Enabling CNV Studies from Single Cells Using Whole Genome Amplification and Low Pass Sequencing

Ioanna Andreou, Silke Huebner, Christin Meerschiff, Annika Piotrowski, Courtney Nadeau, Nan Fang
QIAGEN Strasse 1, 40724 Hilden, Germany

Introduction

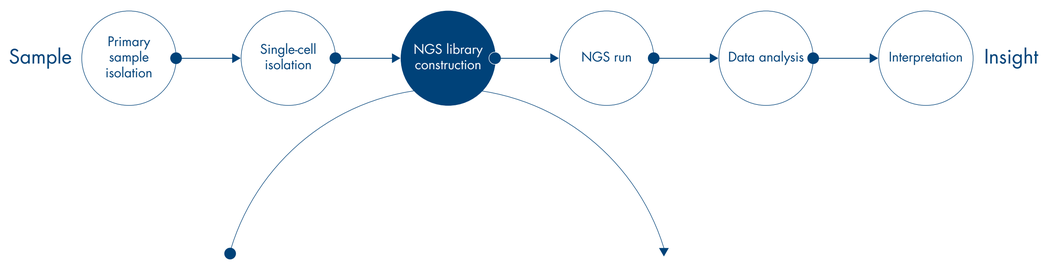
DNA copy number variations (CNVs) play an important role in the pathogenesis and progression of cancer and confer susceptibility to a variety of human disorders. Array comparative genomic hybridization (aCGH) has been used widely to identify CNVs genome-wide, but next generation sequencing (NGS) provides an opportunity to characterize CNVs genome-wide with unprecedented resolution.

We present a workflow using the newly developed QIAseq FX Single Cell DNA Library Kit and sequencing at low depth to enable the detection of both aneuploidy and sub-chromosomal copy number variations.

The QIAseq FX Single Cell DNA Library Kit generated libraries from single cells and low amounts of DNA in less than 4 hours. The kit applies an optimized protocol using QIAGEN's unique multiple displacement amplification (MDA) technology to amplify gDNA from single cells. Amplified DNA is subsequently fragmented using QIAseq FX technology. This incorporates enzymatic DNA fragmentation into a streamlined, optimized protocol that does not require sample cleanup or fragment quantification between fragmentation and adapter ligation, saving time and reducing material loss.

QIAseq FX Single Cell DNA Library Kit

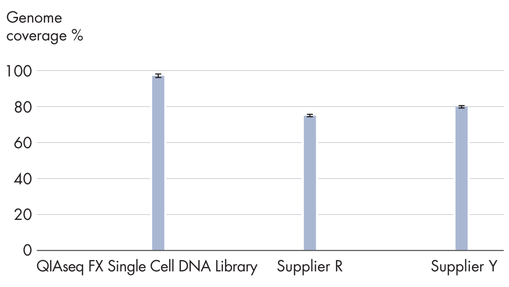
Complete cell-to-library solution



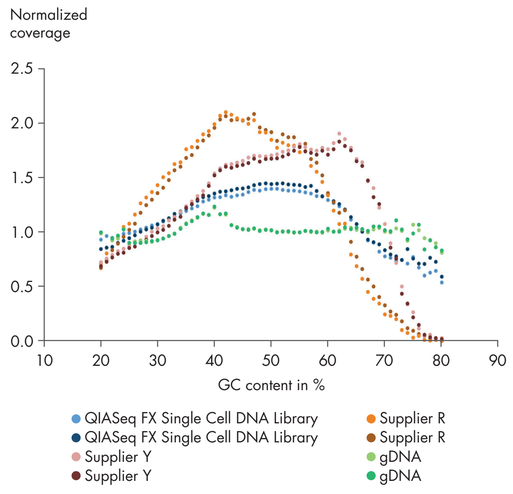
- Single eukaryotic cell
- Single bacterial cell
- Picogram levels of purified DNA

- Whole genome NGS library
  - Illumina-compatible
  - Sequence variants
  - Structural variants
  - Aneuploidy
  - Bacterial genomes

High and Even Genomic Coverage



High genome coverage. Libraries generated from single peripheral blood mononuclear cells (PBMC) using the QIAseq FX DNA Library Kit or kits from two other suppliers and sequenced at low depth using a MiSeq® (Illumina®). Data were analyzed according to Zhang, CZ., et al. (2015) *Calibrating genomic and allelic coverage bias in single-cell sequencing.* Nat. Commun. **6**, 6822.



More uniform coverage, even with GC-rich regions. Libraries were generated from PBMCs using the QIAseq FX DNA Library Kit or kits from two other suppliers and sequenced at low depth using a MiSeq. Data were analyzed using the QIAGEN® CLC Genomics Workbench (v8.5.1).

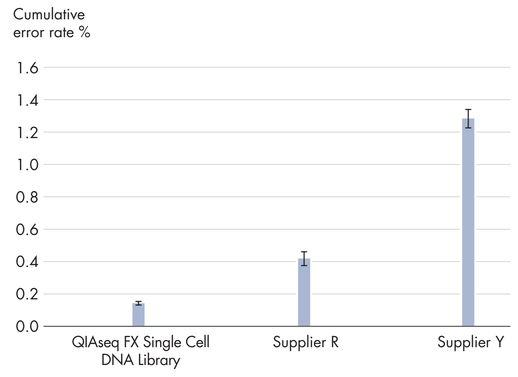
PCR-based kits from other suppliers introduce GC bias and lead to underrepresentation of GC-rich regions. This can be extremely important since features of interest may be located in GC-rich stretches, or researchers may be working with bacterial samples with GC-rich genomes.

High Fidelity and Low Error Rate

Single-cell libraries were prepared from isolated PBMCs using the QIAseq FX Single Cell DNA Library Kit or kits from two other suppliers and sequenced using a MiSeq.

Reads were mapped to the human genome (hg19) and sequence mismatches between NGS data and the reference were computed. All analysis was performed using the QIAGEN CLC Genomics Workbench (v8.5.1). Data plotted are the mean proportion of sequence differences ± standard deviation for three individual libraries prepared with each kit.

Higher numbers mean a greater number of positions where the sequence in the dataset does not match the reference genome. Some of these sites will be normal polymorphisms or mutations, but many of the mismatches with either of the kits from other suppliers will represent false positives introduced during library preparation. These mismatches can increase background when calling variants, and can be identified as false positives in some cases.

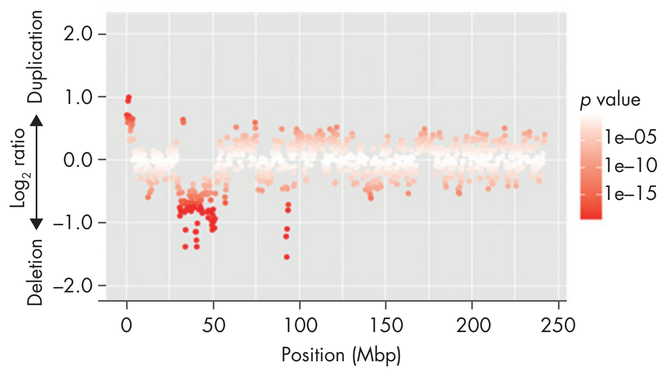


Sequence error rates of several single-cell NGS methods.

Detection of Sub-Chromosomal Copy Number Variations

Single-cell libraries from PBMCs and Jurkat cells were prepared using the QIAseq FX Single Cell DNA Library Kit and were sequenced to 0.1x depth using a MiSeq.

Reads were mapped to the human genome (GRCh38) and the copy number variation of Jurkat versus PBMCs (control diploid cells) was assessed using the methods of Xie, C., et al. (2009) *CNV-seq, a new method to detect copy number variation using high-throughput sequencing*. BMC Bioinformatics.**10**, 80. The $\log_2$ ratio(Jurkat/PBMC) of coverage using a window size of 500 kb for chromosome 2 from a cell with an approximately 25 Mbp deletion.



Detection of small copy-number variations.

Conclusions

We have presented a complete single-cell-to-library solution that delivers whole genome NGS Libraries in a streamlined and optimized workflow.

The QIAseq FX Single Cell DNA Library Kit delivers:

- Maximum genome coverage.
- Greater sequence fidelity to reduce background and false positives.
- Completely PCR-free cell-to-library protocol to minimize bias.
- Libraries from single cells in under 4 hours, using a streamlined protocol.
- Compatibility with both eukaryotic and bacterial cells.
- NGS libraries and amplified gDNA samples that can be archived for follow-up experiments or secondary analyses.

The QIAseq FX Single Cell DNA Library Kit enables the detection of both aneuploidy and sub-chromosomal copy number variations, regardless of their position in the genome. It provides maximum sequence fidelity, minimizing false positives when analyzing sequence variants from the same dataset.

For up-to-date licensing information and product-specific disclaimers, see the respective QIAGEN kit handbook or user manual. QIAGEN kit handbooks and user manuals are available at **www.qiagen.com** or can be requested from QIAGEN Technical Services or your local distributor.

1104402 09/2016

Sample to Insight