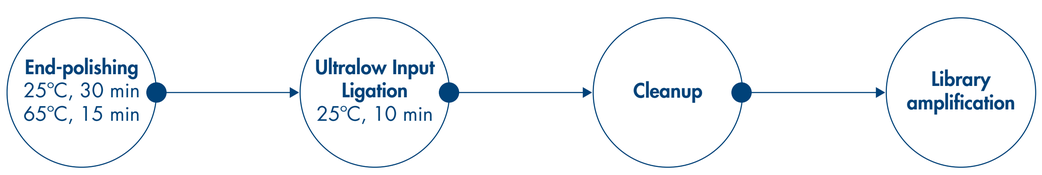
Sensitive and Reliable Variant Detection From Challenging Samples

Katja Heitz, Rumeysa Akinci-Tolun, Jennifer Fostel, Anika Joecker and Nan Fang
QIAGEN GmbH, QIAGEN Strasse 1, 40724 Hilden, Germany

Abstract

The rapidly developing next-generation sequencing (NGS) technologies provide highly sensitive methods in detecting and characterizing variants in clinical samples. However, clinical samples are often limited in quantity, as well as compromised in quality. Such samples are not suitable for standard NGS library construction methods, which commonly require hundreds of nanograms of good-quality DNA. Examples of such challenging clinical samples include laser capture microdissection (LCM) samples, formalin-fixed paraffin-embedded (FFPE) samples and circulating DNA.

Here, we describe an optimized workflow to efficiently convert small amounts of DNA samples into sequencing libraries. The library construction protocol is based on the QIAseq Ultralow Input Library Kit, which has optimized enzyme and buffer formulations that enable high library conversion rate, as well as unbiased library amplification from as little as 10 pg input DNA. In combination with an optimized hybrid-capture–based target enrichment procedure, the workflow described in this poster enables reliable variant detection even with low DNA input amount, enabling sequence insights from challenging sample types.

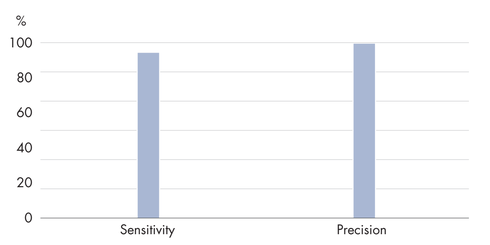


The QIAseq Ultralow Input Library Kit workflow.

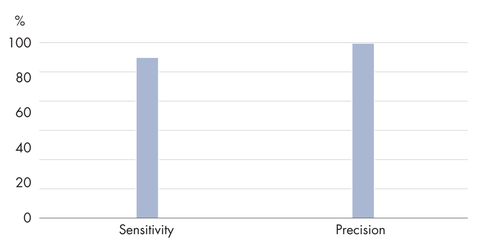
Variant Detection from 1 ng gDNA in WGS and WES

High specificity and sensitivity

The variants in the GIAB samples can be detected with high specificity and sensitivity from 1 ng input DNA using either whole genome sequencing (WGS) or whole exome sequencing (WES) methods. Fragmented GIAB gDNA (1 ng) was constructed into an Illumina sequencing library using the QIAseq Ultralow Input Library Kit. The sequencing libraries were either sequenced directly, or went through the hybrid capture target enrichment procedure, so that only the enriched exons are sequenced. The high-confidence SNPs in the GIAB samples were detected with high sensitivity (>90%) and high precision (>99.50%; or <0.50% false-positive rate) in both WGS and WES experiments.



High SNP calling concordance in WGS with as little as 1 ng GIAB gDNA sample. The genome analysis toolkit (GATK) analysis pipeline was used for variant calling of the GIAB sample that were sequenced on a HiSeq 4000 (Illumina) to an average coverage of 25X. With 1 ng gDNA input, 94.66% of the characterized high-confidence SNPs in the GIAB sample were detected with 99.70% precision (0.30% false-positive rate).

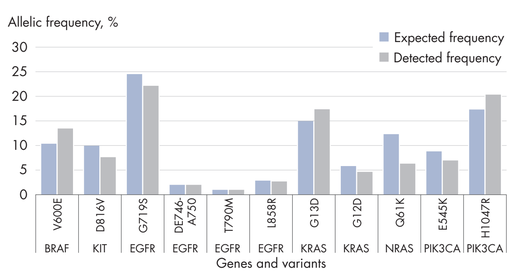


High SNP calling concordance in WES with as little as 1 ng GIAB gDNA sample. The sequencing library was constructed with 1 ng fragmented GIAB gDNA using the QIAseq Ultralow Input Library Kit. Following library construction, exome enrichment was performed with xGen® Exome Research Panel v1.0 (IDT). Genome Analysis Toolkit (GATK) analysis pipeline was used for variant calling of the GIAB WES sample that was sequenced on NextSeq 500 (Illumina) to an average coverage of 86X. With 1 ng initial gDNA input, 91.26% of the characterized high confident SNPs in the GIAB sample were detected with 99.62% precision (0.38% false-positive rate).

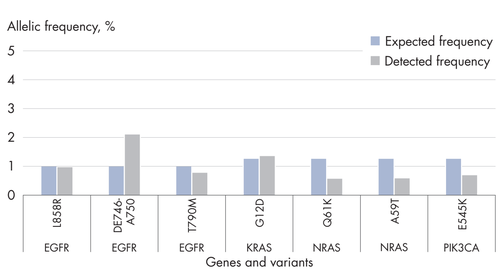
Sensitive Detection of Low-Frequency Variants

Targeted sequencing workflow for sensitive mutation detection

The combination of QIAseq Ultralow Input library prep and xGen Pan-Cancer Panel enables sensitive mutation detection down to 1% allelic frequency with even moderate sequencing coverage. With the FFPE reference sample, all characterized mutations with allelic frequencies from 1–24.5% were accurately detected. With the cfDNA reference sample (harboring 8 characterized mutations at around 1% of allelic frequency), 7 of the 8 mutations were detected with a moderate average coverage of ~500 for the target regions. Sequencing data were analyzed using the CLC Biomedical Workbench (QIAGEN).



Sensitive mutation detection from FFPE reference. All 11 characterized mutations in the FFPE reference samples were accurately detected with the targeted sequencing workflow described above.



Sensitive mutation detection from cfDNA reference. With high adaptor ligation efficiency and target capture uniformity, even the 8 mutations at ~1% allelic frequency were accurately detected with the targeted sequencing workflow described above, with a moderate average coverage of 500.

Sensitive Detection of Low-Frequency Variants

Targeted sequencing workflow for sensitive mutation detection

The combination of QIAseq Ultralow Input library prep and xGen Pan-Cancer Panel enables sensitive mutation detection down to 1% allelic frequency with even moderate sequencing coverage. With the FFPE reference sample, all characterized mutations with allelic frequencies from 1–24.5% were accurately detected. With the cfDNA reference sample (harboring 8 characterized mutations at around 1% of allelic frequency), 7 of the 8 mutations were detected with a moderate average coverage of ~500 for the target regions. Sequencing data were analyzed using the CLC Biomedical Workbench (QIAGEN).

Sensitive mutation detection from FFPE reference. All 11 characterized mutations in the FFPE reference samples were accurately detected with the targeted sequencing workflow described above.

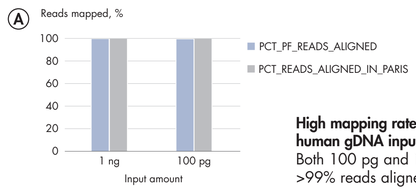
Sensitive mutation detection from cfDNA reference. With high adaptor ligation efficiency and target capture uniformity, even the 8 mutations at ~1% allelic frequency were accurately detected with the targeted sequencing workflow described above, with a moderate average coverage of 500.

1103846 08/2016

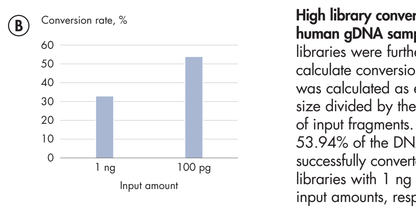
QIAseq Ultralow Input Library Kit For Minimal Input

High mapping rate, conversion rate and uniformity

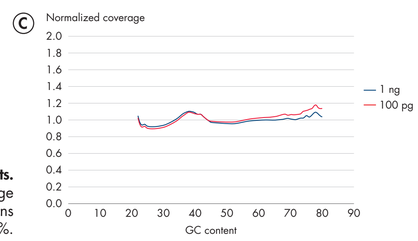
100 pg or 1 ng of fragmented genome-in-a-bottle (GIAB, RM8398; NIST) human reference DNA was constructed into an NGS sequencing library using the QIAseq Ultralow Input Library Kit (QIAGEN) and sequenced on a HiSeq™ 4000 (Illumina) to an average coverage of 23X and 25X, respectively. PicardTools was used to analyze sequencing data and determine library quality and complexity. The sequencing libraries constructed demonstrate almost 100% of reads mapped to the reference, high conversion rate and even coverage of genome regions with different GC contents.



High mapping rate from minimal human gDNA input amount. Both 100 pg and 1 ng samples have >99% reads aligned to the reference.



High library conversion rate with human gDNA sample. The same libraries were further analyzed to calculate conversion rate. This was calculated as estimated library size divided by the total number of input fragments. 33.16% and 53.94% of the DNA fragments were successfully converted to sequencing libraries with 1 ng and 100 pg input amounts, respectively.



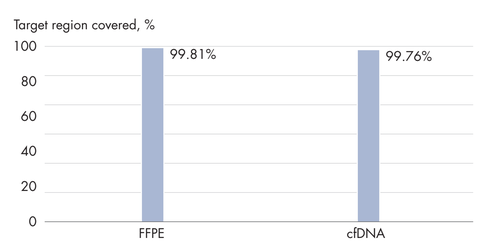
Even coverage of human genomes with different GC contents. The sequencing coverage normalized to the mean coverage ('Normalized Coverage') was calculated for genome regions with GC contents from 20–80%.

QIAseq Ultralow Input Library for Targeted Sequencing

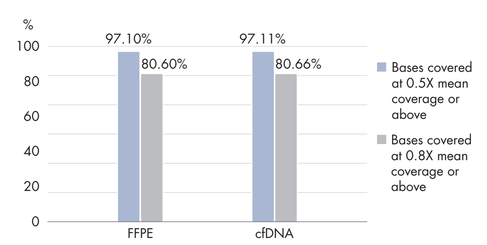
High coverage and uniformity of libraries constructed from FFPE or cell-free DNA samples

A sequencing library constructed using the QIAseq Ultralow Input Library Kit delivers comprehensive and even coverage when used in combination with hybridization-based target enrichment; essential for sensitive, reliable mutation detection.

Sequencing library construction coupled with hybrid-capture–based target enrichment is commonly used to detect mutations in clinically relevant samples such as FFPE or cell-free DNA (cfDNA). We tested this workflow using Quantitative Multiplex Formalin Compromised DNA Standard I or Multiplex I cfDNA Reference Standard Set (both from Horizon Discovery). FFPE or cfDNA DNA standard (20 ng) was constructed into sequencing libraries and the libraries were enriched for the 127 significantly mutated genes (SMGs) using the xGen Pan-Cancer Panel (IDT) and sequenced using a miSeq® (Illumina) sequencer. Data were analyzed using CLC Biomedical Workbench (QIAGEN).



Comprehensive target coverage. Almost 100% of the cancer gene regions targeted by the xGen Pan-Cancer Panel were detected in the sequencing, demonstrating highly efficient library construction and hybrid capture even with limited (20 ng) FFPE or cfDNA samples.



Excellent coverage uniformity. The target-enriched FFPE and cfDNA libraries show high target uniformity, with >97% of bases covered at 0.5X mean coverage or above, and >80% of bases covered at 0.8X mean coverage or above.

Summary: Optimized QIAseq Ultralow Input Library Protocol For Sensitive and Reliable Variant Detection

- Novel QIAseq Ultralow Input Library chemistries enable:
  - Efficient adaptor ligation
  - High library conversion rate
  - Superior coverage uniformity regardless of GC content
  - High-quality sequencing libraries with sub-nanogram input
- For variant detection with WGS and WES:
  - High sensitivity and specificity from only 1 ng of input DNA
- For targeted sequencing based on hybrid capture:
  - High sequence uniformity ensures comprehensive and even coverage
  - Sensitive mutation detection of low frequency variants even with moderate coverage

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.

Sample to Insight