



PacBio Collaborates with Leading Researchers to Establish Long-Read Variant Frequency Consortium

October 18, 2022

The Consortium Will Build a Publicly Available Database as a Resource to Accelerate Insights from Long-Read Human Genome Datasets

MENLO PARK, Calif., Oct. 18, 2022 /PRNewswire/ -- [PacBio](#) (NASDAQ: PACB), a leading developer of high-quality, highly accurate sequencing solutions, today announced the creation of the [Consortium for Long Read Sequencing \(CoLoRS\)](#) that aims to accelerate the utility of long-read human genome datasets. CoLoRS is an open coalition of international researchers focused on creating a comprehensive database of frequency information for all classes of human variation identified using long-read human whole-genome sequencing. High quality long-read data can characterize genetic variation inaccessible to short-read sequencing. As such, CoLoRS plans to critically complement existing databases, help improve the discovery of pathogenic variation, and advance the understanding of the genomic underpinnings of rare disease, where more than half of cases remain unexplained even after short-read genome sequencing.



"PacBio is proud to collaborate with these innovative investigators to build this much needed resource for the genomics research community," said Edd Lee, Director of Human Genomics Segment Marketing at PacBio. "Population frequency is a key tool for interpreting genetic variation. CoLoRS will extend this tool to the variation uniquely detected by HiFi sequencing, particularly structural variants, tandem repeats, and small variants in regions of the genome that are difficult to sequence using other technologies."

The founding members of CoLoRS are leaders from highly respected research hospitals, universities, and laboratories from around the world. Pre-existing datasets provided by consortium members will comprise the initial set of genomes, which will be processed and cataloged using trusted and standardized analysis pipelines. The resulting data will be housed and accessible via National Human Genome Research Institute's (NHGRI) Analysis, Visualization and Informatics Lab-space (AnVIL) which is a cloud-based genomic data sharing and analysis platform. CoLoRS has been awarded supporting funds by the National Institutes of Health Office of Data Science Strategy and NHGRI to help fund cloud-based variant calling and for utilization of the database for NHGRI-funded initiatives such as GREGoR and the *All of Us* Research Program.

"I'm excited to be a part of this consortium of experts in structural variation, genomics, and clinical research to create a database that will enable researchers to realize the full potential of long-read sequencing technology, benefitting their research and the collective understanding of human variation and disease. With this database we will finally be able to consider all types of variation across the entire human genome," said Michael Schatz, Bloomberg Distinguished Professor at Johns Hopkins University.

Recent scientific publications, including those from researchers from the Telomere-to-Telomere consortium, have demonstrated that long-read sequencing can provide unique insights for disease and genome research by covering regions of the genome inaccessible to other technologies. Long-read whole-genome sequencing can detect up to 15,000 more structural variants and 300,000 more small variants, as well as providing significantly higher resolution of tandem repeat regions when compared to short-read sequencing. Structural variants, in particular, account for the majority the base-pair differences between individuals. The CoLoRS database is intended to help researchers by not only providing frequencies of such variants but to also assist future structural variant and tandem repeat genotyping initiatives.

The database is intended to be public, benefiting all researchers, and is expected to be populated with initial data in late 2022. To further expand the power of the database, investigators with raw or summary level HiFi human genome datasets are encouraged to reach out to participate.

For more information, please visit: www.ColorsDb.org.

About PacBio

Pacific Biosciences of California, Inc. (NASDAQ: PACB) is a premier life science technology company that is designing, developing and manufacturing advanced sequencing solutions to help scientists and clinical researchers resolve genetically complex problems. Our products and technology under development stem from two highly differentiated core technologies focused on accuracy, quality and completeness which include our existing HiFi long read sequencing and our emerging SBB™ short read sequencing technologies. Our products address solutions across a broad set of research applications including human germline sequencing, plant and animal sciences, infectious disease and microbiology, oncology, and other emerging applications. For more information, please visit www.pacb.com and follow @PacBio.

PacBio products are provided for Research Use Only. Not for use in diagnostic procedures.

Forward-Looking Statements

This press release may contain "forward-looking statements" within the meaning of Section 21E of the Securities Exchange Act of 1934, as amended, and the U.S. Private Securities Litigation Reform Act of 1995, including statements relating to future availability, uses, accuracy, advantages, quality or performance of, or benefits or expected benefits of using, PacBio products or technologies, including in connection with CoLoRS and its efforts to build

a long-read variant frequency database; the potential of long-read sequencing technology to identify structural and small variants; the increasing prevalence of long-read sequencing utilization in connection with certain areas of research, the means by which the data will be provided, processed, catalogued and available to researchers and the public, the related importance of frequency filtering capabilities, and the anticipated capability of the database to assist in cataloguing and interpreting long-read sequencing data, and other forward-looking statements. Readers are cautioned not to place undue reliance on these forward-looking statements and any such forward-looking statements are qualified in their entirety by reference to the following cautionary statements. All forward-looking statements speak only as of the date of this press release and are based on current expectations and involve a number of assumptions, risks and uncertainties that could cause the actual results to differ materially from such forward-looking statements. Readers are strongly encouraged to read the full cautionary statements contained in PacBio's filings with the Securities and Exchange Commission, including the risks set forth in PacBio's Forms 8-K, 10-K, and 10-Q. PacBio disclaims any obligation to update or revise any forward-looking statements.

Contacts

Investors:


Todd Friedman

ir@pacb.com

Media:

Lizelda Lopez

pr@pacb.com

 View original content to download multimedia: <https://www.prnewswire.com/news-releases/pacbio-collaborates-with-leading-researchers-to-establish-long-read-variant-frequency-consortium-301650342.html>

SOURCE Pacific Biosciences of California, Inc.