



PacBio Announces a New Informatics Analysis Method for Highly Homologous Genes

March 14, 2023

Paraphase Enables Researchers to Analyze Genes That are Critically Relevant but Hard to Genotype

MENLO PARK, Calif., March 14, 2023 /PRNewswire/ -- [PacBio](#) (NASDAQ: PACB), a leading developer of highly accurate sequencing solutions, today announced a new informatics method that genotypes gene paralogs and pseudogenes with high accuracy. The new computational tool, named "Paraphase," enables variant calling, copy number analysis and phasing by identifying the full gene sequence of each of the haplotypes for all genes and pseudogenes of the same gene family. Many medically relevant genes fall into segmental duplications and thus have highly similar gene family members or pseudogenes. The sequence similarity often leads to error prone read alignment and variant calling.



"Through the use of Paraphase, we are able to identify the full sequence of each copy of a gene and, importantly, identify the number of functional and non-functional copies of a gene," said Mike Eberle, Vice President of Computational Biology at PacBio. "This will allow researchers to conduct more accurate carrier analyses and provide a framework for studying the underlying genetics of these complex genomic regions. We believe that applying this method to larger, diverse, population data will ultimately enable researchers to better understand medically important problems, such as silent carriers for spinal muscular atrophy."

Paraphase has been used on several medically relevant genes with highly similar paralogs or pseudogenes, including, *CYP21A2* (21-hydroxylase-deficient congenital adrenal hyperplasia), *TNXB* (Ehlers-Danlos syndrome), *STRC* (hereditary hearing loss and deafness) and *SMN1* and 2 (spinal muscular atrophy). *SMN1* is >99.9 percent similar in sequence to its paralog, *SMN2*, and both genes have variable copy numbers across populations. Mutations in *SMN1* cause spinal muscular atrophy (SMA), a leading cause of early infant death.

High throughput detailed sequence analysis of complete genes is challenging using existing technologies, and identifying silent carriers (having two copies of *SMN1* on one chromosome and zero copies on the other, accounting for 27 percent of carriers in African populations) is impossible without pedigree information. In a recent [peer-reviewed publication](#), Paraphase was able to detect these pathogenic variants for SMA. The study also identified major *SMN1* and *SMN2* sequence haplogroups and characterized their co-segregation through pedigree-based analyses. In addition, the authors identified a pair of haplotypes that can serve as a genetic marker for alleles carrying two copies of *SMN1* in African populations, demonstrating the potential of haplotype-based screening of silent carriers.

"The fact that HiFi sequencing not only allows access to the most difficult regions of the human genome, but also enables calling of all known variant types, such as SNVs, Indels, and SVs including CNVs, plus phasing of these loci, keeps a great promise for applications in rare disease research," said Alexander Hoischen, Ph.D., Associate Professor for Genomic Technologies and Immuno-Genomics at Radboud University Medical Centre.

Paraphase is being extended into a genome-wide generalized paralog caller as more highly homologous genes are included. Paraphase works on whole-genome sequencing and hybrid capture-based enrichment data. It can also be adapted to work with amplicon sequencing data, when the full regions of interest are captured or amplified.

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. All statements other than statements of historical fact are forward-looking statements, including statements relating to the uses, accuracy, coverage, advantages, quality or performance of, or benefits or expected benefits of using, PacBio products or technologies, including Paraphase, and other future events. 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, including, challenges inherent in developing, manufacturing, launching, marketing and selling new products; potential product performance and quality issues; rapidly changing

technologies and extensive competition in genomic sequencing; customers and prospective customers curtailing or suspending activities utilizing PacBio's products; and third-party claims alleging infringement of patents and proprietary rights or seeking to invalidate PacBio's patents or proprietary rights. 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@pacificbiosciences.com

Media:

Lizelda Lopez

pr@pacificbiosciences.com

 View original content to download multimedia: <https://www.prnewswire.com/news-releases/pacbio-announces-a-new-informatics-analysis-method-for-highly-homologous-genes-301770858.html>

SOURCE Pacific Biosciences of California, Inc.