Improving the efficiency of metagenomic analysis of soil samples.

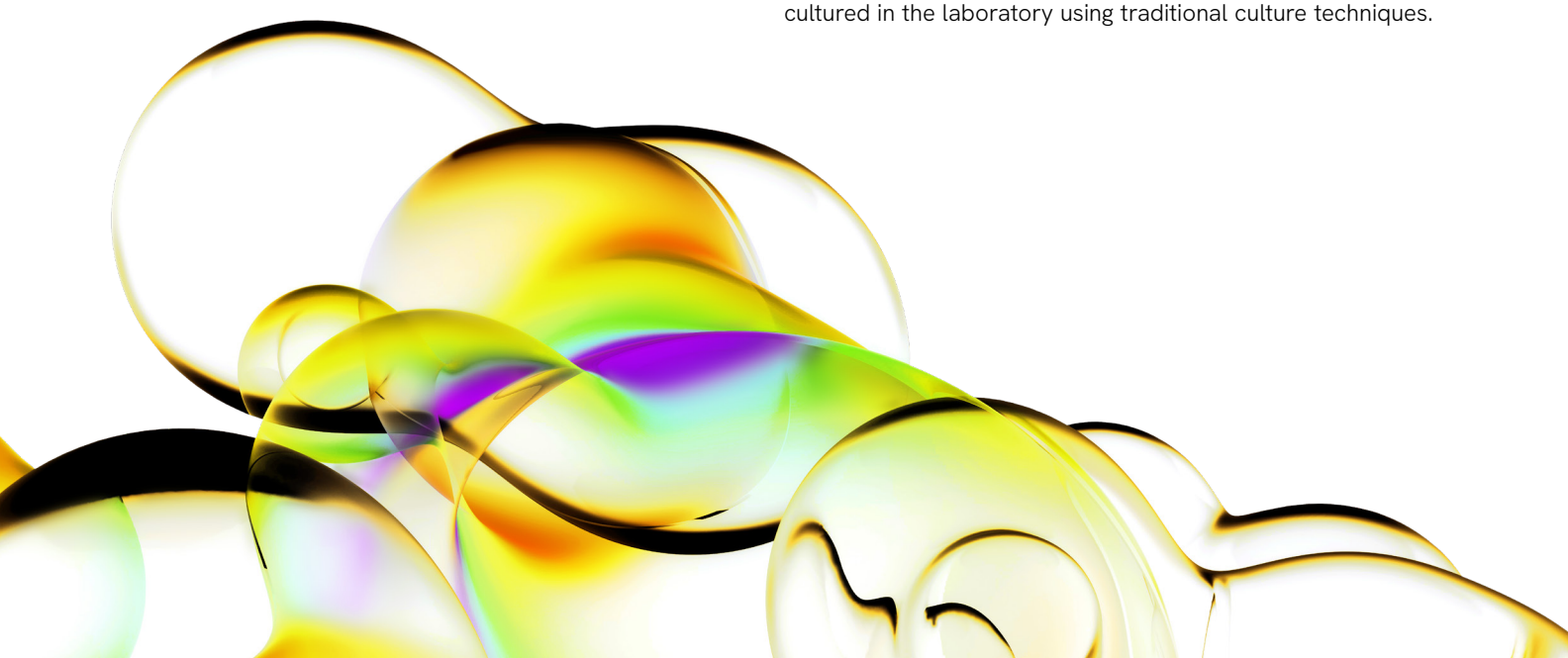
Introduction

Understanding and characterizing soil metagenomic populations can enable regulation of plant performance through improved bioavailability of nutrients, secretion of bioactive compounds, and hormones that can promote or repress plant growth.

Depending on external factors such as land management, crop rotation, use of pesticides and fertilizers, temperature, and pH, all can have an impact on the soil microbiota and the expression and excretion of extracellular metabolites which can directly impact the plant viability and vigour through a symbiotic relationship.

Studies have shown increased concentrations of heavy metals such as Cadmium (Cd) influence the microbial population. Cd-contamination has the potential to reduce microbial diversity by altering the community structure within the soil. Non-contaminated soil has increased microbial diversity, genes and enzymes involved in metabolic pathways of microbial communities than the Cd-contaminated soil.[1] Metabolism of different microorganisms that change in response to different environmental conditions can be studied by analysis of their metabolic footprint and metagenomics sequence analysis.

Previously, studies on the development of microbial communities required the isolation of microbes from soil samples by culture dependent techniques followed by a series of tests for phenotypic evaluation and their identification. Microbial diversity studies conducted in soil have been biased due to the difficulty in culturing many of the microbes. It has been reported that only 1% of soil microbes can be cultured in the laboratory using traditional culture techniques.



To overcome these limitations scientists have adopted culture-independent techniques. These techniques employ molecular biology-based methods, in which the DNA from the soil microbiome is extracted and subsequently subjected to PCR amplification and sequencing reactions.[2]

Here Revvity® and Illumina® describe a streamlined, automated workflow from DNA extraction to sequence-ready libraries for metagenomic shotgun sequencing resulting in high quality data for species identification and functional profiling.

Workflow overview

Automation of an NGS-based metagenomics workflow offers significant advantages over manual sample preparation. Increased throughput and scalability, reduction in human touch-points and error, enhanced consistency and reproducibility, and increased speed all contribute to reliable data production amenable to varying throughputs for metagenomics labs.

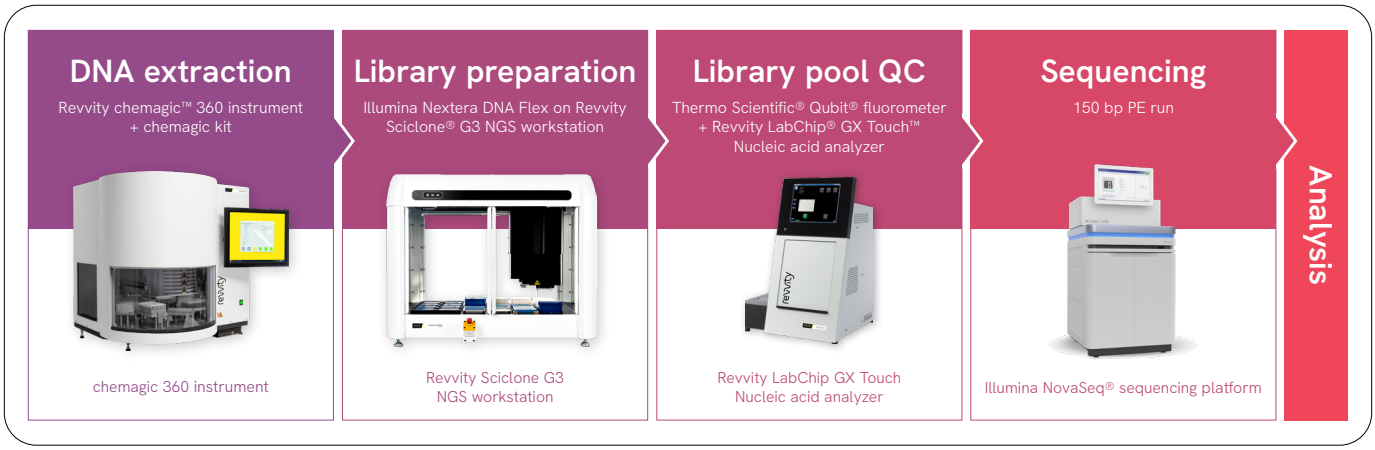


Figure 1. The automated Illumina DNA Prep workflow for metagenomics. Illumina and Revvity collaborated to create a fully automated NGS library preparation workflow for high-throughput metagenomics.

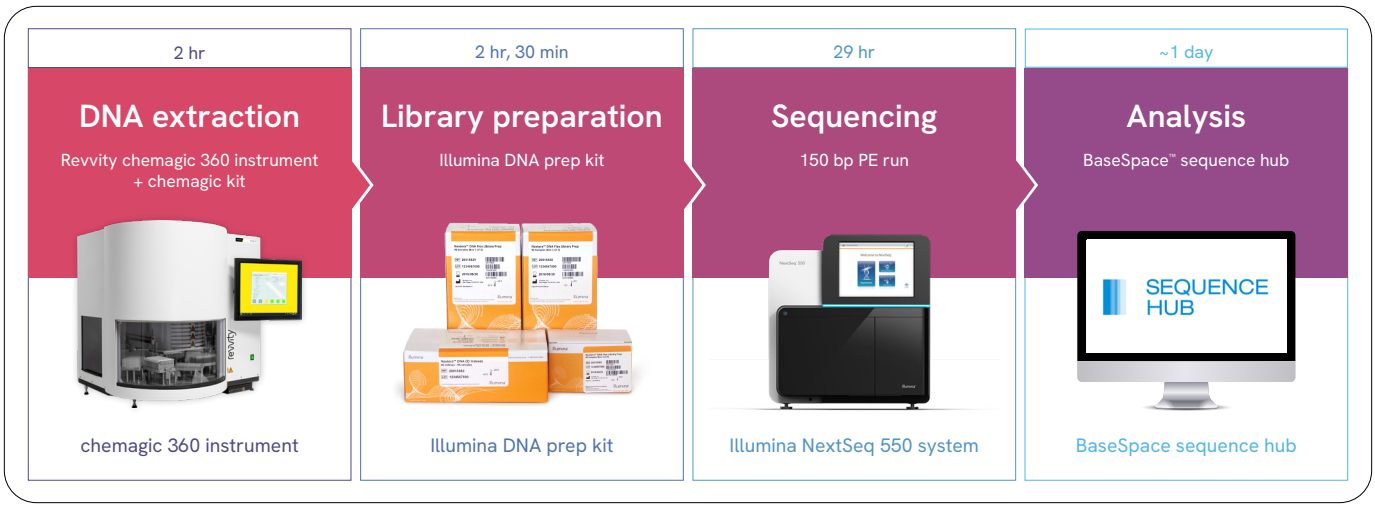


Figure 2. A comprehensive workflow featuring automated DNA extraction for soil metagenomic shotgun sequencing. Illumina and Revvity collaborated to deliver an automated approach for soil genomic DNA extraction while offering flexibility for library preparation and an in-depth solution for metagenomic analysis.

Optimized DNA isolation from soil samples

Correctly representing the genetic diversity of soil samples is a challenging problem, complicated further by biases in DNA extraction and the presence of potential inhibitors like humic substances which might interfere with downstream applications leading to misrepresentation of the microbiota.

In a collaboration with Illumina, we have developed a robust automated system to isolate pure DNA from soil samples to detect a diverse spectrum of organisms, including bacteria, fungi and viruses.

Revvity chemagic technology addresses the challenges of extracting DNA from soil samples suitable for library preparation and sequencing reactions with the chemagic DNA extraction kit for soil samples. With this kit, cells are lysed efficiently in an external incubation step. A subsequent precipitation step removes potential inhibitors.

After centrifugation, DNA is isolated from lysates with the chemagic 360 instrument.[3] This automated magnetic separation procedure uses chemagen M-PVA magnetic beads to isolate and purify nucleic acids. The beads have a high affinity for nucleic acids and low inhibitor binding, resulting in ultra-pure DNA or RNA. chemagen technology features magnetizable rotating rods, combining the transfer and suspension of magnetic beads to extract DNA, preventing further fragmentation. Efficient washing procedures in the automated DNA extraction process eliminate the last traces of inhibitors. This guarantees a high recovery of pure DNA suitable for downstream applications.

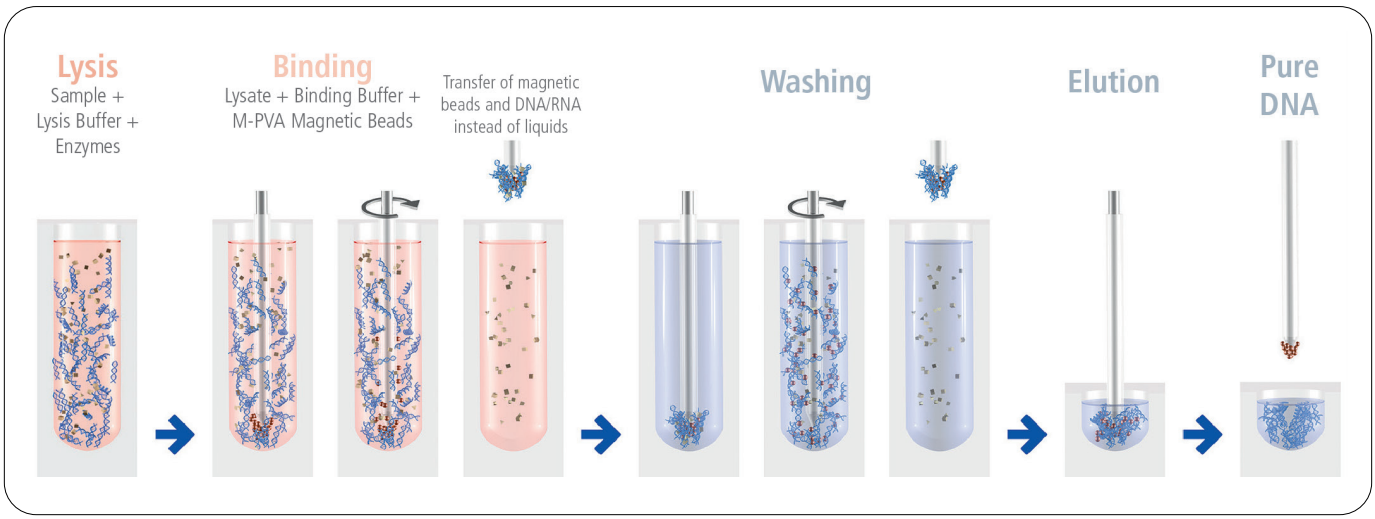


Figure 3. chemagic technology. Uses automated magnetizable rotating rods, combining the transfer and suspension of magnetic beads, to isolate ultrapure, high molecular weight DNA.

Sample collection & DNA isolation

Soil samples from different sources including a bank mud, and strawberry, beet, wheat, and hay fields were collected. The DNA was extracted in four iterations per sample type with the chemagic DNA soil kit H96 (CMG-1054) and the chemagic 360 instrument.

DNA yield & purity

Table 1. DNA yield and purity measured with UV photo spectrometer. Input amount was 250 mg of soil material

Sample type	Mean purity	Mean 260/230 ratio	Mean yield (µg)
Beet Field	1.81	1.30	1.28
Wheat Field	1.82	1.37	1.85
Hay Field	1.76	1.34	2.42
Strawberry Field	1.65	1.20	1.19
Bank Mud	1.82	1.52	1.50

DNA quality assesment

LabChip GX Touch nucleic acid analyzer microfluidics technology delivers rapid capillary electrophoresis analysis for DNA sample quality control.[4] Eluates were analyzed on the LabChip GX Touch nucleic acid analyzer using the genomic DNA chip and reagents. The genomic quality score (gQS) measures the degree of degradation of the sample; a score of 5 indicates intact DNA and 0 indicates degraded DNA. In the gQS, the size distribution of the sample is also analyzed; with increasing degradation the peak shifts to the left and the gQS is reduced.

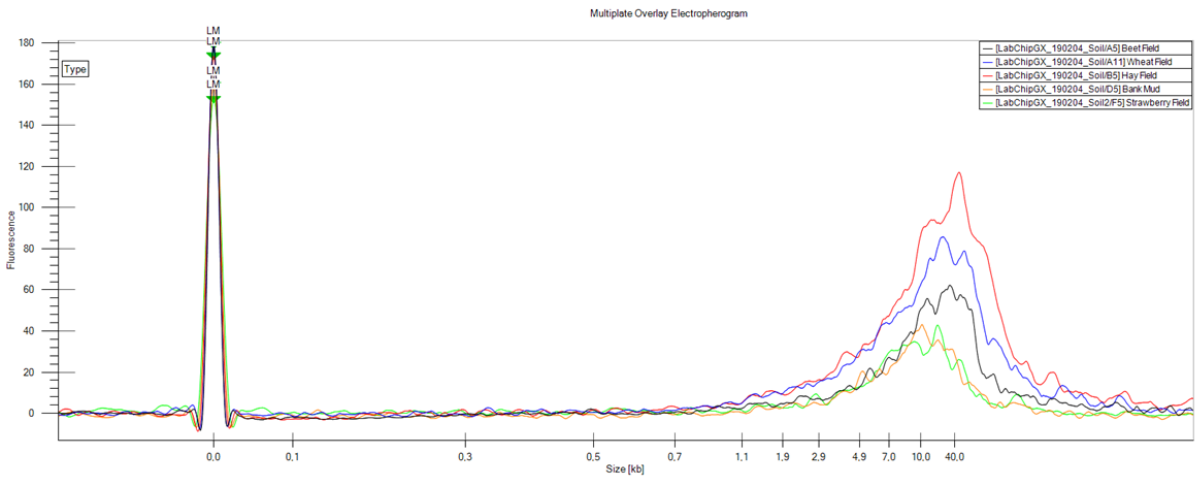


Figure 4. DNA gQS was determined using the LabChip GX Touch nucleic acid analyzer and the genomic DNA chip and reagents. Electropherograms show traces of high quality genomic DNA for beet field (gQS=4.3), wheat field (gQS=4.1), hay field (gQS=4.2), strawberry field (gQS=4.0), and bank mud (gQS=4.0).

Performance in QPCR

DNA eluates obtained from bank mud were tested in a qPCR targeting a bacterial gene.

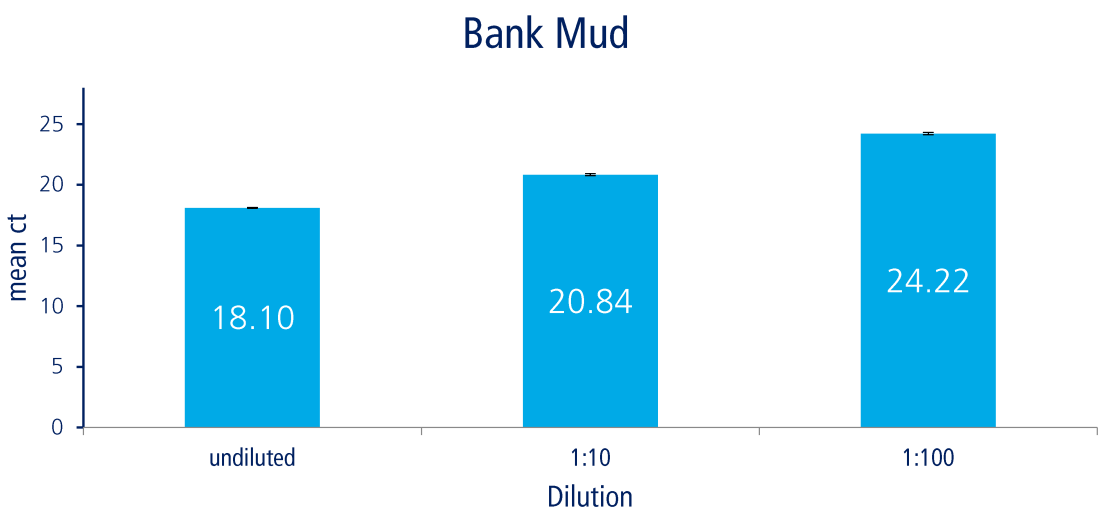


Figure 5. Mean C_T values of eluates from Bank Mud in quantitative PCR for a bacterial target. Eluates were tested undiluted, tenfold diluted and hundredfold diluted. The c_T differences for diluted versus undiluted eluates and tenfold diluted versus hundredfold diluted eluates are close to the perfect value of 3.3 indicating the absence of any inhibitors left in the eluates.

Library preparation

To generate libraries for metagenomic shotgun sequencing, environmental DNA was used as input for the Illumina DNA Prep kit.[5] This kit offers flexibility in terms of the DNA input range (1-500 ng) and multiplexing capacity, generating highly uniform libraries for sequencing (Figure 6).

Additionally, due to the unique tagmentation chemistry from Illumina DNA Prep, saturation of beads using 100-500 ng of DNA eliminates the need for library quantification prior to pooling, rendering this kit both time and cost-effective for high-throughput soil metagenomic studies.

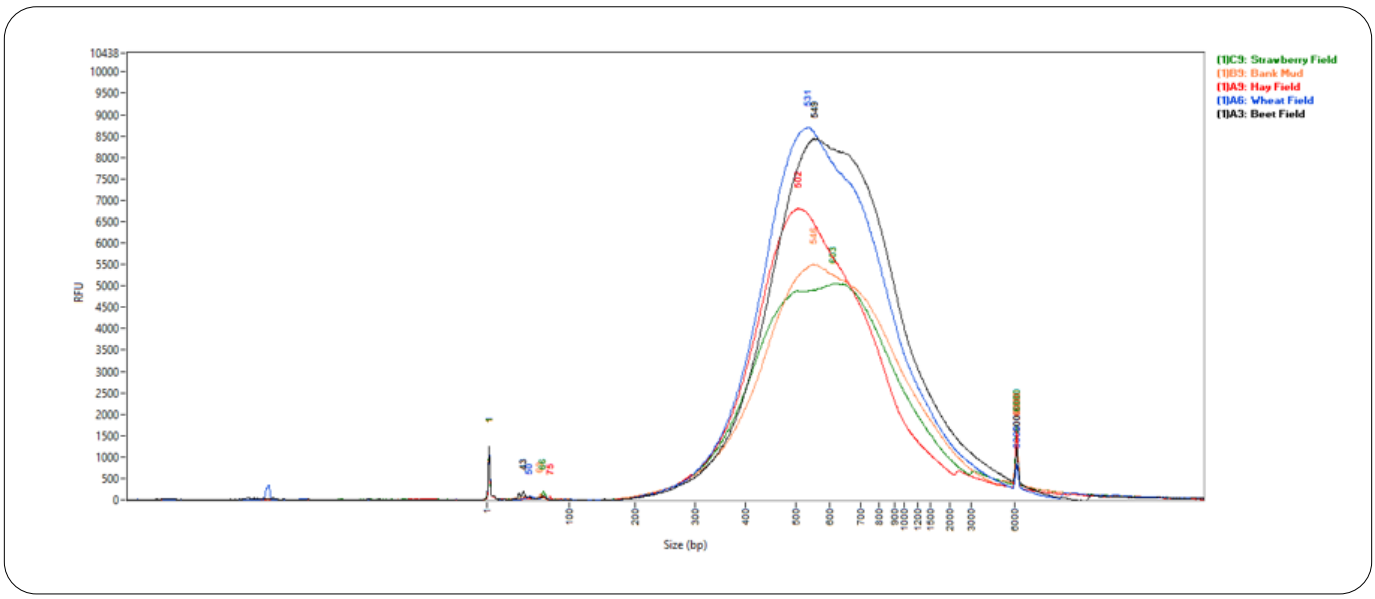


Figure 6. Quality control for Illumina DNA Prep libraries. Electropherograms show traces for soil libraries post-PCR, averaging approximately 550 bp with adapters.

Sequencing

Pooled libraries were sequenced on the Illumina NextSeq™ 550 instrument with a 2x151 run configuration. Over 400M high quality reads were generated from each sequencing run, demonstrating scalability for various metagenomic studies (Table 2).

Table 2. Primary metrics for 2x151 sequencing runs on an Illumina NextSeq 550 instrument. Results show primary metrics (% PF, PF (M), %>=Q30, and error rate) from two sequencing runs using the high output mode.

Pool	%PF	PF (M)	%>=Q30	Error rate (%)
1	78.555	458.37	80.1	2.275
2	82.595	481.97	82.7	1.715

To identify microorganisms present in each environmental sample, taxonomic classification was performed using Kraken 2, a k-mer alignment tool.[6,7] Kraken 2 matches k-mer sequences to a user-directed microbial genome database to assign taxonomy to each read. Using this approach, taxa belonging to various domains were detected, including bacteria, virus, eukaryota and more, highlighting one of the advantages with performing shotgun sequencing over traditional 16S amplicon sequencing for metagenomics analysis (Table 3). Furthermore, this summary table illustrates the need for further characterization of culture-free environmental samples due to the abundance of unclassified reads stemming from limitations in current reference genome databases.

Table 3. Summary of Kraken2 taxonomic classification using Pavian. The abundance of taxa traced to bacterial, viral, fungal, and protozoan lineages are summarized for 5 soil samples. This table also includes the percentages of reads unclassified due to limitations with current reference genome datasets.

Name	Classified reads	Unclassified reads	Microbial reads	Bacterial reads	Viral reads	Fungal reads	Protozoan reads
Bank Mud	18.3%	81.7%	18.3%	17.6%	0.013%	0.0442%	0.0104%
Beet Field	22.5%	77.5%	22.5%	21.8%	0.0134%	0.0451%	0.00938%
Hay Field	23.7%	76.3%	23.6%	22.8%	0.0148%	0.0534%	0.0101%
Strawberry Field	19.1%	80.9%	19.1%	18.4%	0.0128%	0.0458%	0.0103%
Wheat Field	25.1%	74.9%	25.1%	24.4%	0.0136%	0.0501%	0.0091%

Metagenomic shotgun sequencing provides a relatively unbiased assessment of microbial composition and supports high resolution of detection at the species level (Figure 7). Robust detection of genera commonly found in soil were detected, such as *Streptomyces* and *Bradyrhizobium*, which participate in clinically-relevant antibiotic production and nitrogen fixation, respectively.[9,10] The ability to identify specific species within each genus enables increased precision in associating microorganisms with their functional potential.

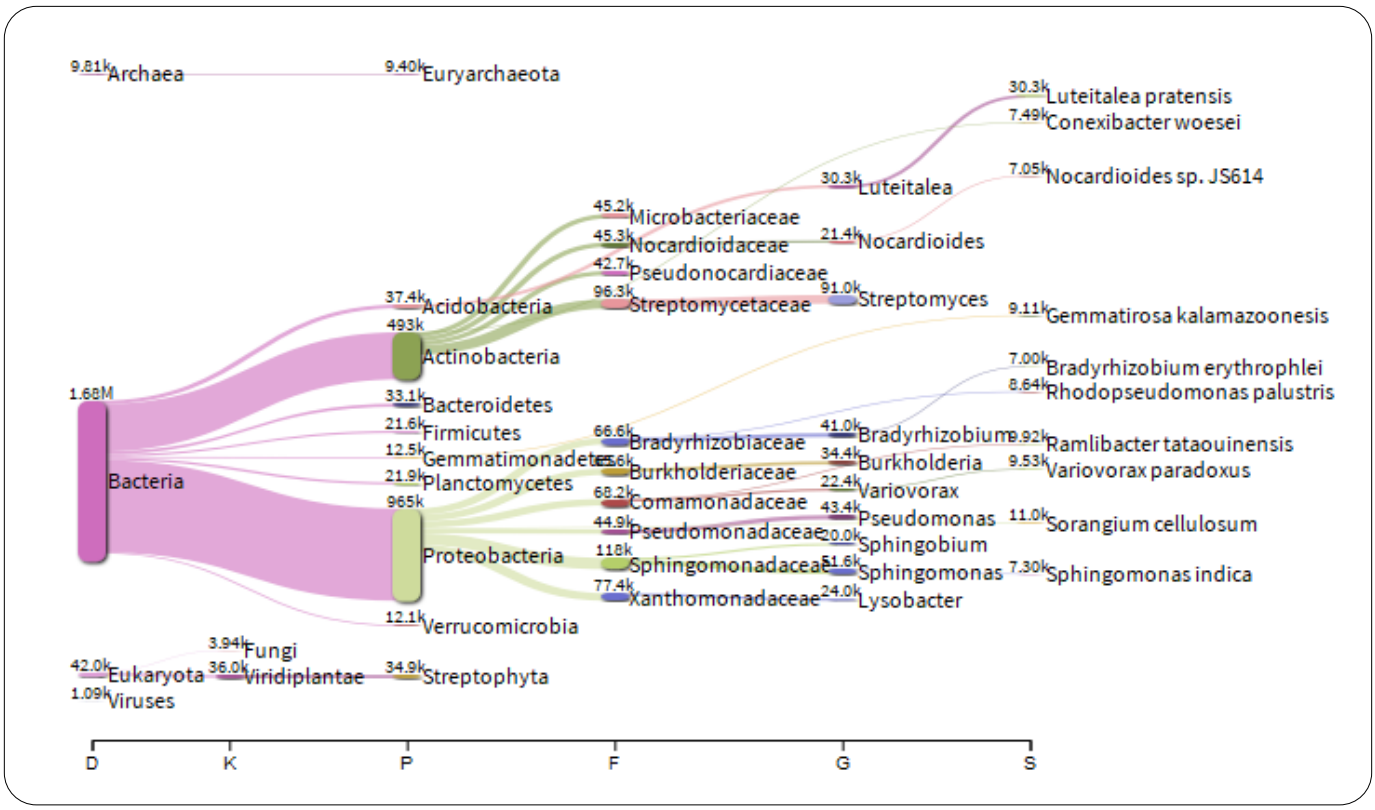


Figure 7. Visualization of taxonomic classification using Pavian.[8] Output from Kraken2-mediated taxonomic classification was used as input for Pavian, showing presence of highly abundant microbes from the hay field soil.

Soil provides a home to a large variety of microorganisms which contribute to the most diverse habitat on earth.[11] The large biodiversity of soil is important for the synergistic interactions between microbes which carry out many important environmental and clinically-relevant functions, such as plant growth and antibiotic production. Thus, it is not surprising when compared to stool microbial communities, the alpha-diversity of soil samples is significantly greater (Figure 8).

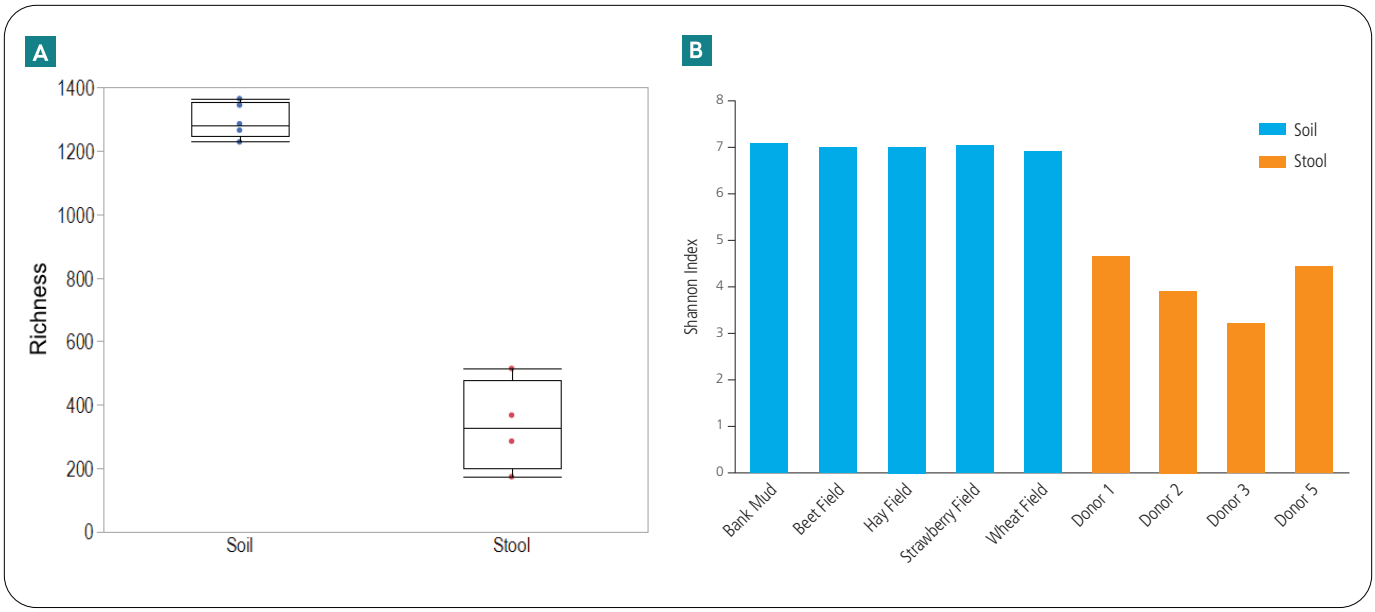


Figure 8. Alpha diversity measurements comparing soil to stool. (A) The richness, a measure of the number of species detected, was calculated for 5 soil samples (bank mud, beet field, hay field, strawberry field, and wheat wield) and 4 stool samples (Donor 1, Donor 2, Donor 3, Donor 5). **(B)** The Shannon index, a measure of richness and evenness, was calculated for the same samples. Both metrics show greater diversity in the soil samples when compared to the stool samples.

Overall this workflow provides an efficient and unbiased assessment of the microbial composition present in soil samples. Using chemagic technology for automated soil extraction in combination with Illumina DNA Prep kit for shotgun sequencing, species-level resolution was achieved. Hierarchical clustering was able to resolve key differences in microbial composition between soil and stool samples, enabling identification of top species that drive the distinct cluster patterns (Figure 9).

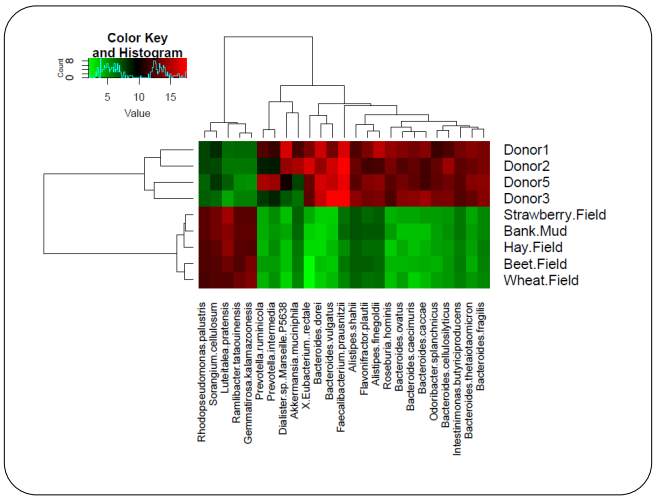


Figure 9. Hierarchical clustering visualized with heatmap. At the species level, the microbiome in the soil and stool samples cluster distinctly from each other, demonstrating key differences in microbial composition between sample types.

Summary

The automation of the Illumina DNA Prep metagenomic shotgun workflow using the Revvity chemagic technology for automated soil extraction provides an efficient and unbiased assessment of the microbial composition present in soil samples, making it an excellent solution for high-throughput metagenomics laboratories.

References

1. Feng et al. Metagenomic analysis of microbial community and function involved in cd-contaminated soil, BMC Microbiol. 2018; 18: 11. Published online 2018 Feb 13. doi: 10.1186/s12866-018-1152-5

2. Fatima et al. An Improved Method for Soil DNA Extraction to Study the Microbial Assortment within Rhizospheric Region Molecular Biology International Volume 2014, Published 15 September 2014, Article ID 518960

3. Revvity (2016). Compact, high volume, high throughput nucleic acid isolation. Document #CT6/30/0116-01. Revvity (2016). Automated, high performance electrophoresis for genomics. Document #011784B_01.

4. Revvity (2016). Automated, high performance electrophoresis for genomics. Document #011784B_01.

5. Illumina (2017). Illumina® DNA Prep Kit Data Sheet. Accessed March 28, 2019.

6. Wood DE, Salzberg SL: Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome Biol. 2014, 15: R46. Published online 2014 Mar 2. doi: 10.1186/gb-2014-15-3-r46

7. Kraken2 Metagenomics. https://basespace.illumina.com/apps/6863857/Kraken2-Metagenomics?preferredversion. Accessed March 28, 2019.

8. Breitweiser FP, Salzberg SL. Pavian: Interactive Analysis of metagenomics data for microbiomics and pathogen identification. BioRxiv 084715 [Preprint]. October 31, 2016 [cited 2019 Mar 28]. Available from https://dx.doi.org/10.1101/084715.

9. Manteca A and Sanchez J. Streptomyces Development in Colonies and Soils. Appl Environ Microbiol. 2009; 75: 2920-2924. Published online 2009 May. doi: 10.1128/AEM.02288-08

10. Hara S et al. Identification of Nitrogen-Fixing Bradyrhizobium Associated With Roots of Field-Grown Sorghum by Metagenome and Proteome Analyses. Front Microbiol; 10; 407. Published online 2019 Mar 12. doi: 10.3389/fmicb.2019.00407

11. Nesme J et al. Back to the Future of Soil Metagenomics. Front Microbiol. 2016; 7: 73. Published online 2016 Feb 10. doi: 10.3389/fmicb.2016.00073

Revvity, Inc.
940 Winter Street
Waltham, MA 02451 USA
www.revvity.com