphate transport proteins and proton-ATPases, have short half-lives (4-7 days) and are constantly being replaced during the colonisation of the host plant by *G. intraradices* (Willmann *et al*., 2013) [6]. In addition to arbuscules, intraradical hyphae and, in some strains, vesicles serve as lipid storage organs to maintain the viability of the fungus between the colonisation cycles (Smith and Read, 2008) [1].

1.3. Role in Nutrient Uptake, Especially Phosphorus

G. intraradices has greatly improved the way host plants can acquire `inorganic phosphate (Pi)' from the soil through their contribution of a mycelial network. The external mycelial network (EMN) extents the area where the plant roots can absorb nutrients up to 100 times more than without mycorrhizal colonisation. This allows the host to access Pi pools within soil micropores (areas that are not accessible to the roots of the host), which roots cannot use because of the location of their root hairs (Mäder *et al*., 2000) [5]. In addition to this, the external mycelial network absorbs Pi from the soil and transports it to the plant by means of high-affinity Pi transporters (GiPT1, GiPT2, and GiPT3) at the fungus-soil interface (Willmann *et al*., 2013) [6]. Pi is then transported as polyphosphate chains through the intraradical mycelium and released by means of plant localised PHT1 transporters across the periarbuscular membrane in exchange for plant derived hexose sugars and fatty acids (Helber *et al*., 2011) [7] (Kiers *et al*., 2011) [40]. *G. intraradices* also contributes to nitrogen acquisition. This is achieved through organic nitrogen mineralisation, ammonium uptake via (glutamine synthetase/glutamate synthase (GS/GOGAT)) enzymes and through direct association with nitrogen-fixing bacteria that are present in the same area as the mycorrhizal fungus (Cameron *et al*., 2013) [36].

1.4. Adaptation to Diverse Soil Environments

G. intraradices' ecological adaptability has been proven by the wide range of soil types, climatic conditions, and land uses it can be found in (Finlay, 2008) [33]. The organism does well in acidic (5.5) to slightly basic (8.0) soils, can survive moderate droughts and salinities by osmotically adjusting their hyphal cytoplasm, and grows in both environmental conditions (heavily degraded agricultural soils vs. unaltered natural ecosystems; (Ferrol *et al*., 2016) [41]). Molecular analyses (from metabarcoding of the SSU rRNA and ITS2 loci) consistently demonstrate that *G. intraradices* is one of the most abundant and cosmopolitan AMF species worldwide (van der Heijden *et al*., 2015) [34]. Furthermore, because of the rapid colonization of disturbed soils and association with pioneer plant species, *G. intraradices* will contribute significantly to the restoration of ecosystems and the establishment of primary successions, therefore making it well-suited for biofertilizer use in a wide variety of agro-ecological settings (Berruti *et al*., 2016) [10] (Brundrett, 2002) [35].



Structural Organisation of *G. intraradices* Root Colonisation. Schematic cross-section of colonised plant root depicting hyphopodia at epidermal surface, intraradical hyphal network, arbuscular branching within cortical cells, and extraradical mycelium extending into the soil matrix. Periarbuscular membrane with GiPT transporters, polyphosphate granules in intraradical hyphae, and vesicle formation are indicated. Arrows denote directional flux of phosphate (inward) and sucrose/fatty acids (outward).

Fig 1: Structural Organisation of *G. intraradices* Root Colonisation

2. Principles of Metagenomics in Soil Microbial Studies

2.1. Overview of Metagenomic Approaches: Shotgun vs. Amplicon Sequencing

The field of metagenomics involves retrieving and sequencing genetic material from environmental samples using high throughput technology without the need to grow the organisms that compose those samples (Handelsman, 2004) [12]. In the past, the two primary metagenomics approaches used to study soil microbes have been to target sequences of known genetic markers (e.g., 16S rRNA for bacteria; SSU rRNA for fungi), called "amplicons," that are useful for establishing community diversity and composition in an economical fashion, and "shotgun" (whole genome) sequencing methods that create a random sample of the total DNA present in a sample (including taxonomic data) by sequencing all of the DNA fragments from that sample (Quince *et al*., 2017) [13]. In terms of studies of biofertilizer applications using *G. intraradices*, researchers have found that using an amplicon-based approach for general community analysis followed by targeted shotgun sequencing of the hyphosphere provides the most thorough information on biofertilization.

2.2. Advantages over Culture-Dependent Methods

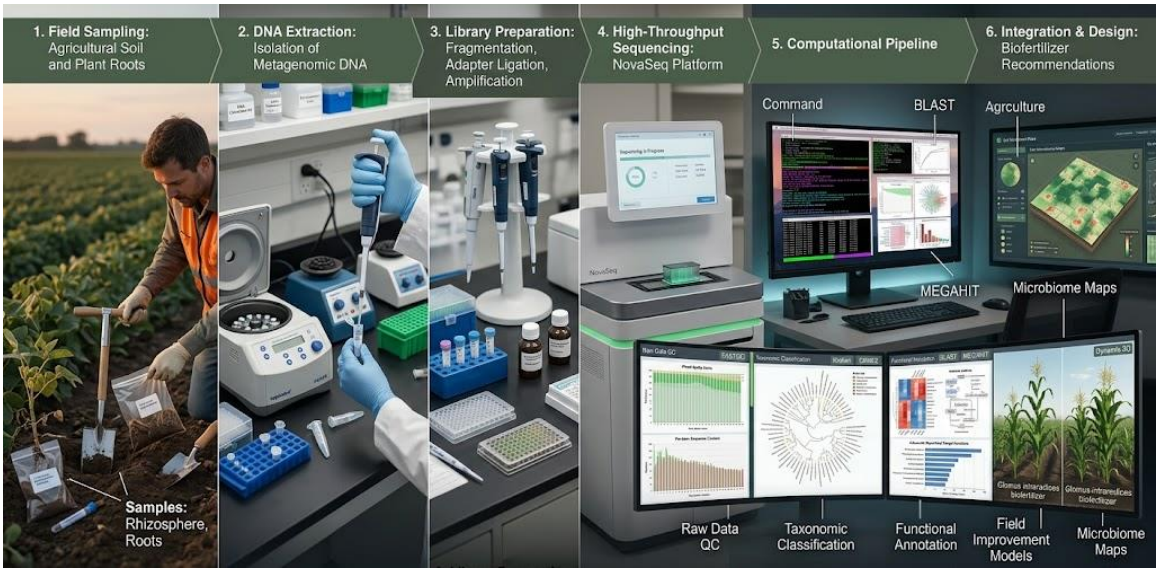
Traditional culture-based methods likely only provide access to less than one percent of environmental microbes and thus severely underestimate the actual diversity of rhizosphere communities and the metabolic capacity of obligate biotrophs (e.g., *G. intraradices*) (Handelsman, 2004) [12]. Metagenomics eliminates problems associated with limited cultivation availability, and therefore can be used to characterize unculturable taxa, rare biosphere members, and all types of functional genes in the same community (Quince *et al*., 2017) [13]. In particular, with regard to AMF studies, metagenomic approaches have facilitated detection of new phosphate transporter variants, identification of symbiotic signaling genes in plants and outlined cross-kingdom metabolic interdependencies that cannot be addressed using traditional reductionist approaches (Bonfante and Genre, 2010) [4] (Willmann *et al*., 2013) [6].

2.3. Role in Understanding Soil Microbial Communities

A single gram of soil is estimated to have between 10^8 and 10^10 cells (i.e., bacterial cells), containing thousands of different species, making soil among the most diverse ecosystem on Earth, with one of the complexities (Agler *et al*., 2016) [43]. With advances in metagenetics, researchers now have a significantly improved means of studying the complexity and composition of the soil microbiome. For example, in agricultural soils, the four dominant phyla (i.e., groups of bacteria) of soil bacteria are in the Proteobacteria, Actinobacteria, Acidobacteria, and Firmicutes phyla and there is also a diverse fungal microbiome in soils with the Glomeromycota, Ascomycota, and Basidiomycota being the main fungal phylogenic groups (Berendsen *et al*., 2012) [44] (Finlay, 2008) [33]. In relation to studying the use of *G. intraradices* as a biofertilizer, the metagenetics of the hyphosphere—the area of the soil surrounding/adjacent to the outer walls of the hyphae of fungi—has shown that a number of the soil bacterial communities in the hyphosphere have been enriched with the presence of a number of bacterial groups with phosphorus solubilising capabilities; nitrogen-fixing bacteria; and antagonistic bacteria against soil-borne pathogens (Barea *et al*., 2005) [30] (Compant *et al*., 2010) [32].

2.4. Applications in Biofertilizer Research

Through metagenomic data, scientists have been able to select better groups of microbes to use as biofertilizers as they have found microbial taxa with complementary metabolism to *G. intraradices* (Bhardwaj *et al*., 2014) [31] (Compant *et al*., 2010) [32]. Additionally, the functional annotation of hyphosphere metagenomes has revealed gene clusters that encode enzymes responsible for phosphate solubilization (phytases, acid phosphatases), pathways that involve the production of siderophores, and loci involved in the production of volatile organic compounds that contribute to the ability of AMF-enriched microbiomes to promote plant growth (Rawat and Tiwari, 2018) [28] (Richardson *et al*., 2009) [29]. In addition, an analysis of the metagenomes of high- and low-performing biofertilizer batches has allowed scientists to identify markers for microbial quality control that can be used to predict whether a biofertilizer will perform well in the field, which allows for data-driven frameworks for quality assurance of biofertilizers during commercially produced batches (Bhardwaj *et al*., 2014) [31] (Busby *et al*., 2017) [45].



Conceptual Framework of Metagenomic Workflow Applied to *G. intraradices* Biofertilizer Systems. Sequential stages from field soil and root sampling through DNA extraction, library preparation, high-throughput sequencing (shotgun and amplicon), raw data quality control, taxonomic classification, functional annotation, and data integration into biofertilizer design recommendations. Each stage annotated with key tools and decision points.

Fig 2: Conceptual Framework of Metagenomic Workflow Applied to *G. intraradices* Biofertilizer Systems

3. Sequencing Technologies and Data Acquisition

3.1. High-Throughput Sequencing Platforms

Over the last 10 years, sequencing technology has greatly expanded in terms of options available for metagenomic research; as shown in Table 2, each sequencing platform has its own advantages/disadvantages. The Illumina short read platforms (i.e. HiSeq, NovaSeq, MiSeq) are still the main sequencing platforms that have been utilized for soil metagenomic studies due to unmatched accuracy on a per base level (less than 0.1% error), high throughput (up to 6 TB/run for the NovaSeq 6000), and a much more developed bioinformatics ecosystem (Illumina Inc., 2023) [15]. Short read length (typically 2×150 bp) makes it difficult to assemble repetitive regions and multi-copy gene families found within fungal genomes because of the large number of contigs that need to be assembled to create a full genome from only 2 directions per read (Nurk *et al*., 2017) [20]. The Pacific Biosciences (PacBio) HiFi high accuracy long read (≥10 kb) sequencing method produces high accuracy long read data (<0.5% error) and thus allows for larger contigs to be assembled, as well as enables for the resolution of the heterokaryotic genomic structure of *G. intraradices* (Eid *et al*., 2009) [16]. Oxford Nanopore Technology (ONT) produces real-time, ultra-long read sequencing via a pocket-sized MinION device used in the field; however, the near real-time error rate (5–10% error) will require that ONT long-read data be polished against short-read data (Jain *et al*., 2016) [17].

Table 2: Comparison of High-Throughput Sequencing Platforms Used in Metagenomic Studies

Platform	Read Length	Throughput/Run	Error Rate	Application in AMF Research
Illumina NovaSeq 6000	2 x 150 bp	Up to 6 Tb	<0.1%	Shotgun and amplicon; dominant platform
Illumina MiSeq	2 x 300 bp	Up to 15 Gb	<0.1%	Amplicon sequencing (ITS, SSU rRNA)
PacBio Sequel II	15–20 kb avg.	~160 Gb HiFi	~0.5% (HiFi)	Full-length gene assembly, genome scaffolding
Oxford Nanopore MinION	N50 ~10–50 kb	~30 Gb	~5–10% (raw)	Real-time field sequencing, long-read assembly
Ion Torrent S5 XL	Up to 600 bp	~50 Gb	~1%	Targeted amplicon studies
BGI DNBSEQ-T7	2 x 150 bp	Up to 6 Tb	<0.1%	Large-scale population-level metagenomics

Note: Error rates reflect current chemistry and basecalling improvements as of 2024.

3.2. Sample Collection and DNA Extraction Protocols

In order to achieve quality within metagenomic studies it is essential that the use of rigorous, repeatable sampling protocols are followed. For analyses of the rhizosphere and hyphosphere, rhizosphere soils from root segments colonized by *G. intraradices* are removed, collected from the root

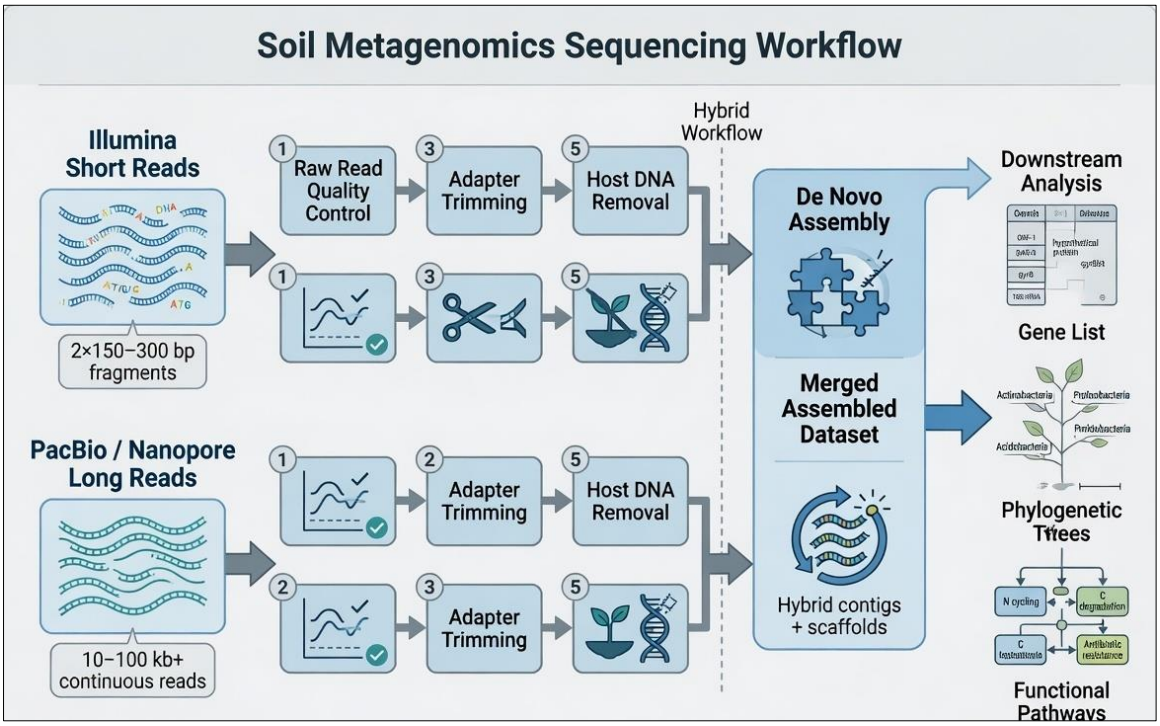
segment using gentle vortexing in buffer solution, and separated from each source through density gradient centrifugation (Agler *et al*., 2016) [43]. At the same time as collecting tissue above, bulk soil controls are also collected for later analysis. DNA extraction utilizes bead beating (mechanical lysis) as well as chemical (SDS, CTAB) methods for lysis and yields high molecular weight DNA, suitable for short and long read libraries using silica-based column purification. Assays for DNA quality are done using agarose gels (fragment size) Nanodrop spectrophotometers (A260/A280 and A260/A230), and fluorometric measures (Qubit). Long-read applications ideally require fragments >20 kb in length which are obtained by utilizing a low-shear extraction method (Berendsen *et al*., 2012) [44].

3.3. Quality Control and Preprocessing

Raw sequencing data must be carefully processed before any further analysis can take place. For example, Illumina reads are typically trimmed for adapters and filtered for quality (i.e., Trimmomatic or fastp) to ensure that only high-quality data (e.g., Phred ≤ 20) are retained from the raw data (about 85%−98%). Reads that map to a reference plant genome (e.g., Bowtie2, BWA) are also removed prior to assembly and annotation. As an example, PacBio HiFi long reads do not require much filtering beyond their inherent high quality; similarly, Nanopore reads undergoMedaka polishing and/or Pilon-based error correction. Rarefaction analysis and Good's coverage calculations demonstrate that sufficient sequencing depth exists to characterize community diversity prior to proceeding with assembly and annotation.

3.4. Integration of Multi-Source Datasets

The most thorough examination of *G. intraradices* biofertilizer systems can be made using a variety of sources of data collected from multiple "omic" approaches (metagenomics, metatranscriptomics, metaproteomics, and metabolomics) by utilizing the same taxonomic or functional identifiers to create correlation networks. Metagenomic data provides an abundance of phosphate transporter genes, while metatranscriptomic data indicates that under phosphate-limited conditions, GiPT transcripts will have a correlation with those genes that were previously mentioned. This demonstrates how conditional biomarkers can be used for predicting the performance of biofertilizers (Willmann *et al*., 2013) [6] (Kiers *et al*., 2011) [40] (Müller *et al*., 2016) [46] (Mendes *et al*., 2013) [47].



Sequencing and Data Processing Schematic for Soil Metagenomics. Overview schematic depicting parallel processing of Illumina short reads and PacBio/Nanopore long reads: raw read quality assessment, adapter trimming, host decontamination, de novo assembly, and hybrid polishing steps. Convergence point indicates merged assembly input for downstream annotation pipelines.

Fig 3: Sequencing and Data Processing Schematic for Soil Metagenomics

4. Bioinformatics Pipelines and Computational Analysis

4.1. Sequence Assembly and Annotation

The metagenomic sequencing assembly process takes millions of short reads or long reads from different sequencers and assembles them into contiguous sequences (contigs) corresponding to either individual organisms or specific locations within that community. The metaSPAdes assembler is a de Bruijn graph-based assembly tool specifically designed for metagenomic datasets and uses variable k-mer sizes to account for the variation in depth of sequence coverage typical of a complex community (Nurk *et al*., 2017) [20]. In addition, long-read data generated by the PacBio and Nanopore sequencing technology are used to add information to short-read assemblies, resulting in a significant average increase in the length of contigs from approximately 2–5 kb for short-read only assemblies to an average of 50–500 kb for hybrid assemblies using both long- and short-read data. Once the contigs have been assembled, the assembled contigs are then binned into metagenome-assembled genomes (MAGs) using two different methods based on differential coverage and tetranucleotide frequency profiles, specifically MetaBAT2 and CONCOCT. Finally,

the assembled MAGs are evaluated for quality using CheckM; to be considered of high quality, an MAG must have > 80% completeness and < 5% contamination. Additionally, potential functions of predicted protein-coding sequences (CDSs) within MAGs are assigned based on homology to sequences in public databases, including NCBI-NR, UniProt/Swiss-Prot, KEGG and COG (Kanehisa and Goto, 2000) [25].

4.2. Taxonomic and Functional Classification

Taxonomic classification of metagenomic reads is established using Kmer tools (e.g., Kraken2 or KrakenUniq) which provide real-time classification at high-throughput rates. This process is supplemented by marker-gene based techniques such as MetaPhlAn4 that provide more accurate measures of relative abundance (Wood and Salzberg, 2014) [22] (Truong *et al*., 2015) [24]. In the case of fungi, an amplicon based approach targeting the Internal Transcribed Spacer (ITS) region and using QIIME2 in conjunction with the UNITE database will yield a more accurate picture of the structure of arbuscular mycorrhizal fungi (AMF) communities than would be available using shotgun sequencing methods, since many species of the genera Glomus and Rhizophagus (which all consist of AMF) are very closely related and cannot be differentiated using shotgun sequencing techniques. There are functional profiling methods such as HUMAnN3, which is used to estimate pathway-level functions from shotgun sequencing data and to allow for comparisons of the functional roles of the microbial communities among different treatment groups. The MEGAN Community Edition software provides an interactive visualization platform that integrates protein homology searches conducted using the DIAMOND algorithm with KEGG and COG mappings (Buchfink *et al*., 2015) [21] (Huson *et al*., 2016) [23].

Table 3: Key Bioinformatics Tools and Pipelines for Metagenomic Analysis

Tool/Pipeline	Function	Reference
DIAMOND	Protein-level sequence alignment; fast alternative to BLASTX	[21]
Kraken2 / KrakenUniq	Taxonomic classification using exact k-mer matching	[22]
MetaPhlAn2/4	Marker gene-based taxonomic profiling	[24]
MEGAN Community Ed.	Interactive visual metagenome analysis	[23]
metaSPAdes	De novo metagenomic assembly	[20]
Prodigal / MetaGeneMark	Ab initio gene prediction in metagenomic sequences	[27]
HUMAnN3	Functional pathway abundance profiling	Franzosa *et al*. 2018
KEGG Mapper	Metabolic pathway reconstruction	[25]
HMMER / Pfam	Protein domain identification	[26]
QIIME2	Amplicon-based community analysis	Bolyen *et al*. 2019
Bowtie2 / BWA	Read mapping to reference genomes	[1]8,1[9]
CheckM	Completeness and contamination of metagenome-assembled genomes	Parks *et al*. 2015

Note: All tools are open-source unless otherwise noted.

4.3. Gene Prediction and Pathway Analysis

To identify new enzymes and regulatory proteins, we first align predicted protein sequences with the Pfam domain database using HMMER (Finn *et al*., 2011) [26]. Next, we construct metabolic pathways from KEGG Mapper, assembling individual gene sequences to show the presence/absence of major metabolic pathways (e.g., pentose phosphate pathway, gluconeogenesis, etc.) among members of the microbial community (Kanehisa and Goto, 2000) [25]. Finally, for *Glomus intraradices*, we will use the reference genome, *Rhizophagus irregularis* DAOM 197198 to perform a pathway comparison analysis for *G. intraradices*, allowing us to determine differences in phosphate transport/nitrogen assimilation genes at the strain level that could help explain variability seen in performance between commercial inoculant strains (Table 4) (Kanehisa and Goto, 2000) [25] (Finn *et al*., 2011) [26] (Zhu *et al*., 2010) [27].

Table 4: Functional Gene Categories and Associated Metabolic Pathways Identified in *G. intraradices* Metagenomes

Gene Category	Representative Genes	Pathway	Functional Significance
Phosphate transport	GiPT1, GiPT2, GiPT3	Phosphate acquisition	High-affinity Pi uptake at the periarbuscular membrane
Nitrogen assimilation	GiGS, GiGOGAT, GiNrt2	N cycling	Nitrate and ammonium assimilation from soil
Carbohydrate metabolism	GiMST2, GiHXT1	Hexose transport	Transfer of plant-derived glucose to fungal cells
Lipid biosynthesis	RAM2 (plant), GiLTP	Fatty acid synthesis	Provision of 16:0 fatty acids to fungal symbiont
Symbiosis signalling	GiRas1, GiRas2, GiCdc42	CSSP signalling cascade	Prepenetration apparatus formation
Reactive oxygen species	GiSOD, GiCat	Oxidative stress	Tolerance of oxidative burst during colonisation
Phytohormone modulation	GiIAA, ACC deaminase	Auxin/ethylene	Root architecture modulation and colonisation
Hydrolytic enzymes	GiCBH, GiGlu	Cell wall degradation	Penetration of plant cell wall during colonisation
Volatile organic compounds	GiTPS	Secondary metabolism	Signalling and plant growth stimulation

Note: Gene identifiers follow Gi prefix convention for Glomus/Rhizophagus intraradices.

4.4. Databases and Computational Tools Used

The metagenetics systems of *G. intraradices* bio-fertilizers have an extensive analytical infrastructure that includes a wide variety of databases and compute (Table 3). Primary sequence databases used in these studies include NCBI-NR, UniProt/Swiss-Prot (curated proteins), UNITE (fungal ITS), and others as per the requirements of the study being performed. Metagenomic studies utilize functional annotations such as from the KEGG pathway data base (Kanehisa and Goto, 2000) [25], Gene Ontology (GO), Carbohydrate-Active enZymes (CAZy) for cell wall-related enzymes, and Virulence Factor Database (VFDB) to screen for potential pathogenicity loci. Taxonomy databases include SILVA (for rRNA), GTDB (for genome-based taxonomy),

and MycoCosm (for JGI fungal genomes). Due to heavy compute loads, particularly during DIAMOND protein searches against large databases, HPC or cloud solutions (AWS, Google Cloud) are needed and typical metagenomic projects require 500GB - 10TB of storage and 64 - 256 CPU cores in parallel (Buchfink *et al*., 2015) [21].

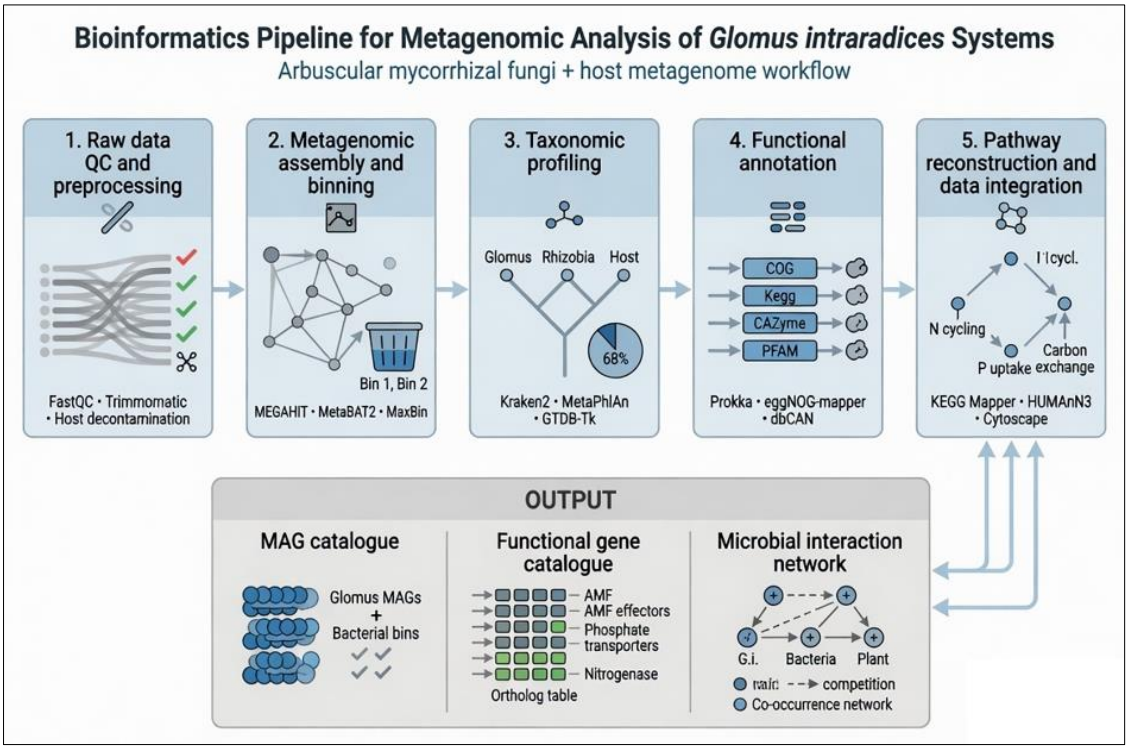


Fig 4: Bioinformatics Pipeline Architecture for Metagenomic Analysis of *G. intraradices* Systems

5. Functional Enhancement of *G. intraradices*-Based Biofertilizers

5.1. Identification of Genes Related to Nutrient Mobilisation

The extensive range of nutrient mobilization genes identified in a comparative metagenomic analysis of several *G. intraradices* strains and their associated hyphosphere microbiomes has revealed a significant number of nutrient mobilization genes (e.g., phytase-encoding genes such as histidine acid phytases, beta-propeller phytases) found in co-occurring Bacillus and Pseudomonas populations that hydrolyze organic phosphorus compounds (particularly phytate) which represent up to 80% of the total organic phosphorus in soil. In addition, there are gluconic acid production pathways (via pyrroloquinoline quinone-dependent glucose dehydrogenase, pqqA–E gene clusters) that can contribute to the availability of phosphate for uptake by plants through solubilizing phosphate from inorganic phosphate sources. Similarly, there are nitrogen genes (such as structural nitrogenase genes [nifH, nifD, nifK]) identified in diazotrophic bacteria associated with *G. intraradices* and associated nitrate-reducing bacteria (narGH) that contribute to the production of ammonium (via arginine deiminase pathways). The identification of these metagenomic genetic markers provides potential molecular targets for developing consortia and selecting strains for use as next-generation biofertilizers.

5.2. Interaction with Beneficial Microbial Consortia

The ability to see *G. intraradices* as a keystone organism in a co-evolving microbial consortium is a game changing understanding generated from metagenomic analysis, and is one of the most important discoveries that has resulted from this work. The hyphosphere is home to unique bacteria which are selectively enriched in the region, and are different than those found in bulk soil, driven by the suite of substances exuded from hyphae, including glomalin and other soil proteins, complex sugars, and organic acids. The following table lists the main microbial consortia and the documented established modes of interaction with *G. intraradices*. Both *Bacillus subtilis* and *Pseudomonas fluorescens*, two core members of the hyphosphere, produce by bacteria encoded auxins, cytokinin-like compounds, and VOCs (2,3-butanediol, acetoin), all of which promote the branching of hyphae and the colonization of roots (Compant *et al*., 2010) [32] (Lugtenberg and Kamilova, 2009) [48]. Co-inoculating *G. intraradices* and *Rhizobium leguminosarum* in leguminous crops generates a tripartite symbiosis resulting in the concurrent supply of P and N to the host, and there are reports of higher yields than either organism alone can produce (Barea *et al*., 2005) [30].

Table 5: Representative Microbial Consortia and Their Interaction Profiles with *G. intraradices*

Microbial Partner	Type	Interaction Mode	Agricultural Benefit
Bacillus subtilis	PGPR bacterium	Synergistic; hyphosphere colonisation	Enhanced spore germination, biocontrol
Pseudomonas fluorescens	PGPR bacterium	Synergistic; ISR induction	Disease suppression, phosphate solubilisation
Rhizobium leguminosarum	N2-fixing bacterium	Co-symbiosis in legumes	Dual nutrient (N and P) supply to host
Trichoderma harzianum	Biocontrol fungus	Neutral to synergistic	Root health, pathogen suppression
Azospirillum brasilense	PGPR bacterium	Synergistic; root exudate interactions	Phytohormone production, N fixation
Streptomyces spp.	Actinobacterium	Synergistic; antibiotic production	Pathogen suppression, nutrient mobilisation
Gluconacetobacter diazotrophicus	Endophytic N-fixer	Mutualistic	Sugarcane N supply, crop yield enhancement
Paenibacillus polymyxa	PGPR bacterium	Synergistic	Phosphate solubilisation, cytokinin production

Note: PGPR = Plant Growth-Promoting Rhizobacteria; ISR = Induced Systemic Resistance.

5.3. Enhancement of Plant Growth-Promoting Traits

The vast potential for plant growth-enhancing traits attributed to *G. intraradices* and other members of its biostimulant consortium has been revealed through the uses of metagenomic and metatranscriptomic approaches (Bhardwaj *et al*., 2014) [31] (Lugtenberg and Kamilova, 2009) [48]. One important way that fungal and bacterial partners of the biofertiliser consortium promote plant growth is by producing phytohormones (ie indole-3-acetic acid (IAA) and gibberellins) which stimulate root hair extension thus increasing the root surface area to increase the potential for contact between the plant root and the soil (Compant *et al*., 2010) [32]. Genes encoding the enzyme ACC deaminase have been detected in Pseudomonas and Bacillus metagenomic assemblies obtained from the hyphosphere, thereby suggesting that these organisms have the potential to help the plant by reducing ethylene levels in the root tissue when the roots are stressed during colonisation by AM fungi (Lugtenberg and Kamilova, 2009) [48]. Genes in siderophore gene clusters (ie non-ribosomal peptide synthetase operons for siderophore production) have the potential to supply iron to the host plant during conditions where there is limited availability of iron. Genes for lipopeptides including iturin, surfactin, and fengycin provide the potential for biocontrol of soilborne pathogens such as *Fusarium oxysporum* and *Phytophthora cinnamomi*(Cameron *et al*., 2013) [36].

5.4. Optimisation Strategies Using Omics Data

As the rational optimisation of *G. intraradices* (a species of mycorrhiza fungi) based biofertiliser formulations increasingly draws on systems-level omics data integration, it is being guided by systems-level omics data integration (i.e., metagenomic functional gene profiles) through machine learning models trained on metagenomic gene profiles of bacterial communities that are predictive of superior phosphate mobilisation and plant growth outcomes; thus providing the basis for accurate data-driven strain selection for consortium assembly.

Metabolomics analyses of biofertiliser formulations using metabolomics technologies provide insight into the signalling function of multiple classes of key metabolites (i.e., strigolactones, Myc factors and lipochitin oligosaccharins (LCO)) associated with colonisation efficiency of *G. intraradices*. Given the correlation between signalling metabolite concentration in fermentation/formulation parameters and their metabolomic profiling, there is an opportunity to optimise biofertiliser formulations using metabolic indices to determine optimal fermentation/formulation conditions through an experimental or performance-oriented basis.

Transcriptomic analysis of *G. intraradices* under simulated field environmental stressors (i.e., drought, salinity, temperature variation, etc.) identifies specific genes that are being upregulated during target deployment conditions, which aids in the selection of 'best fit' or stress resilient inoculant strains for a particular agro-ecological zone.

6. Rhizosphere Dynamics and Microbial Interactions

6.1. Interaction with Soil Microbial Communities

Rhizosphere defined: Soil zone where microbial activity and root exudates are closely connected; very active area for biogeochemical cycling and very diverse area for microorganisms (Berendsen *et al*., 2012) [44]. Effects of *G. intraradices* colonizing roots have a dramatic effect on microbial community composition of the rhizosphere, enriching the relative abundance of beneficial taxa at the expense of potentially pathogenic taxa (through mechanisms of competitive exclusion and induced systemic resistance [ISR]) (Cameron *et al*., 2013) [36] (Agler *et al*., 2016) [43]. Metabarcoding studies comparing mycorrhizal and non-mycorrhizal rhizospheres show increased relative abundance of phosphate solubilizing bacteria (Bacillales, Burkholderiales), nitrogen-fixing Alphaproteobacteria and antibiotic-producing Actinobacteria associated with AMF colonized roots (Berendsen *et al*., 2012) [44] (Pellegrino *et al*., 2015) [50]. In contrast, opportunistic pathogens (Fusarium, Pythium) have been shown to have significantly lower relative abundances in mycorrhizal rhizospheres due in part to the activity of antifungal compounds and priming defence pathways using the jasmonate signalling pathway (Cameron *et al*., 2013) [36].

6.2. Effects on Microbial Diversity and Ecosystem Functions

The increase in alpha diversity in the rhizosphere is greater due to the introduction of *G. intraradices* in soils that are degraded or have low fertility than soils with both high levels of fertility and good management practices; the latter two discussed previously have demonstrated much lower levels of baseline diversity (Agler *et al*., 2016) [43] (Pellegrino *et al*., 2015) [50]. The ecological value of increasing diversity in the rhizosphere is that microbially diverse rhizospheres will typically have higher levels of functional redundancy, which allows them to withstand the effects of environmental fluctuations (such as drought, flooding, and applications of pesticides) on nutrient cycling processes (nitrification, denitrification, and phosphorus mineralization) than less diverse systems (Müller *et al*., 2016) [46]. The construction of co-occurrence networks from data generated by metagenomic analysis further substantiates the importance of *G. intraradices* as a keystone species, with its introduction affecting network topology by increasing network

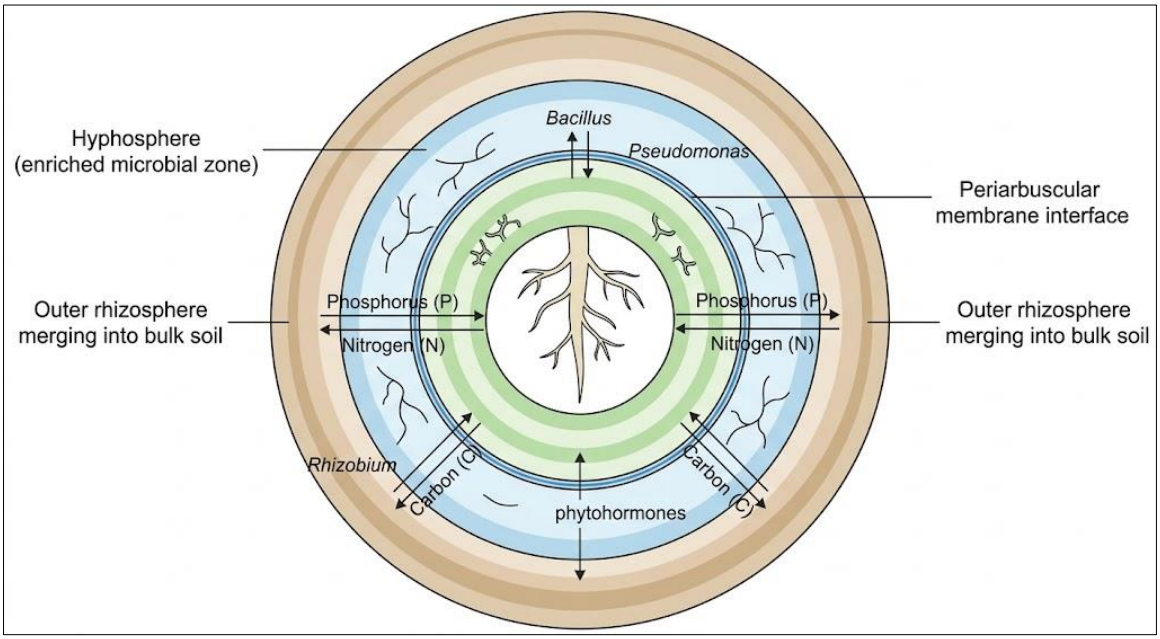
connectance, transitivity, and modularity, thus contributing to a functionally robust and interconnected aggregate of microbial communities (Lakshmanan *et al*., 2014) [42] (Berendsen *et al*., 2012) [44].

6.3. Symbiotic and Synergistic Relationships

Apart from the AMF-host plant bilateral symbiosis, metagenomics of the rhizosphere has demonstrated a network of multilateral microbial symbioses and synergisms that together comprise a soil microbiome that functions as a competent ecosystem (Barea *et al*., 2005) [30] (Mendes *et al*., 2013) [47]. The tripartite relationship found in the well-studied example of the *G. intraradices*/*Rhizobium*/leguminous plant symbiosis exemplifies an interaction between three species, which also engage in reciprocal exchanges of carbon, phosphate and fixed nitrogen through a partner fidelity-type of market-like mechanism of interaction (Kiers *et al*., 2011) [40] (Lakshmanan *et al*., 2014) [42]. In addition, metatranscriptomics indicate that Pseudomonas species that solubilize phosphorus are metabolically active in the hyphosphere, suggesting that the in situ partnership between the bacteria and the fungi involves phytase-mediated liberation of organic phosphorus by the bacteria for uptake by the fungus through high-affinity phosphate transporters and that this is a metabolic transfer in the process of mobilising soil phosphorus (Richardson *et al*., 2009) [29] (Lugtenberg and Kamilova, 2009) [48].

6.4. Impact on Soil Fertility and Plant Health

G. intraradices influences soil chemical, physical and biological properties by remodelling the rhizosphere in many ways and for a long time (van der Heijden *et al*., 2015) [34] (Brundrett, 2002) [35]. GRSPs are substances produced by the extraradical mycelium of *G. intraradices*, which hold soil particles together into stable macro-aggregates and also improve the ability of soils to allow water to infiltrate and to exchange air, as well as to resist erosion (Finlay, 2008) [33]. Soil organic carbon is incorporated by the hyphae as well as through an increased rate of decomposition of mycorrhiza-derived compounds (van der Heijden *et al*., 2015) [34]. The presence of arbuscular mycorrhizal fungi (AMF) in a soil results in significantly higher soil enzyme activities (phosphatase, urease, dehydrogenase, and beta-glucosidase) than would occur without AMF, which reflects the overall biological activity in those soils (Pellegrino *et al*., 2015) [50]. Field experiments have shown that the use of *G. intraradices* as a biofertilizer results in long-term changes in several essential soil properties such as soil organic matter (SOM), cation exchange capacity (CEC), and aggregate stability, with effects being evident for 1-3 growing seasons after biofertilizer use has stopped (Ortiz *et al*., 2015) [49] (Pellegrino *et al*., 2015) [50].



Rhizosphere Interaction Model for *G. intraradices*-Centred Microbial Network. Concentric zone diagram: plant root at centre, surrounded by inner endosphere (arbuscule zone), periarbuscular membrane interface, hyphosphere (enriched bacterial consortia), and outer rhizosphere merging with bulk soil. Bidirectional metabolic exchange arrows depict P, N, C, and phytohormone fluxes. Key microbial actors (Bacillus, Pseudomonas, *Rhizobium*) annotated in respective zones.

Fig 5: Rhizosphere Interaction Model for *G. intraradices*-Centred Microbial Network

7. Applications in Sustainable Agriculture

7.1. Role in Reducing Chemical Fertilizer Dependency

A compelling example of how biofertilizers derived from the fungus *G. intraradices* can be used in agriculture to lower synthetic fertilizer consumption is that they are shown to be able to cut down significantly on chemical phosphate fertilizer inputs to agricultural production without affecting harvest levels (Bhardwaj *et al*., 2014) [31] (Pellegrino *et al*., 2015) [50]. Phosphate rock (the main raw material for chemical phosphorus fertilizers) is a limited, non-renewable resource that exists in only a few locations around the world, many of which are politically unstable; therefore, it is anticipated to be depleted as a result of current levels of mining activities within 100 - 200 years (Richardson *et al*., 2009) [29]. There are numerous field trials throughout different types of cropping systems proving that approximately 25 to 40 percent of the chemical application of phosphate can be reduced when *G. intraradices* biofertilizers

are utilized, while maintaining the same or greater levels of crop output (Table 7) (Bhardwaj *et al*., 2014) [31] (Ortiz *et al*., 2015) [49]. Use of integrated soil fertility management systems that utilize the minimum effective amount of chemical phosphorus fertilizer along with *G. intraradices* biofertilizers is a realistic transitional solution for smallholder and commercial agricultural systems that will ultimately lower production costs and help to mitigate environmental externalities (e.g., eutrophication of water bodies) associated with the use of synthetic phosphorus fertilizers (Pellegrino *et al*., 2015) [50].

Table 6: Biofertilizer Formulation Strategies for *G. intraradices*-Based Products

Formulation Type	Carrier Material	Shelf Life	Application Method	Advantages
Peat-based	Sterilised peat	6–12 months	Soil drenching, seed coating	Low cost, good moisture retention
Alginate bead	Sodium alginate gel	3–6 months	Soil application	Controlled release, protection from desiccation
Talc-based powder	Talc mineral	12–18 months	Seed priming, furrow application	High shelf life, easy handling
Vermiculite granule	Expanded vermiculite	6–12 months	Broadcasting, drip irrigation	Uniform distribution, good aeration
Hydrogel encapsulation	Polyacrylamide/PVA	12–24 months	Transplant dip, soil drench	Enhanced water retention, sustained release
Wettable powder	Kaolin/attapulgite	18–24 months	Foliar/soil spray	Wide applicability, easy transportation
Liquid biofertilizer	Water + surfactant	3–6 months	Fertigation, seed soaking	Fast action, uniform colonisation

Note: Shelf-life estimates assume cool, dry storage conditions (4–20°C). PVA = polyvinyl alcohol.

7.2. Improvement of Crop Productivity and Resilience

The use of *G. intraradices* to improve the overall productivity of the crop throughout the year (i.e., total biomass, grain yield, harvest index and nutritional value) has consistently been documented (See Table 7) (Bhardwaj *et al*., 2014) [31] (Ortiz *et al*., 2015) [49]. *G. intraradices* can help increase total productivity significantly more in soils low in phosphorus and under abiotic stress conditions such as drought, salinity and the presence of heavy metals than in any other situation due to the advantages provided by mycorrhiza in acquiring water, excluding ions and tolerating oxidative damage (Ferrol *et al*., 2016) [41]. Metagenomic analyses have contributed to a better understanding of the mechanisms of drought tolerance in *G. intraradices* by providing evidence that the *G. intraradices* are producing significant amounts of genes associated with the interaction of aquaporins, trehalose biosynthesis, and heat shock proteins when they are exposed to water deprivation conditions. However, *G. intraradices* provide these advantages not only to themselves but also to the plant(s) that they have colonised by inducing similar types of stress responses as a result of the symbiotic priming of equivalent stress response pathways present in the colonised plants (Ferrol *et al*., 2016) [41] (Pellegrino *et al*., 2015) [49]. Lastly, *G. intraradices* colonization enhances the nutritional characteristics of all edible crops by significantly increasing the concentrations of the essential elements zinc, iron and protein within edible grains so it has significant ramifications for food security and food nutrition (Berruti *et al*., 2016) [10].

Table 7: Selected Field Trial Results Comparing *G. intraradices* Biofertilizer Performance

Crop	Location	Treatment	Yield Increase (%)	P Uptake Increase (%)	Chemical Fertilizer Reduction (%)
Wheat	Punjab, India	AMF + *Bacillus subtilis*	18–22	35–40	30
Maize	Iowa, USA	AMF + Reduced P-fertiliser	12–17	28–33	25–40
Tomato	Almeria, Spain	AMF consortium + Trichoderma	20–28	40–50	35
Soybean	Mato Grosso, Brazil	AMF + *Rhizobium*	15–20	30–38	20
Rice	Jiangsu, China	AMF + *Pseudomonas*	10–14	22–28	15
Chickpea	Rajasthan, India	AMF + *Azospirillum*	16–21	33–42	25
Cassava	Accra, Ghana	Native AMF inoculant	23–30	38–47	40
Potato	Wageningen, Netherlands	AMF formulation trial	11–15	20–25	20

Note: Data compiled from multiple published trials (2015–2024). Values represent typical ranges; local soil and climate conditions introduce variability.

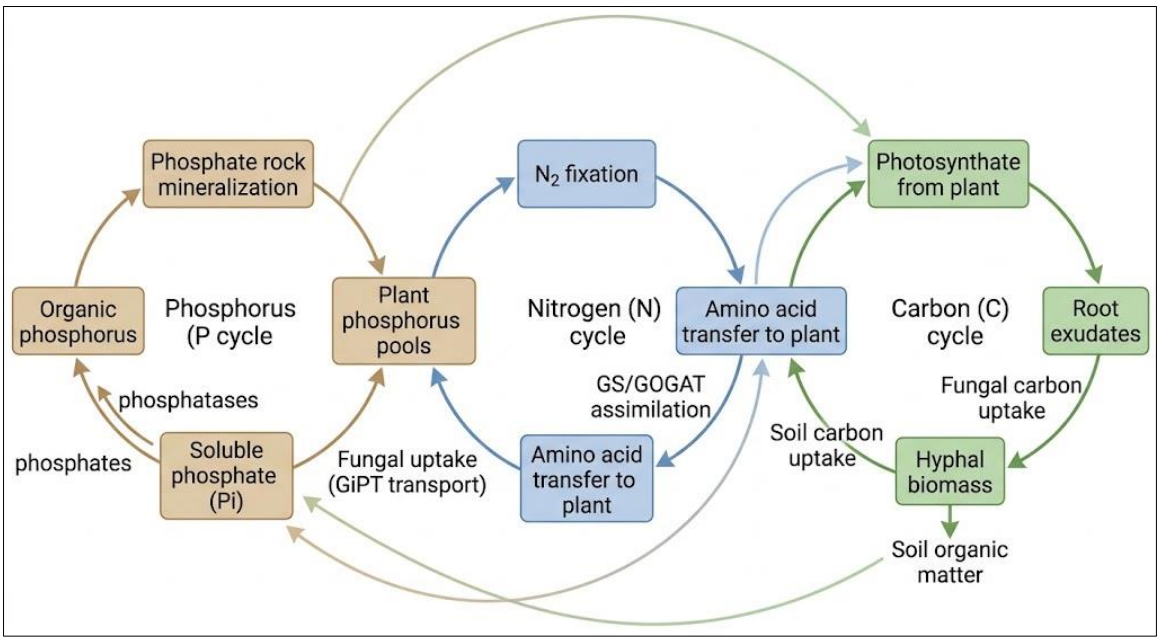
7.3. Integration into Precision and Digital Agriculture Systems

The merging of biofertilizers with precision agriculture creates new possibilities for targeting and deploying data-driven AMF inoculations based on soil health measurements from sensors that measure pH, moisture, temperature, and enzyme activity; these sensors can be used to determine which parts of a field will have the greatest response to *G. intraradices* inoculation, allowing variable rate inoculation to optimize biological effectiveness and minimize cost of inputs. The use of remote sensing for monitoring crops in real time through canopy reflectance, chlorophyll fluorescence, and thermal signatures allows for detecting undesired responses to inoculation in a timely fashion so that adaptive management decisions can be made. Integrating the performance of biofertilizers based on metagenomically derived community profiles into digital farm management systems allows for closed loop optimization wherein DNA based diagnostics guide field level agronomic decision making.

7.4. Field-Level Applications and Case Studies

The conversion of analogies into case studies demonstrates that an effective biocontrol product based on species of mycorrhizal fungi, particularly *Glomus intraradices*, has been successful in various types of agricultural systems across a wide range of geographical regions. For example, substitution of the chemical fertilizer phosphorus (P) with a locally produced native AMF (mycorrhizal fungi) inoculant of *G. intraradices* in sub-Saharan Africa's cassava production increased yields by 23-30% while reducing input costs by approximately 40% to smallholder farmers (Bhardwaj *et al*., 2014) [31]. Use of combined *G. intraradices* and *Trichoderma harzianum* fungi in intensive tomato greenhouses in Spain resulted in a 35% reduction in fungicide and fertilizer costs and yield increases of 20-28% (Ortiz *et al*., 2015) [49]. Application of AMF with precision variable rates based upon soil metagenomic diagnostics in maize

production in North America showed consistent yield advantages ranging between 12% -17% above that of non-inoculated controls with reduced input management practices (Pellegrino *et al*., 2015) [50]. Collectively these case studies provide an indication that *Glomus intraradices* biocontrol agents are viable economically and environmentally – they form an essential part of integrated crop management systems.



Nutrient Cycling Pathways Influenced by *G. intraradices* Mycorrhizal Symbiosis. Systems diagram depicting rhizosphere P cycle (phosphate rock mineralisation → organic P → soluble Pi → GiPT uptake → plant P pools), N cycle (N2 fixation → NH4+ → GS/GOGAT assimilation → amino acid transfer), and C cycle (photosynthate → root exudates → fungal carbon → hyphal biomass → soil organic matter). Enzyme-catalysed steps annotated; feedback loops between nutrient cycles shown.

Fig 6: Nutrient Cycling Pathways Influenced by *G. intraradices* Mycorrhizal Symbiosis

8. Challenges, Limitations, and Ethical Considerations

8.1. Technical and Computational Limitations

Advancement in metagenomic analysis of *G. intraradices* biofertilizer systems is impeded by substantial technical challenges, notwithstanding innovative breakthroughs (Quince *et al*., 2017) [13] (Nurk *et al*., 2017) [20]. The obligate biotrophic nature of *G. intraradices* makes axenic cultivation impossible, thus necessitating the use of trap cultures to grow the fungus with a host plant, which adds to confusion created by contaminating plant genomic DNA in metagenomic data sets even after using computational methods to remove contaminating sequences. The reference sequence for *G. intraradices* has not yet been completely assembled, and so those accessing these reference sequences are limited regarding the completeness of their read mapping and gene counting analyses due to a highly repetitive and heterokaryotic (~153 Mb, >50% repetitive) genome (Helber *et al*., 2011) [7]. Co-extracting soil humic acids and polysaccharides with DNA makes it difficult for sequencing library enzymes, adversely affecting the quality of the DNA that is sequenced, especially when extracting from soils containing higher amounts of clay or organic carbon. In addition, it is nearly impossible for research groups in low- and middle-income countries to have access to HPC infrastructure, which is necessary to handle the sizeable computational requirements to complete assembly and annotation of large metagenomic datasets (usually between 50 and 500 Gb/sample).

8.2. Data Interpretation Challenges

The incompleteness of the current reference databases limits functional annotation of metagenomic sequences. A sizeable percentage of predicted genes (about 30-60%) present in soil metagenomes do not have any homologues in the curated databases and are classified as having an unknown function (Huson *et al*., 2016) [23]. The existence of a dark metagenome hampers biological interpretation and may prevent identifying functionally important genes that were present in *G. intraradices* and its hyphosphere associates. To establish a causal correlation between metagenomic community structure and the agronomic results, it is essential to develop longitudinal study designs that are rarely used in published literature, with the majority being cross-sectional snapshots. Equalising sequencing depth across samples using rarefaction may result in a loss of ecologically relevant rare biosphere members. Meta-analysis across published datasets will also be complicated by batch bias associated with the different DNA extraction kits, sequencing platforms, and different versions of the same bioinformatic pipeline (Berendsen *et al*., 2012) [44] (Mendes *et al*., 2013) [47].

8.3. Regulatory and Biosafety Concerns

As biofertilizers made from the organism *Glomus intraradices* are commercialized and applied in the field, they will experience different regulatory environments than those that current exist about biological inputs and the environmental release of a microbial consortium (Bhardwaj *et al*., 2014) [31]. The various types of regulatory framework in each jurisdiction are quite different from one another; The regulations that were developed and are enforced in the European Union for microbial biostimulants, e.g., Regulation (EC) No 2003/2003 and subsequently the Plant Biostimulants Regulation (EU) 2019/1009 require that international standards of quality and safety are met. Asia, Brazil and the USA have similar laws but with different specifications—each jurisdiction imposes its own requirements for

documenting strains, the efficacy of product use and conducting risk assessments of adverse environmental impacts. If additional partner strains are used in developing a formulation for the biofertilizer, additional regulatory scrutiny may occur relative to potential gene transfer; the dislocation of native AMF communities, and the ecological impact of non-target organisms. In recent years, regulatory agencies have started to develop frameworks based on metagenomics-based environmental monitoring as tools to conduct post-application surveillance of the introduced strains from biofertilizers in the soil ecosystem.

8.4. Economic Feasibility and Scalability

The production of *G. intraradices* biofertilizers at a commercial scale is restricted by the plant's obligate biotrophic nature, as production must occur on living host root systems, which is much more costly and requires more space than using fermentation to produce bacterial biofertilizers (Berruti *et al*., 2016) [10] (Bhardwaj *et al*., 2014) [31]. However, root organ culture (ROC) systems using Ri-transformed carrot roots can produce *G. intraradices* at an *in vitro* level in a closed (i.e., contamination-free), controlled setting, although substantially at higher costs than that of competing bacterial-based products. In addition, metagenomic quality control pipelines for detecting contamination and verifying genetic identity of inoculum require additional labour and resources, which adds cost and time to the quality assurance process. According to economic models developed from various locations, *G. intraradices* biofertilizers will only produce favourable cost-benefit ratios when they are used along with significant reductions in the amount of chemical fertilizer. Thus, it is vitally important to demonstrate the ability to reduce inputs while maintaining yields during commercial validation trials (Bhardwaj *et al*., 2014) [31] (Pellegrino *et al*., 2015) [49].

9. Future Perspectives and Research Directions

9.1. Integration with AI and Systems Biology

The study of the *Glomus intraradices* mycorrhizal fungi and its interactions with other soil microbes and plants has witnessed an increase in the application of artificial intelligence (AI) and machine learning techniques to create metagenomic data for research purposes. Predictive bio-fertilizer performance models, strains for selection, and the optimization of field sampling have also seen new and improved methodologies through the use of artificial intelligence and machine learning (Lakshmanan *et al*., 2014) [42] (Müller *et al*., 2016) [46]. In addition, deep learning architectures have been applied to metagenomic datasets as a means of improving the identification of non-linear relationships between AMF colonisation and crop yield response. For example, convolutional neural networks for identifying sequence motifs and graph neural networks for analysing microbial interaction networks are two forms of deep learning architecture which hold promise for the identification of non-linear predictors of AMF colonisation and crop yield response from complex multi-omics datasets (Mendes *et al*., 2013) [47]. In addition, systems biology frameworks can model the *G. intraradices*-rhizosphere community as a dynamic metabolic network parameterised by metagenomics-derived stoichiometries to predict nutrient and metabolite fluxes under different soil conditions. The use of flux balance analysis (FBA) on community-based metabolic models has the potential to provide an in silico means of designing optimal consortia for testing prior to expensive field trials.

9.2. Advances in Multi-Omics Approaches

Utilizing single-nucleus and single-cell omics methods together with single-cell RNA sequencing developed for fungal cells (scRNA-seq) and droplet-based microfluidics would enable identification of transcriptional variances within individual hyphae of *G. intraradices* inhabiting different soil microhabitats, which has not been possible using bulk metatranscriptomic techniques (Müller *et al*., 2016) [46] (Mendes *et al*., 2013) [47]. Additionally, incorporating spatial metagenomic techniques that utilize in situ sequencing and imaging mass spectrometry will allow for an understanding of microbial community structure and distribution of metabolites over rhizosphere millimeter-scale gradients of *G. intraradices* and other microorganisms. Furthermore, advances in long-read/direct RNA sequencing (Oxford Nanopore), which do not require reverse transcription and can identify RNA modifications (e.g., N6-methyladenosine (m6A)), provides an additional epigenetic mechanism for regulating expression of genes that have yet to be discovered in AMF investigations (Jain *et al*., 2016) [17].

9.3. Development of Next-Generation Biofertilizers

The rational development of next generation biofertilizers based on *G. intraradices* will take advantage of data obtained from metagenomic & multi-omics studies that are generating new functional characteristics for biofertilizers beyond those currently provided by traditional inoculants. New forms of synthetic ecology approaches will allow for defined and minimal microbiome consortia composed of *G. intraradices* and selected keystone bacterial partners as identified by using network analyses, to provide consistent performance in multiple soil types and reduce formulation complexity (Busby *et al*., 2017) [45] (Lakshmanan *et al*., 2014) [42]. To enhance the shelf life of biofertilizers based on *G. intraradices*, several different encapsulation technologies such as alginate-clay nanocomposite matrices and biofilm-producing carriers will be developed. Targeted gene editing tools such as CRISPR are being developed, to enable targeted genetic modification of bacterial consortium partners, improving phytase activity, improving stress tolerance or removing potential virulence genes from consortium partners, with the intention of moving toward regulatory approval in the future across multiple jurisdictions.

9.4. Identification of Research Gaps and Innovation Pathways

There are still many critical areas of research that have not been fully addressed in the field of *G. intraradices* biofertilizer science. The details of how host specificity and preferential compatibility between plants and fungi are controlled at the molecular level are incomplete. This limits our current ability to predict colonisation models (Parniske, 2008) [38] (Wang and Qiu, 2006) [39]. We need long-term monitoring over many decades to understand how the application of *G. intraradices* as a biofertiliser affects the diversity of native AMF communities of fungi and the functioning of the soil ecosystem in general. Currently published literature is dominated by studies that only follow plants for one growing season (van der Heijden *et al*., 2015) [34] (Brundrett, 2002) [35]. Another major problem hindering our understanding of AMF is the issue of the 'dark

metagenome'; approximately 60% of genes from the vast majority of metagenomics datasets have no assigned functional characteristics, therefore there is a huge unexplored potential resource of biological functionality (Huson *et al*., 2016) [23]. The restrictive nature of equitable access to metagenomics technologies, substandard bioinformatics infrastructure, and the like; especially in low- and middle-income countries where smallholder farmers would benefit most from using biofertiliser; necessitates initiatives for international capacity-building in research and open-data sharing projects, amongst other things, to overcome this inequity (Busby *et al*., 2017) [45] (Mendes *et al*., 2013) [47].

Table 8: Comparative Analysis of *G. intraradices* with Other Common Biofertilizer Microorganisms

Organism	Type	Primary Nutrient	Host Specificity	Technology Readiness Level
Glomus intraradices	AMF	Phosphorus (P)	Broad (>80% plants)	TRL 7–9 (commercial)
Rhizobium spp.	N2-fixing bacteria	Nitrogen (N)	Legumes only	TRL 9 (widely commercial)
Azospirillum brasilense	PGPR	N (associative fixation)	Cereals, grasses	TRL 8–9
Bacillus subtilis	PGPR/biocontrol	P solubilisation	Non-specific	TRL 8–9
Pseudomonas fluorescens	PGPR	P solubilisation, ISR	Non-specific	TRL 7–8
Azotobacter chroococcum	Free-living N-fixer	Nitrogen (N)	Non-specific	TRL 7–8
Trichoderma harzianum	Biocontrol fungus	Biocontrol, P	Non-specific	TRL 8–9
Glomus mosseae	AMF	Phosphorus (P)	Broad	TRL 7–8
Frankia spp.	Actinorhizal bacteria	Nitrogen (N)	Actinorhizal shrubs	TRL 5–6

Note: TRL = Technology Readiness Level (1 = basic research; 9 = fully commercial). ISR = Induced Systemic Resistance.

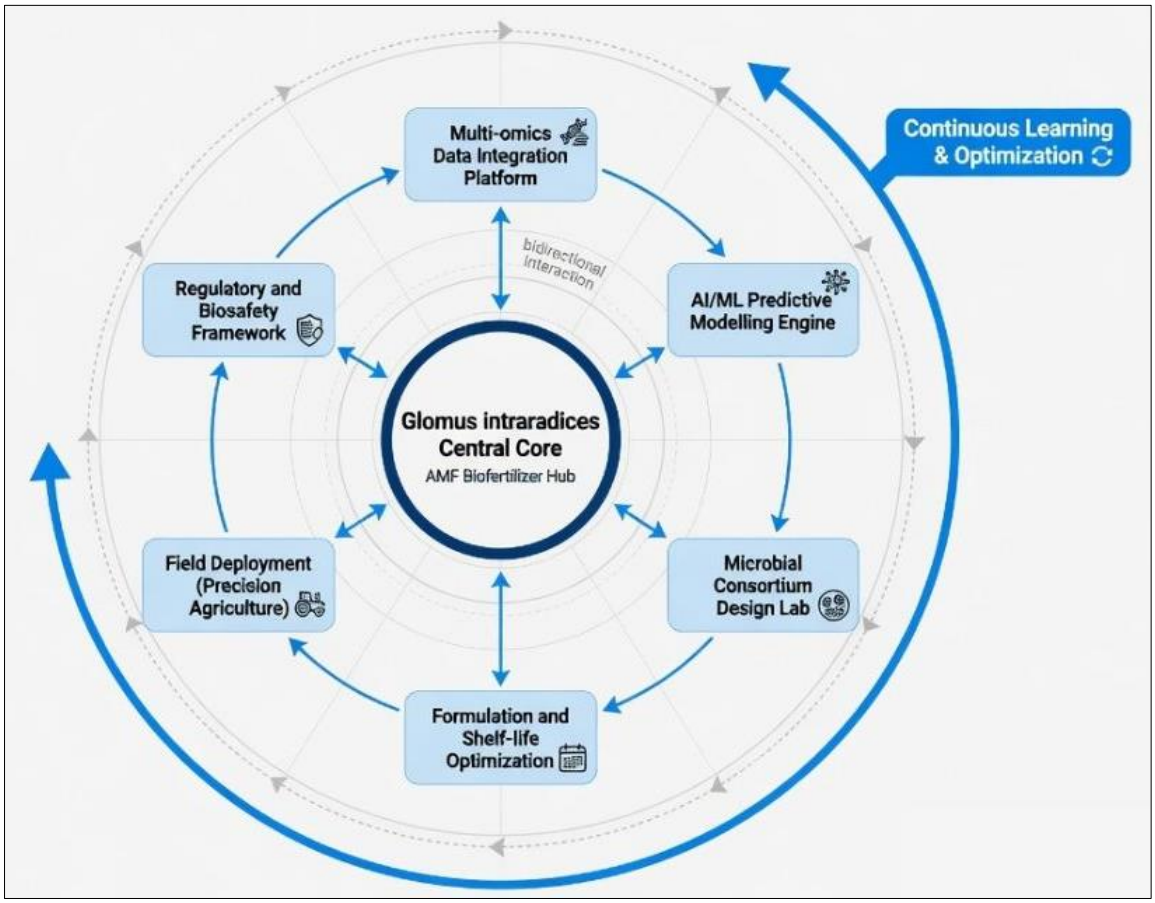


Fig 7: Integrated Framework for Enhanced *G. intraradices* Biofertilizer System Design

10. Conclusion

The application of metagenomic and bioinformatic approaches to bio-fertilizer systems that contain *G. intraradices* has changed the way we understand the mechanisms, design and implementation of biological soil amendments. In this review, we have shown through systematic studies that when combined with an integrated bioinformatic analysis pipeline (that includes assembly, taxonomic classification, functional annotation and the integration of multi-omic data) next generation sequencing technologies (including Illumina short-read, PacBio HiFi, and Oxford Nanopore long reads) provide unparalleled functional insight into *G. intraradices* and its associated microbial community in the rhizosphere. Among the major findings of this research are: the identification of novel gene variants of phosphatase transporters; identification of new synergistic members of the bacterial consortium associated with the hyphosphere; and the elucidation of mechanisms for the metabolic transfer of phosphorous to the plant, which supports the enhanced phosphorus acquisition capacity of AMF-based bio-fertilizers.

The agricultural implications of *G. intraradices* biofertilizers

are significant. See Table 7 for field trials that support *G. intraradices* biofertilizers reducing the requirement for chemical phosphate fertiliser by 25-40% while still producing equivalent crop yields for wheat, maize, legume, and horticultural systems in a range of regions worldwide. Other benefits associated with AMF biofertilizers include soil ecosystem co-benefits such as increased microbial diversity, improved aggregate stability, increased soil organic carbon, and reduced survival of soilborne pathogens. The combination of precision agriculture digital technologies with metagenomics-guided deployment of biofertilizers provides an exceptionally strong framework for developing and implementing a much more targeted, cost-effective, and environmentally sustainable nutrient management program.

Practical recommendations arising from this synthesis include: (1) the implementation of multi-platform hybrid sequencing strategies that integrate Illumina short reads with PacBio or Nanopore long reads to get the best genome assembly quality; (2) standardize hyphosphere sampling protocols and bioinformatics pipelines so that studies can be compared; (3) require metagenomic quality control metrics to be included in biofertilizer product registration dossiers; (4) create AI-assisted strain selection platforms that are trained on publicly available metagenomic datasets; and (5) set up open-access, globally federated AMF metagenome repositories to speed up knowledge exchange and make technology available to everyone. Future research should focus on long-term ecological monitoring of biofertilizer applications, the elucidation of the dark metagenome via functional metagenomics, and translational studies in underrepresented smallholder agricultural settings. The integration of metagenomics, systems biology, artificial intelligence, and precision agriculture establishes *G. intraradices*-based biofertilizer technology as an important part of the next generation of sustainable food systems.

References

1. Smith SE, Read DJ. Mycorrhizal Symbiosis. 3rd ed. Academic Press; 2008.
2. Schüssler A, Schwarzott D, Walker C. A new fungal phylum, the Glomeromycota: phylogeny and evolution. Mycological Research. 2001;105(12):1413–1421.
3. Giovannetti M, Avio L, Fortuna P, *et al*. At the root of the wood wide web: self-recognition and non-self incompatibility in mycorrhizal networks. Plant Signaling & Behavior. 2006;1(1):1–5.
4. Bonfante P, Genre A. Mechanisms underlying beneficial plant-fungus interactions in mycorrhizal symbiosis. Nature Communications. 2010;1:48.
5. Mäder P, Vierheilig H, Streitwolf-Engel R, *et al*. Transport of 15N between interacting mycorrhizal plants via hyphal networks. New Phytologist. 2000;146(3):405–415.
6. Willmann M, Gerlach N, Buer B, *et al*. Mycorrhizal phosphate uptake pathway in maize: vital for growth and cob development on nutrient poor agricultural and greenhouse soils. Frontiers in Plant Science. 2013;4:533.
7. Helber N, Wippel K, Sauer N, *et al*. A versatile monosaccharide transporter that operates in the arbuscular mycorrhizal fungus Glomus sp. is crucial for the symbiotic relationship with plants. The Plant Cell. 2011;23(10):3812–3823.
8. Remy W, Taylor TN, Hass H, Kerp H. Four hundred-million-year-old vesicular arbuscular mycorrhizae. Proceedings of the National Academy of Sciences of the USA. 1994;91(25):11841–11843.
9. Varela BJ, Lesbarrères D, Ibáñez R, Green DM. Environmental and host effects on skin bacterial community composition in Panama-native frogs. Frontiers in Microbiology. 2018;9:298.
10. Berruti A, Lumini E, Balestrini R, Bianciotto V. Arbuscular mycorrhizal fungi as natural biofertilizers: let's benefit from past successes. Frontiers in Microbiology. 2016;6:1559.
11. Dowd SE, Zaragoza J, Rodriguez JR, *et al*. Windows .NET network distributed basic local alignment search toolkit (W.ND-BLAST). BMC Bioinformatics. 2005;6:93.
12. Handelsman J. Metagenomics: application of genomics to uncultured microorganisms. Microbiology and Molecular Biology Reviews. 2004;68(4):669–685.
13. Quince C, Walker AW, Simpson JT, Loman NJ, Segata N. Shotgun metagenomics, from sampling to analysis. Nature Biotechnology. 2017;35(9):833–844.
14. Myllykangas S, Buenrostro J, Ji HP. Overview of sequencing technology platforms. In: Bioinformatics for High Throughput Sequencing. Springer; 2012. p. 11–25.
15. Illumina Inc. An introduction to next-generation sequencing technology [Internet]. 2023 [cited 202[4]. Available from: https://www.illumina.com/science/technology/next-generation-sequencing.html
16. Eid J, Fehr A, Gray J, *et al*. Real-time DNA sequencing from single polymerase molecules. Science. 2009;323(5910):133–138.
17. Jain M, Olsen HE, Paten B, Akeson M. The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. Genome Biology. 2016;17(1):239.
18. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler Aligner. Bioinformatics. 2009;25(14):1754–1760.
19. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nature Methods. 2012;9(4):357–359.
20. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. metaSPAdes: a new versatile metagenomic assembler. Genome Research. 2017;27(5):824–834.
21. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. Nature Methods. 2015;12(1):59–60.
22. Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome Biology. 2014;15(3):R46.
23. Huson DH, Beier S, Flade I, *et al*. MEGAN Community Edition—interactive exploration and analysis of large-scale microbiome sequencing data. PLOS Computational Biology. 2016;12(6):e1004957.
24. Truong DT, Franzosa EA, Tickle TL, *et al*. MetaPhlAn2 for enhanced metagenomic taxonomic profiling. Nature Methods. 2015;12(10):902–903.
25. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Research. 2000;28(1):27–30.
26. Finn RD, Clements J, Eddy SR. HMMER web server: interactive sequence similarity searching. Nucleic Acids Research. 2011;39(suppl 2):W29–W37.
27. Zhu W, Lomsadze A, Borodovsky M. Ab initio gene identification in metagenomic sequences. Nucleic Acids

Research. 2010;38(12):e132.
28. Rawat N, Tiwari VK. Phosphate solubilizing microbes: mechanism and their role in phosphate solubilization and uptake. Journal of Pharmacognosy and Phytochemistry. 2018;7(2):1785–1793.
29. Richardson AE, Barea JM, McNeill AM, Prigent-Combaret C. Acquisition of phosphorus and nitrogen in the rhizosphere and plant growth promotion by microorganisms. Plant and Soil. 2009;321(1–2):305–339.
30. Barea JM, Pozo MJ, Azcón R, Azcón-Aguilar C. Microbial co-operation in the rhizosphere. Journal of Experimental Botany. 2005;56(417):1761–1777.
31. Bhardwaj D, Ansari MW, Sahoo RK, Tuteja N. Biofertilizers function as key player in sustainable agriculture by improving soil fertility, plant tolerance and crop productivity. Microbial Cell Factories. 2014;13(1):66.
32. Compant S, Clément C, Sessitsch A. Plant growth-promoting bacteria in the rhizo- and endosphere of plants: their role, colonization, mechanisms involved and prospects for utilization. Soil Biology and Biochemistry. 2010;42(5):669–678.
33. Finlay RD. Ecological aspects of mycorrhizal symbiosis: with special emphasis on functional diversity of interactions involving the extraradical mycelium. Journal of Experimental Botany. 2008;59(5):1115–1126.
34. van der Heijden MGA, Martin FM, Selosse MA, Sanders IR. Mycorrhizal ecology and evolution: the past, the present, and the future. New Phytologist. 2015;205(4):1406–1423.
35. Brundrett MC. Coevolution of roots and mycorrhizas of land plants. New Phytologist. 2002;154(2):275–304.
36. Cameron DD, Neal AL, van Wees SCM, Ton J. Mycorrhiza-induced resistance: more than the sum of its parts? Trends in Plant Science. 2013;18(10):539–545.
37. Govindarajulu M, Pfeffer PE, Jin H, *et al*. Nitrogen transfer in the arbuscular mycorrhizal symbiosis. Nature. 2005;435(7043):819–823.
38. Parniske M. Arbuscular mycorrhiza: the mother of plant root endosymbioses. Nature Reviews Microbiology. 2008;6(10):763–775.
39. Wang B, Qiu YL. Phylogenetic distribution and evolution of mycorrhizas in land plants. Mycorrhiza. 2006;16(5):299–363.
40. Kiers ET, Duhamel M, Beesetty Y, *et al*. Reciprocal rewards stabilize cooperation in the mycorrhizal symbiosis. Science. 2011;333(6044):880–882.
41. Ferrol N, Tamayo E, Vargas P. The heavy metal paradox in arbuscular mycorrhizas: from mechanisms to biotechnological use. Journal of Experimental Botany. 2016;67(22):6253–6265.
42. Lakshmanan V, Selvaraj G, Bais HP. Functional soil microbiome: belowground solutions to an aboveground problem. Plant Physiology. 2014;166(2):689–700.
43. Agler MT, Ruhe J, Kroll S, *et al*. Microbial hub taxa link host and abiotic factors to plant microbiome variation. PLOS Biology. 2016;14(1):e1002352.
44. Berendsen RL, Pieterse CMJ, Bakker PAHM. The rhizosphere microbiome and plant health. Trends in Plant Science. 2012;17(8):478–486.
45. Busby PE, Soman C, Wagner MR, *et al*. Research priorities for harnessing plant microbiomes in sustainable agriculture. PLOS Biology. 2017;15(3):e2001793.
46. Müller DB, Vogel C, Bai Y, Vorholt JA. The plant microbiota: systems-level insights and perspectives. Annual Review of Genetics. 2016;50:211–234.
47. Mendes R, Garbeva P, Raaijmakers JM. The rhizosphere microbiome: significance of plant beneficial, plant pathogenic, and human pathogenic microorganisms. FEMS Microbiology Reviews. 2013;37(5):634–663.
48. Lugtenberg B, Kamilova F. Plant-growth-promoting rhizobacteria. Annual Review of Microbiology. 2009;63:541–556.
49. Ortiz N, Armada E, Duque E, Roldán A, Azcón R. Contribution of arbuscular mycorrhizal fungi and/or bacteria to enhancing plant drought tolerance under natural soil conditions. Journal of Plant Physiology. 2015;174:87–96.
50. Pellegrino E, Bedini S, Nuti M, Ercoli L. Impact of AM fungi application on maize production and soil microbial community under different farming systems. Science of the Total Environment. 2015;531:1–11.

How to Cite This Article
Falquet N. Metagenomic and bioinformatics approaches to enhance Glomus intraradices-based biofertilizer systems. J Agric Digit Res. 2026;7(1):83-97.
doi:10.54660/JADR.2026.7.1.83-97.