

Longest Reads and Maximum Output: GENEWIZ's Best Practices for PacBio Sequel Sequencing

Dave Corney, PhD

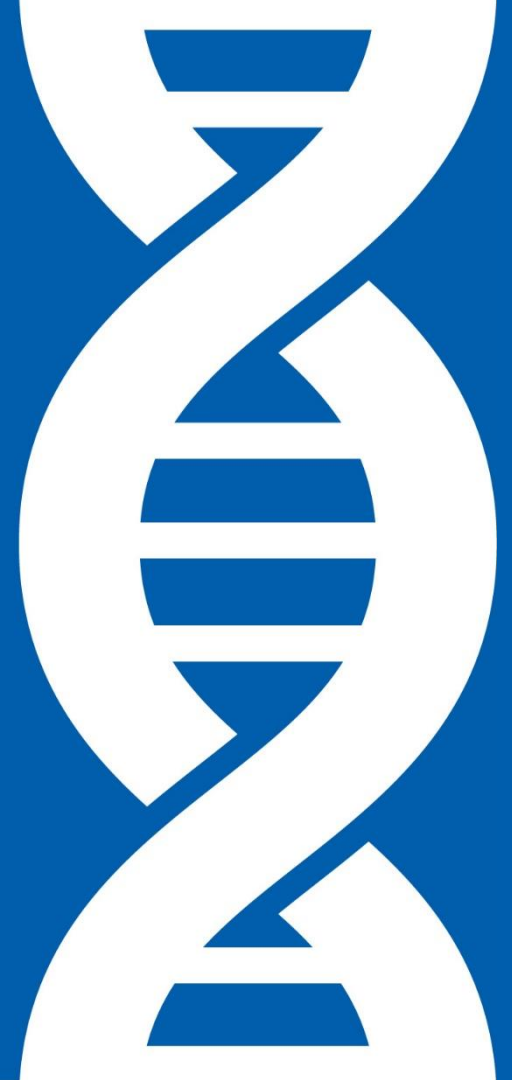
Senior Scientist

Next Generation Sequencing, GENEWIZ



GENEWIZ[®]

Solid science. Superior service.

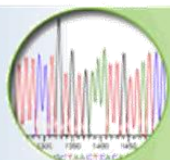


GENEWIZ – a full-service sequencing provider



Gene Synthesis

- Gene Synthesis Services
- TurboGENE™
- US-based synthesis



DNA Sequencing

- DNA Sequencing Services
- Difficult Template Sequencing
- Sequencing from Bacteria and Phage
- Primer Walking



Molecular Biology

- DNA Cloning
- Construction of shRNA Plasmids
- Plasmid DNA Preparation
- Site-Directed Mutagenesis



Next Generation Sequencing

- DNA Sequencing (whole genome, exome, targeted)
- RNA-Seq (small/micro, long non-coding, mRNA)
- ChIP-Seq
- Metagenomics
- Bioinformatics Solutions



Genomic Services

- Mutation Analysis
- SNP Discovery Sequencing
- Fragment Analysis
- Transgenic Mouse Genotyping



GLP Regulatory-Compliant Services

- Confirmatory DNA Sequencing
- SNP/Mutation analysis
- Validation and clinical sample processing
- Plasmid Preparation
- Small Genome and Amplicon NGS



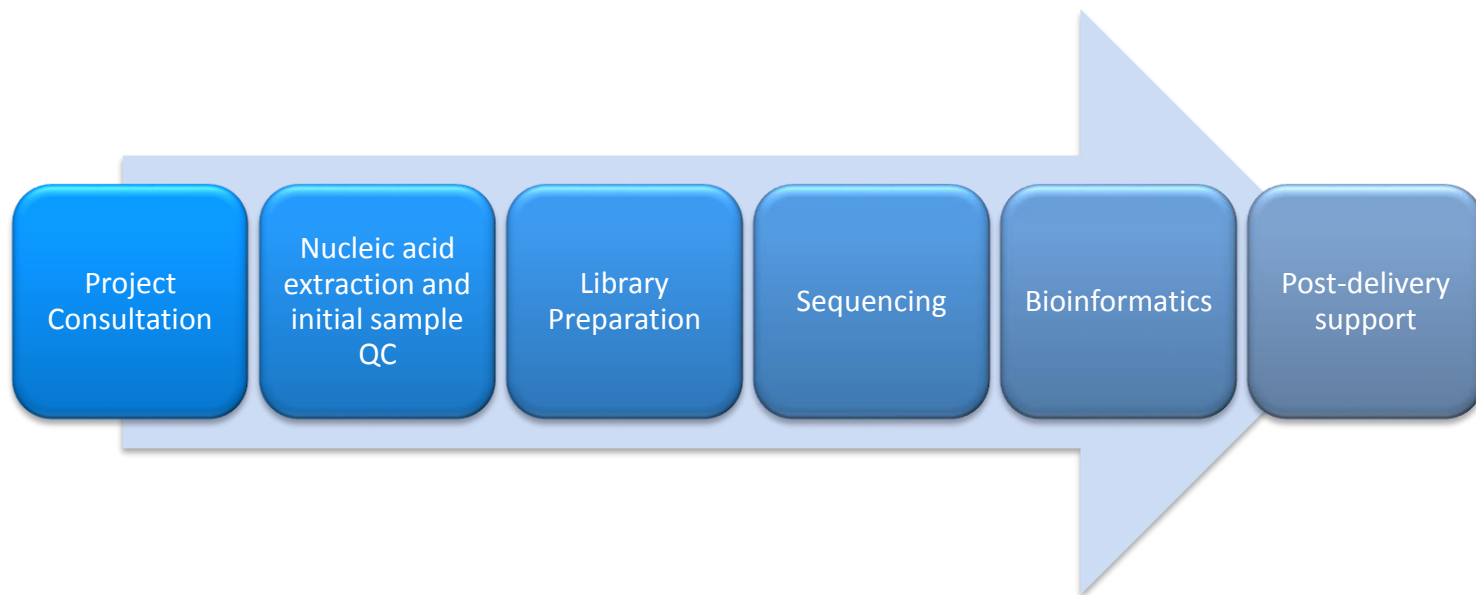
illumina®



 PACBIO®



GENEWIZ's sample-to-result workflow

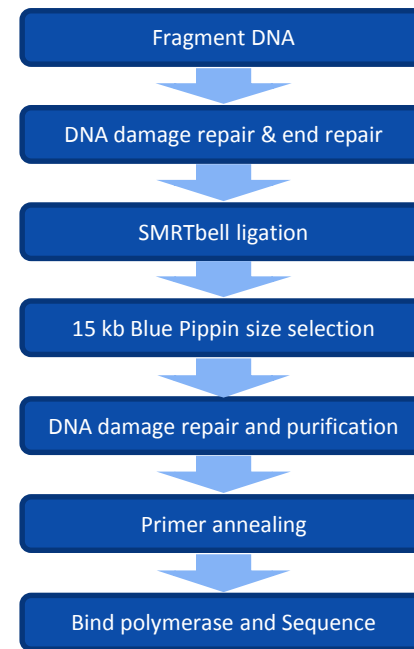
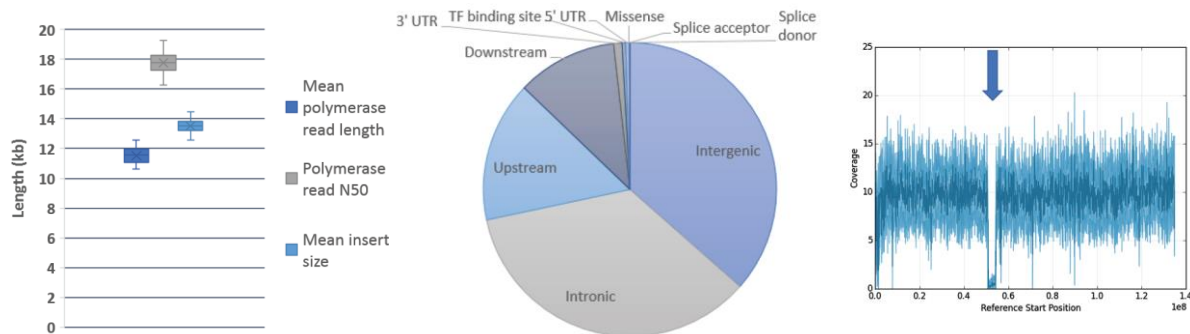


Importance of comprehensive project consultation

- Many scientists unfamiliar with PacBio sequencing
 - Smaller number of scientists are incredibly well-educated
- Expectations may not be aligned, or correct, e.g.:
 - Common misconception is a high error rate
 - Good enough DNA for Illumina likely not suitable for PacBio
 - Handle raw BAM files like an Illumina aligned BAM file (“just give me the fastq file”)
- In-depth customer consultation required
 - Project goals
 - Benefits of PacBio, and other GENEWIZ capabilities and platforms
 - Educate customers about DNA quality and handling
 - Bioinformatic approaches

Production-scale WGS for SV detection

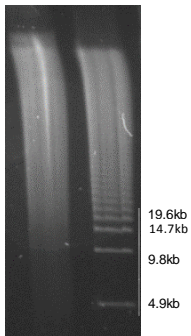
- Pharma client interested in detecting SVs in glioblastoma
- Provided high-quality HMW genomic DNA
- Prepared 20 kb libraries with 15 kb Blue Pippin size selection
- Sequence to ~9x using 6 hr movies and v1.2.1 chemistry



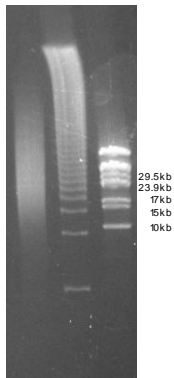
Importance of sample quality

- Customer wanted to sequence 2.5 Gb genome to 20x
- Provided with partially degraded which failed library prep
- Solution: customer submitted cell pellets for extraction by GENEWIZ

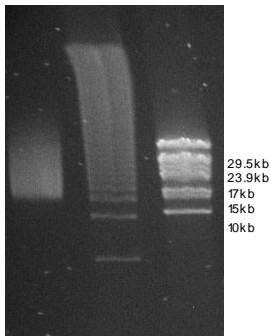
Extracted sample



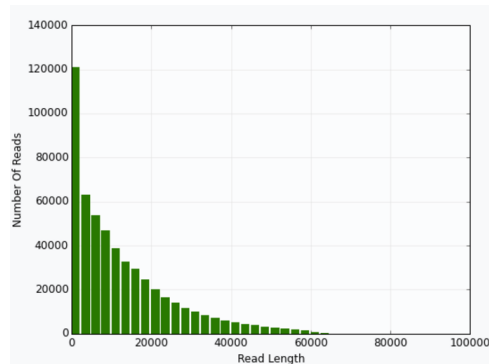
Sheared DNA



Final 20 kb library



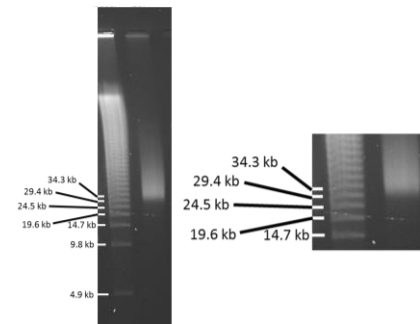
Polymerase read length



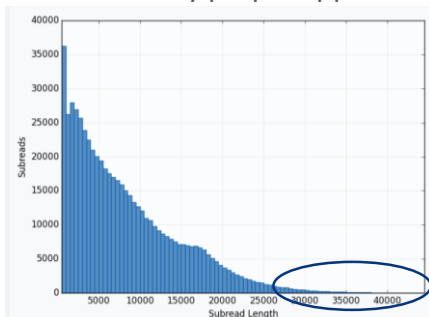
Analysis Metric	Value
Data output	6.86 Gb
Polymerase reads	532,023
Polymerase RL (mean)	12,887 bp
Polymerase RL N50	22,750 bp
Subread length (mean)	9,834 bp

Pushing the boundaries: >30 kb libraries

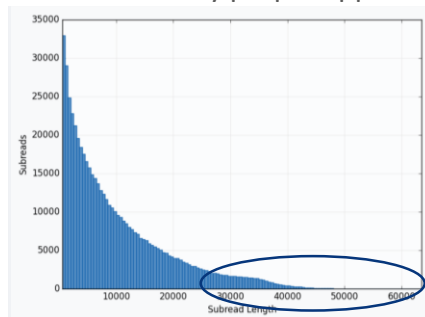
- 50 Mb organism with 6x5 kb transgene cassette
- Customer was only able to provide 10 µg DNA
- Size of gDNA 15-50 kb
- Combined gentle shear and no-shear to prepare library



20 kb library prep mapped



30 kb library prep mapped



Metric	Value
Coverage	123x
Polymerase RL mean	10,996 bp
Subread mean	10,357 bp
Max mapped polymerase RL	80,432 bp
Max mapped subread	63,000 bp
Contig N50	4.77 Mb
Max contig	7.31 Mb

Conclusions

- Work closely with clients during planning stages to ensure expectations are aligned
 - Client education
 - Understand project goals and be sure experiment will address them
 - Coordinate with bioinformatics team
- Transparent communication throughout project
- Work closely with your PacBio FAS

Solid science. Superior service.