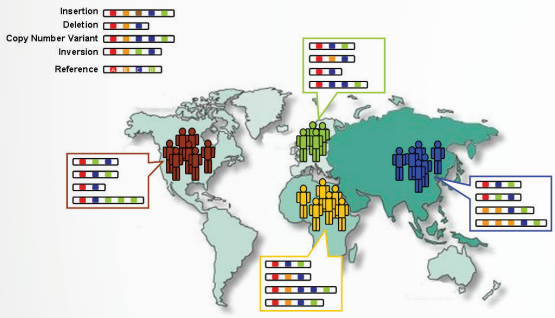
The Most Comprehensive View of the Human Genome

To understand the genetic factors underlying disease and address missing heritability, researchers require a more comprehensive understanding of all the variation in the human genome. Single Molecule, Real-Time (SMRT®) Sequencing delivers the read lengths, unbiased coverage and accuracy needed for accessing the complete size spectrum of sequence variant types, from single nucleotides to complex structural variants. PacBio's long single-molecule reads also provide direct variant phasing information across full-length genes and chromosome haplotype blocks. With SMRT Sequencing, you can now access novel variant types and regions of the human genome that were previously inaccessible, and gain new insight into the genetic basis of disease.

Create Gold-Standard Population References

- Increase power by matching your reference to the genetic background of your study population
- Access novel types of genetic variation and difficult to characterize regions
- Improve variant calling with a more comprehensive reference
- Cost-effectively validate novel variants with a highly accurate orthogonal platform



Genetic diversity varies both within and between populations[1]

"To get a medical-grade genome...we need to have the most accurate and complete genome for each individual. We believe that the PacBio SMRT machines will help us reach this goal."

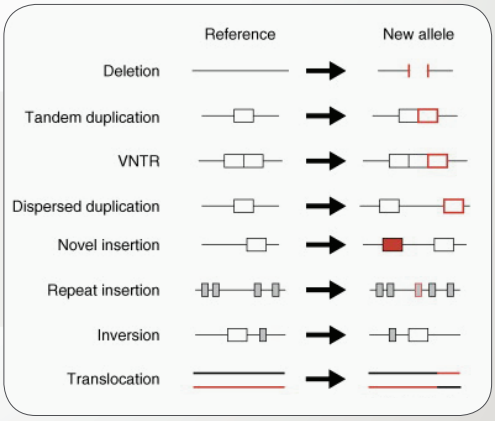
— Craig S. Venter, HLI Co-Founder[1]

Resolve Structural Variation

- Uncover the missing heritability linked to structural variation
- Identify breakpoints at the sequence level to unravel the genetic etiology of disease
- See the unbiased range of structural variation of all sizes, types, and GC content in the genome

"We now have access to a whole new realm of genetic variation that was opaque to us before."

— Professor Evan Eichler, University of Washington[2]

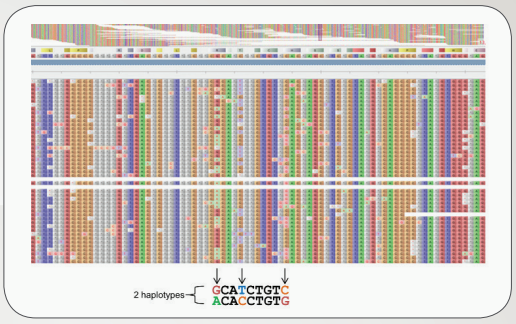


"Non SNP variants ranging from small indels to large CNVs and inversions accounted for 74% of the total number of variant bases".[3,4]

Characterize Complex Regions Underlying Genetic Disease

"We basically can see the entire picture. We're not looking under a lamppost for the keys. It's daylight, and we can see the whole neighborhood. So we're gonna find the keys."

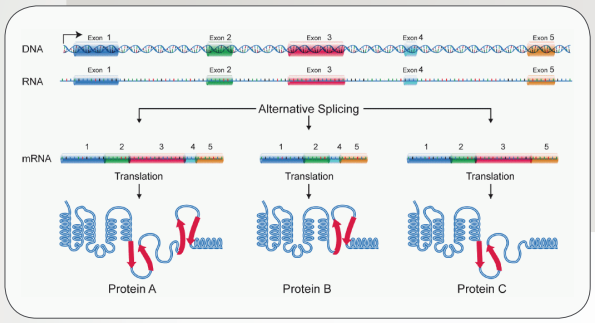
— Dan Geraghty, Fred Hutchinson Cancer Center[5]



Resolution of two allelic copies of the MUC5AC gene[8]

- Sequence previously unsequencable loci associated with genetic disease
- Accurately and definitively phase polymorphisms over entire genes, such as HLA
- Multiplex across many samples to scale your research cost-effectively

Profile the Complexity of the Transcriptome with Iso-Seq™ Sequencing



Alternative splicing can have important implications for protein structure and function[9].

- Discover novel genes and gene isoforms
- Directly sequence full-length transcripts to eliminate the need for transcript reconstruction
- Differentiate isoform expression between cells, tissues, and disease states

"The power of long-read sequencing is really to be able to capture all of the information in its intact form without trying to solve a jigsaw puzzle that you may have put together wrong."

— Mike Snyder, Stanford[6]

Key References

1. Venter, C. (2014) Human Longevity Inc. Adds PacBio(R) Sequencing to Enable Additional Insight Into Human Genetic Variation [Press release]. Retrieved January 16, 2015 from http://investor.pacificbiosciences.com/releasedetail.cfm?ReleaseID=874612
2. University of Washington Health Sciences/UW Medicine. (2014, November 10). Thousands of never-before-seen human genome variations uncovered. *ScienceDaily*. Retrieved January 16, 2015 from www.sciencedaily.com/releases/2014/11/141110124233.htm
3. Hurles, ME, Dermitzakes, ET & Tyler-Smith, C. (2008) The functional impact of structural variation in humans. *Trends in Genetics*. **24** (5), 238-245.
4. Gonzaga-Jauregui, C, Lupski, JR, Gibbs, RA. (2012) Human genome sequencing in health and disease. *Annu Rev Med.* **63**, 35-61.
5. Geraghty, D. (2014) *Major Sequencing Projects Should Be Done with Long Reads, Says Dan Geraghty.* Retrieved January 16, 2015 from http://mendelspod.com/podcast/major-sequencing-projects-should-be-done-long-reads-says-dan-geraghty
6. Snyder, M. (2014) The Rise of Long Reads: Mendelspod Podcast Series. Retrieved January 16, 2015 from http://blog.pacificbiosciences.com/2014/09/the-rise-of-long-reads-mendelspod.html
7. Jthiele. 1000 Genomes Project. (2008) http://en.wikipedia.org/wiki/1000_Genomes_Project#mediaviewer/File:Genetic_Variation.jpg
8. Guo X et al. (2014) Genome reference and sequence variation in the large repetitive central exon of human MUC5AC. *Am J Respir Cell Mol Biol*. **50** (1), 223-32.
9. National Human Genome Research Institute. Bioinformatics: Finding genes. (2013) http://www.genome.gov/25020001
10. Chaisson, MJP, et. al. (2014) Resolving the complexity of the human genome using single-molecule sequencing. *Nature*. doi:10.1038/nature13907.
11. Au, KF et al. (2013) Characterization of the human ESC transcriptome by hybrid sequencing. *PNAS* **110** (50), E4821-E4830.
12. Berlin, K et.al. (2014) Assembling large genomes with single-molecule sequencing and locality sensitive hashing. *BioRxiv*. doi: http://dx.doi.org/10.1101/008003.
13. Loomis, EW et al. (2013) Sequencing the unsequenceable: Expanded CGG-repeat alleles of the fragile X gene. *Genome Res*. **23**(1), 121-128.
14. Mayor, NP et al. (2014) OR57:Generation of 252 HLA Class I genomic sequences in a single sequencing reaction using DNA barcodes and single molecule real-time (SMRT) DNA sequencing technology. *Human Immunology*. **75** (Suppl), 49. doi:10.1016/j.humimm.2014.08.060.

PACIFIC BIOSCIENCES®

PN: VM101-012015