Metagenomics Core Control Set

Maximize insights from every sample with Sequins™ internal standards

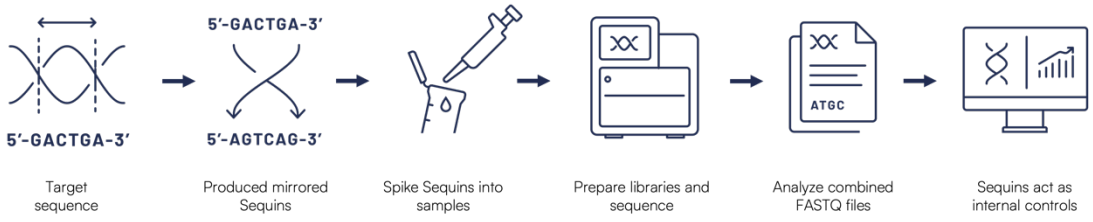
Background

Next generation sequencing (NGS) enables comprehensive and high-throughput analysis of microbial communities directly from diverse sample types, providing unique insights into the taxonomic composition, functional potential and their associated impact on health and disease. However, numerous variables influence the resolution and completeness of microbial profiles. Sequencing errors can introduce artefacts, and together with PCR amplification bias during library preparation, can skew the representation of certain taxa or genes. During downstream bioinformatic analysis, outcomes may vary between software tools and parameter settings, leading to misinterpretation of microbial diversity. These, and other variables, accumulate throughout the NGS workflow, confounding the accurate assessment of microbial communities and their biological roles.

Introduction to Sequins

Sequins (sequencing spike-ins) are synthetic nucleic acid controls that directly mirror naturally occurring sequences. Because Sequins retain the same nucleotide composition as the natural sequence, they enable accurate representation of genomic complexity without compromising integrity of the sample and results. Sequins perform equivalently throughout sequencing workflows, providing a true measure of control.

Sequins' innovative design enables production of synthetic mirrored sequences that directly represent almost any genomic feature, in any organism with a reference genome. Common and ecologically significant features, such as microbial markers or virulence-associated genes and features from analytically challenging regions of metagenomes, can be represented. By combining Sequins in precise stoichiometric ratios, quantification of low-abundance species can also be assessed. Sequins are simply 'spiked-in' to a sample prior to library preparation and progressed together through a workflow. Sequins controls can then be distinguished from native microbial DNA in subsequent data by virtue of their synthetic sequence. This enables normalization and comparison between samples, runs, laboratories, chemistries, and sequencers.



Schematic showing the design and use of Sequins in an NGS workflow.

Sequins for Metagenomics

The rapid advancements in NGS, have made Shotgun Metagenomics a powerful tool for a wide range of applications in microbial ecology and pathogen detection. Not only does metagenomics allow the comprehensive study of non-culturable or difficult-to-culture microbes, it is also instrumental in detecting novel and dangerous pathogens, monitoring pathogens of public health importance, and screening for antimicrobial resistance genes.

Despite its advantages, metagenomics faces challenges due to the inherent complexity of microbial communities, technical biases in NGS, and accuracy of species detection and quantification at low abundances. The extensive diversity and size of microbial genomes within a sample, can complicate accurate identification and quantification. To enhance the robustness of metagenomic analysis and comparisons between different metagenome samples, the inclusion of metagenome reference standards is crucial. Traditional mock microbial communities, where multiple microbes have been individually cultured and combined at known abundances to form a community, have been used as process controls. However, these are limited in their application as they cannot be added directly to samples without potentially interfering with downstream analyses.

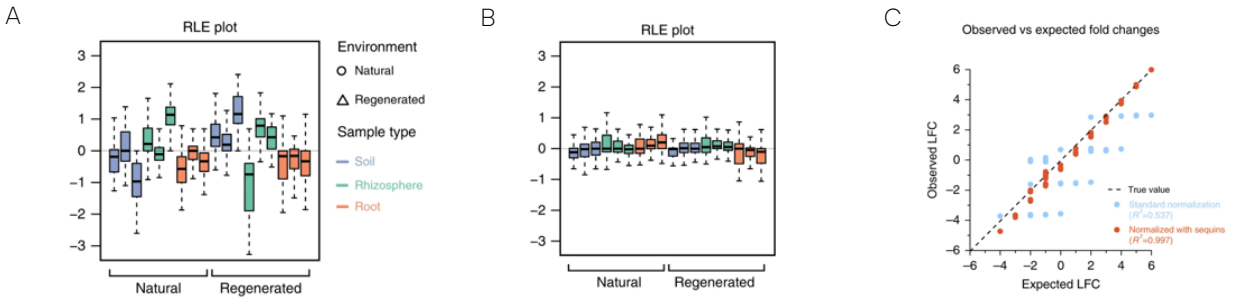
The Sequins Metagenomics Core Control Set is a comprehensive, easy-to-use, pre-configured control set with built-in redundancy that ensures integrity of results. It contains synthetic sequences that collectively represent 52 microbial species at varying abundances to emulate the complexity found in a natural microbial community. Features include: quantitative ladders; wide representation of domains including bacteria, archaea and eukaryotes; GC% content range (27% to 71%); single tube blend; demonstrated on various samples using Illumina and Oxford Nanopore technologies. Sequins are spiked into a sample at approximately 1% of total calculated input to a 250ng library.

Sequins Metagenomics Core Control Set content

Genus	No. Species	Genus	No. Species	Genus	No. Species
Aciduliprofundum	1	*Fusobacterium*	1	*Porphyromonas*	1
Acinetobacter	2	*Helicobacter*	1	*Roseobacter*	1
Anabaena	1	*Klebsiella*	1	*Salmonella*	1
Bacillus	2	*Lactobacillus*	1	*Shigella*	1
Buchnera	1	*Legionella*	1	*Staphylococcus*	1
Caldicellulosiruptor	1	*Leuconostoc*	1	*Streptococcus*	1
Candida	1	*Listeria*	2	*Synechocystis*	1
Candidatus Caldiarchaeum	1	*Magnetococcus*	1	*Thermotogan*	1
Candidatus Carsonella	1	*Metallosphaera*	1	*Thermus*	1
Candidatus Korarchaeum	1	*Methanopyrus*	1	*Toxoplasma*	1
Chlamydia	1	*Mycobacterium*	1	*Treponema*	1
Chlorobium	1	*Neisseria*	2	*Vibrio*	1
Corynebacterium	1	*Nitrosopumilus*	1	*Vulcanisaeta*	1
Desulfitobacterium	1	*Nostoc*	1	*Wolinella*	1
Desulfovibrio	1	*Oenococcus*	1	*Yersinia*	1
Ehrlichia	1	*Persephonella*	1		

Sequins controls are directly spiked into samples prior to library preparation to account for workflow variability and provide both identification and quantification of microbes, while also enabling normalization for inter-sample and inter-run comparison. Sequins are compatible with most sequencing technologies, including short- and long-read approaches, and have been used with a range of library preparation methods. A Sequins-specific FASTA file is provided for simple concatenation with a reference genome for alignment and further analysis using standard pipelines and tools.

Sequins Performance



Sequins provide superior normalization and quantification performance. The relative log expression plot for metagenomics samples from Soil, Rhizosphere and Root in Natural and Regenerated conditions, showed marked improvement when normalized with Sequins (A and B). Following normalization using Sequins, there was a significantly higher corelation of Log-Fold Changes for microbial species with expected change, versus using standard normalization (C). Modified from Hardwick et al. (2018) Synthetic microbe communities provide internal reference standards for metagenome sequencing and analysis. *Nat Commun* 9, 3096. (used with permission from publisher).

Benefits of Sequins for Metagenomics

Normalization both within and between samples, users, equipment and locations	The use of Sequins controls for normalization mitigates the heterogeneity of samples to enable interoperability across users and laboratories.
Workflow monitoring and optimization	Sequins are subjected to the same technical influences and errors as the samples they are combined with, enabling the evaluation of workflow performance.
Absolute quantification	Sequins standards are combined at varying concentrations and the resultant ladder enables quantitative analysis in addition to measuring sensitivity, specificity and limit of detection.

Product ordering information

Product	Catalog Number	Description
Metagenomics Core Control Set	PN-10008	Metagenomics Core Control Set (24 samples*)
Metagenomics Core Control Set	PN-10009	Metagenomics Core Control Set (48 samples*)

based on a 1% spike-in; 250ng library input

Contact Information

Ordering and Support

enquiries@sequins.bio
support@sequins.bio

Key Publications

The utility of Sequins in various microbial metagenomics applications has been demonstrated in several publications, including Hardwick et al. (2018), Wang et al. (2022), Vyshenska et al. (2023), Gunter et al. (2022) and Pallenberg et al. (2022). In these studies, sequins made calculating absolute bacterial abundances in environmental or clinical samples possible while accounting for variability in intermediate processing steps such as customized target enrichment, whole genome amplification and centrifugation. This underscores the breadth of metagenomics applications from environmental surveillance and food safety to agrigenomics, water quality, and health and disease.

Deveson, I., Chen, W., Wong, T. *et al.* Representing genetic variation with synthetic DNA standards. *Nat Methods* 13, 784—791 (2016).

Hardwick SA, Deveson IW, Mercer TR. Reference standards for next-generation sequencing. *Nat Rev Genet*. 2017 Aug;18(8):473-484. doi: 10.1038/nrg.2017.44. Epub 2017 Jun 19. PMID: 28626224.

Hardwick, S.A., Chen, W.Y., Wong, T. *et al.* Synthetic microbe communities provide internal reference standards for metagenome sequencing and analysis. *Nat Commun* 9, 3096 (2018).

Wang, C., Zhang, L., Jiang, X. *et al.* Toward efficient and high-fidelity metagenomic data from sub-nanogram DNA: evaluation of library preparation and decontamination methods. *BMC Biol* 20, 225 (2022).

Gunter, H.M., Youlten, S.E., Madala, B.S. *et al.* Library adaptors with integrated reference controls improve the accuracy and reliability of nanopore sequencing. *Nat Commun* 13, 6437 (2022).

Pallenberg ST, Pust M, Rosenboom I, Hansen G, Wiehlmann L, Dittrich A, Tümmler B, 2022. Impact of Elexacaftor/Tezacaftor/Ivacaftor Therapy on the Cystic Fibrosis Airway Microbial Metagenome. *Microbiol Spectr* 10:e01454-22.

Vyshenska D, Sampara P, Singh K, Tomatsu A, Kauffman WB, Nuccio EE, Blazewicz SJ, Pett-Ridge J, Louie KB, Varghese N, Kellom M, Clum A, Riley R, Roux S, Eloe-Fadrosh EA, Ziels RM, Malmstrom RR. 2023. A standardized quantitative analysis strategy for stable isotope probing metagenomics. *mSystems* 8:e01280-22.