BECKMAN COULTER
Life Sciences

Long-Read Sequencing: A Brief Overview and the Role of Bead-Based DNA Extraction

Introduction to Long-Read Sequencing

Genomic sequencing has come a long way since the Human Genome Project (**https://www.genome.gov/human-genome-project/What**). This was the first large-scale initiative to utilize sequencing to understand how human genetics influenced the susceptibility, development, and progression of disease. Sequencing technology has vastly improved since the 1990s, especially as academic and clinical researchers around the world have rapidly implemented third generation, better known as long-read sequencing.

Long-read sequencing lets researchers characterize high molecular weight single DNA fragments in the 10s to 100s and even 1000s of kilobases with accurate base calling.[1] This approach is applicable to all sample types and species,[1-3] including genomes historically difficult to sequence. While long-read sequencing has only been around for about a decade, it is now ubiquitous in scientific, industrial, and clinical research labs around the world.

Advantages of Long-Read Sequencing

Long-read sequencing is fostering strategic gains in multiple research fields. Oncology and rare disease are two examples that have harnessed the power of long-read sequencing to elucidate previously unknown mechanisms of disease. This is enabling scientists and physician-researchers to postulate and study new genomic medicine approaches. **Table 1** lists a few major advantages of long-read sequencing and how they are applied in the field.

Long-Read Advantage	Areas of Applicability
Discovery of structural variants[4]	Disease etiology
	Polyploidy identification[7]
	Epigenetic characterization[8]
	Cellular and tumor heterogeneity
	HLA mapping[9]
Minimizes amplification biases[5]	Viral pathogen screening[10]
	Native RNA sequencing
Detection of repetitive regions of the genome[6-7]	Reference genome construction
	Facilitates *de novo* assemblies
	Cell line authentication[11]

Table 1. Major advantages and emerging applications for long-read sequencing.

The applications listed above have translational and clinical research consequences. One area of importance that is often overlooked is cell line authentication. Implications include errors in data reproducibility, cellular contamination, and acquisition of genetic anomalies during cell passaging.[11] In drug development, epigenetic characterization helps to create new targets and model mechanisms of action.[12] HLA mapping is critical for patient stratification during cell therapies and informs on possible rejection of organ transplantations.[9] These are just a few of the ways long-read sequencing is helping to advance important research.

Importance of Sample Preparation

Long-read sequencing, like other sequencing technologies, is still limited by the quality of the DNA obtained from the initial sample, therefore extraction remains an important consideration. When determining the best approaches to prepare long-read sequencing samples, a plethora of distinctive characteristics fundamentally impact data quality. Sample attributes like species, collection method, type (e.g., tissue, cell, liquid biopsy), quality and quantity can pose roadblocks when isolating nucleic acids. Another key consideration is the shearing of DNA fragments during the extraction and library construction portion of workflows. High molecular weight genomic DNA (gDNA) capture and retention is pivotal for optimal read length and sequence coverage.[13]

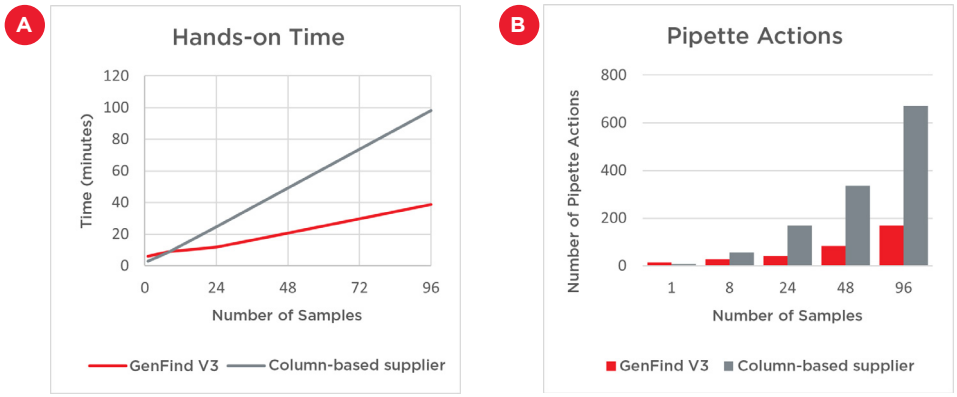
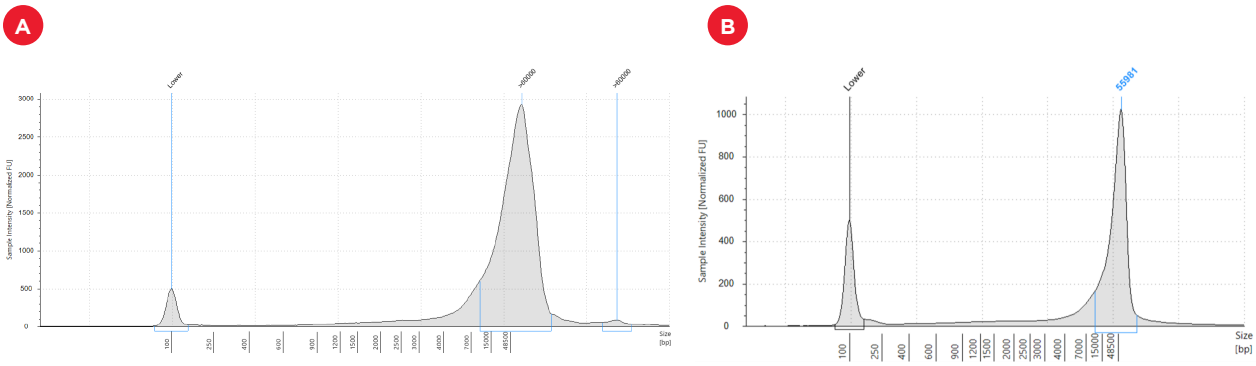
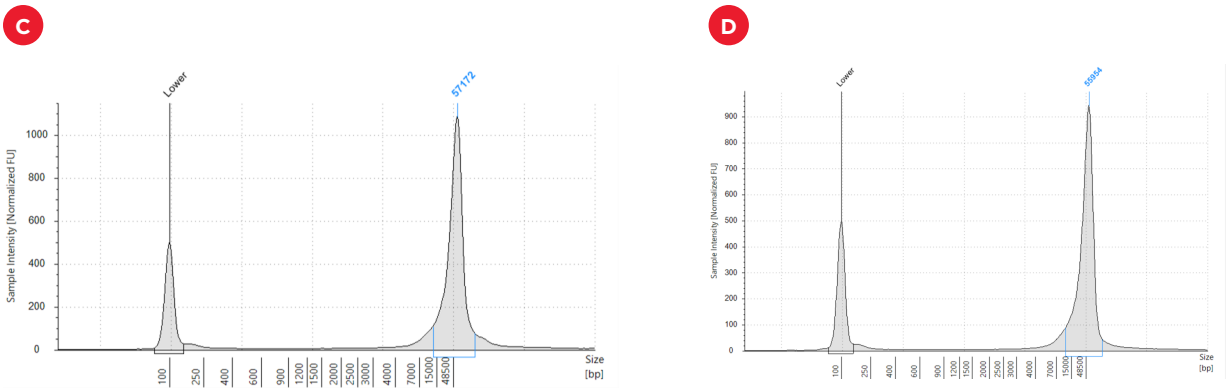
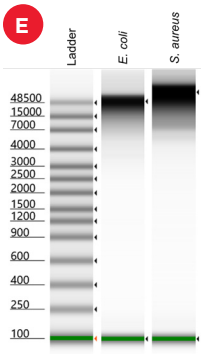


Figure 1. Bead-based extraction requires less hands-on time with fewer pipette actions compared to column-based kits. **(A)** Represents hands-on time to extract gDNA for 1 to 96 samples using GenFind V3. **(B)** The total number of pipette actions required for 1, 8, 24, 48 and 96 samples. Pipette actions include discarding the supernatant and sample dispensing or mixing.

Bead-based and spin columns are two widely used methods for nucleic acid extraction. However, in higher throughput labs, reproducibility necessitates automation integration of workflows. One benefit of automation integration is the reduction of hands-on time (**Figure 1**) for bead-based extraction. Beckman Coulter Life Sciences empowers bead-based workflows with highly accurate and reliable automation that can reduce hands-on time to free up staff for other productive lab activities. **Figure 2** below displays how the bead-based and automation-friendly GenFind V3 kit provides users with an excellent option for extracting high molecular weight gDNA from various species and sample types.







F

Sample	Conc. (ng/µL)	Yield (µg)	DIN	260/280
Jurkat cells	281.5	11.3	9.0	2.0
Saliva	32.4	4.9	8.2	1.9
Fresh whole blood (citrate)	33.7	3.6	9.5	1.8
Fresh whole blood (heparin)	28.5	3.4	9.5	1.8
E. coli	57.3	2.3	8.9	1.9
S. aureus	80.5	3.2	9.6	1.9

Figure 2. GenFind V3 recovers high molecular weight and high-quality DNA from multiple sample types. Representative electropherograms and Agilent Genomic DNA Screen Tape of extracted DNA from **(A)** 1.4 million Jurkat cells, **(B)** Saliva, **(C)** Fresh whole blood in a citrate tube, **(D)** Fresh whole blood in a heparin tube and **(E)** Gram positive (*S. aureus*) and gram negative (*E. coli*) bacteria. Images show the GenFind V3 kit can extract high molecular weight DNA >48.5 kb. Jurkat cells **(A)** show a predominate extracted DNA length of >60 kb which constitutes the upper limit for measurement of gDNA on the TapeStation. **(F)** Table listing sample concentrations, yields, DNA integrity (DIN) and 260/280 ratios.

GenFind V3 – Enabling Long-Read Sequencing Research

As long-read sequencing has gained popularity, the GenFind V3 kit (C34881) has played a role in empowering important discoveries in various fields. Here are three examples of how Beckman Coulter Life Sciences is supporting researchers pushing the limits of long-read sequencing:

- The need for powerful bioinformatics and software algorithms is crucial to further the utility of long-read sequencing, however, *in silico* work must be corroborated and thus necessitates the use of physical samples. In one report by Wick et al., GenFind V3 was able to yield enough DNA to facilitate three different simultaneous sequencing approaches that were crucial to the researchers building a computational assembly pipeline using six bacterial strains as a benchmark.[14]
- Disease prevalence, including the identification of drug-resistant bugs, is a concern for public health officials worldwide. We find GenFind V3 potentiates research of samples obtained from Irish hospitals to accommodate a hybrid long-read sequencing approach. The authors were able to quantify the presence of *Enterococcus* spp. bearing plasmids for linezolid resistant genes.[15]
- Surveillance of food supplies is paramount to every nation and includes monitoring of marine species in coastal environments. One study in Norway was able to document the presence of antibiotic resistant *Klebsiella pneumoniae* in mollusks using GenFind V3 for long-read sequencing.[16] A concluding remark from the team indicates their approach enables foodborne pathogen surveillance of marine reservoirs.

Conclusion

Scientific innovation is rapidly moving technologies from concept to market and long-read sequencing is benefiting from over a decade of success. It continues to carve and simultaneously expand its place in the scientific toolbox as the speed, efficiency and cost of sequencing new genomes becomes more responsive to researcher needs. However, there is still much work left to standardize assays and data analysis algorithms to ensure reproducibility and data integrity. Thus, the need to extract high molecular weight nucleic acids should not become a workflow bottleneck.

Beckman Coulter Life Sciences is supporting researchers tapping into the power of long-read sequencing by providing quality genomic solutions for nucleic acid extraction. The data from samples and cited references above are just a few ways we can empower these workflows. For more information, please visit **Beckman.com**.

References

1. Rhie A, McCarthy SA, Fedrigo O, et al. Towards complete and error-free genome assemblies of all vertebrate species. Nature 2021;592:737–46.
2. Hotaling S, Sproul JS, Heckenhauer J, et al. Long reads are revolutionizing 20 years of insect genome sequencing. Genome Biol Evol 2021;13:8.
3. Mantere T, Kersten S, Hoischen A. Long-read sequencing emerging in medical genetics. Front Genet 2019;10:426.
4. Thibodeau ML, O'Neill K, Dixon K, et al. Improved structural variant interpretation for hereditary cancer susceptibility using long-read sequencing. Genet Med 2020;22:1892–7.
5. Depledge DP, Srinivas KP, Sadaoka T, et al. Direct RNA sequencing on nanopore arrays redefines the transcriptional complexity of a viral pathogen. Nat Commun 2019;10:754.
6. Du H, Liang C. Assembly of chromosome-scale contigs by efficiently resolving repetitive sequences with long reads. Nat Commun 2019;10:5360.
7. Amarasinghe SL, Su S, Dong X, et al. Opportunities and challenges in long-read sequencing data analysis. Genome Biol 2020;21:30.
8. Sakamoto Y, Zaha S, Suzuki Y, et al. Application of long-read sequencing to the detection of structural variants in human cancer genomes. Comput Struc Biotechnol J 2021;19:4207-16.
9. Matern BM, Olieslagers TI, Groeneweg M, et al. Long-read nanopore sequencing validated for human leukocyte antigen class I typing in routine diagnostics. J Mol Diagn 2020;22:912-9.
10. Boldogkői Z, Moldován N, Balázs Z, et al. Long-read sequencing – A powerful tool in viral transcriptome research. Trends Microbiol 2019;27:578-92.
11. Zaaijer S, Gordon A, Speyer D, et al. Rapid re-identification of human samples using portable DNA sequencing. eLife 2019;6:e27798.
12. Ganesan A, Arimondo PB, Rots MG, et al. The timeline of epigenetic drug discovery: from reality to dreams. Clin Epigenet 2019;11:174.
13. Ou S, Liu J, Chougule KM, et al. Effect of sequence depth and length in long-read assembly of the maize inbred NC358. Nat Commun 2020;11:2288.
14. Wick RR, Judd LM, Cerdeira LT, et al. Trycycler: consensus long-read assemblies for bacterial genomes. Genome Biol 2021;22:266.

15. Egan SA, Shore AC, O'Connell B, et al. Linezolid resistance in Enterococcus faecium and Enterococcus faecalis from hospitalized patients in Ireland: high prevalence of the MDR genes optrA and poxtA in isolates with diverse genetic backgrounds. J Antimicrob Chemoth 2020;75:1704-11.

16. Håkonsholm F, Hetland MAK, Svanevik CS, et al. Antibiotic sensitivity screening of Klebsiella spp. and Raoultella spp. isolated from marine bivalve molluscs reveal presence of CTX-M-producing K. pneumoniae. Microorganisms 2020;8:1909.

Beckman Coulter makes no warranties of any kind whatsoever express or implied, with respect to this protocol, including but not limited to warranties of fitness for a particular purpose or merchantability or that the protocol is non-infringing. All warranties are expressly disclaimed. Your use of the method is solely at your own risk, without recourse to Beckman Coulter. Not intended or validated for use in the diagnosis of disease or other conditions. This protocol is for demonstration only and is not validated by Beckman Coulter.

© 2021 Beckman Coulter, Inc. All rights reserved. Beckman Coulter, the stylized logo, and the Beckman Coulter product and service marks mentioned herein are trademarks or registered trademarks of Beckman Coulter, Inc. in the United States and other countries. All other trademarks are the property of their respective owners.

For Beckman Coulter's worldwide office locations and phone numbers, please visit Contact Us at **beckman.com**

21.11.3197