Application Note

QIAscout® allows detection of subpopulation-specific variants using whole genome amplification and low-pass sequencing, followed by census-based variant identification

Ruth Kläver, Andrea Janosch, Charline Bemmann, Claudia Kappmeier, Dorothee Unt, Frank Reinecke, Divya Vijay Pratheek, Norbert Hochstein

QIAGEN GmbH, QIAGEN Strasse 1, 40724 Hilden, Germany

This application note demonstrates that single cells isolated using the QIAscout are suitable for whole genome amplification and library preparation using the QIAseq™ FX Single Cell DNA Library Kit, followed by low-pass sequencing, allowing the identification of subpopulation-specific variants. Furthermore, it shows how the QIAscout instrument enables targeted isolation of single cells from a large, mixed cell population.

Introduction

While bulk cell analysis is critical for understanding a biological system as a whole, it is an under-representation of the inherent biological complexity. Bulk analysis leads to "cellular averages," masking minor genomic differences across individual cell subpopulations. Single-cell analysis is capable of bringing into focus the individual contribution of every cell, without obscuring a biological response that may otherwise occur when cells are assessed in bulk. Although single-cell research is currently gaining momentum, the lack of affordable methods to precisely isolate a single cell from a heterogeneous cell population without manipulating the cellular microenvironment remains a pervasive challenge.

The need to dissect cellular heterogeneity at single-cell resolution is fundamental to many aspects of biological research. Specific selection and isolation of single cells from a mixed cell population form the basis of genetic screens or generation of stable transformants and cell lines. Due to the biological heterogeneity of a cell population, tools must be developed and optimized to detect cell type-specific variants and to reliably identify isolated single cells based on these specific variants.

In an initial experiment, we successfully applied the novel QIAscout technology to isolate 44 single cells of two different cell lines (22 cells each from LoVo and SW48) and detected cell type-specific variants using whole genome amplification, followed by low-pass sequencing. In a second ▷

experiment, we selectively isolated single cells of one cell type (LoVo) from a heterogeneous cell population (LoVo and SW48) using the QIAscout. Pyrosequencing® technology was used to confirm the successful and specific isolation of single cells by detecting the presence of three LoVo-specific SNPs and the absence of three SW48-specific SNPs identified in the initial experiment.

This new single-cell isolation platform (Figure 1) is ideal for various cell types, such as adherent or suspension cells, primary cells or cell lines and fluorescent cells. Additionally, the QIAscout 12,000-Microraft Array singularizes cells in an environment similar to a standard cell culture dish. Microrafts carrying individual cells can be dislodged and transferred to reaction tubes using a magnetic wand, without the risk of cross-contamination at any given stage. Subsequently, the cells can be further processed using a variety of downstream applications, such as NGS and PCR or even be further cultivated.

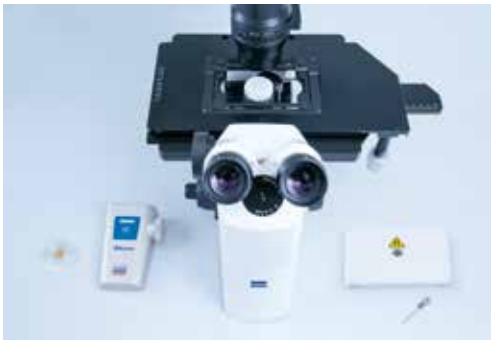


Figure 1. The QIAscout system.

The QIAscout workflow includes the following steps:

- Seeding and cultivating cells in medium on the QIAscout array
- Placing release device containing release needle on the microscope objective
- Placing array on the microscope stage
- Identifying microrafts containing single cells
- Piercing microraft containing cell of interest
- Transferring microraft to secondary vessel using a magnetic wand
- Processing the collected cell for downstream application or further cultivation

Material and Methods

Experiment 1: Identification of cell line-specific variants using whole genome amplification and low-pass sequencing, followed by census-based variant identification

a) Sample preparation and sequencing

To identify cell line-specific SNPs, 22 cells each from LoVo and SW48 cell lines respectively were isolated using the QIAscout and analyzed by next-generation sequencing (NGS) using the QIAseq FX Single Cell DNA Library Kit and a low-pass sequencing method [1].

Two QIAscout arrays were prepared by loading 2 ml of cell culture media containing 6000 cells from each cell line (one array loaded with LoVo cells and another with SW48 cells). Cells in the array were left overnight in a standard cell culture incubator. On the next day, the QIAscout was installed on a 10x objective of an inverted microscope (Axio Vert.A1, Zeiss®) and the array was screened using brightfield microscopy. Twenty-two cells per array were dislodged andtransferred to a reaction tube containing 4 µl PBS (= starting volume for whole genome amplification).

All the single cells collected were further prepared for NGS using the QIAseq FX Single Cell DNA Library Kit (refer to *QIAseq FX Single Cell DNA Library Kit Handbook* for further details). In summary, whole genome amplification followed by NGS library preparation was performed for all samples. All 44 libraries were quantified and pooled equally prior to the sequencing run. The entire workflow was conducted within 4 hours with only ~40 minutes hands-on time. A low-depth next-generation sequencing (mean sequencing depth per sample 0.6x) of 44 prepared libraries was performed on an Illumina® NextSeq® instrument.

b) Analysis of sequencing results

Analysis of sequencing results was performed using the "census-based variant calling" method as described by Zhang et al. [1]. Census-based variant calling requires that acceptable variants be observed in at least two independent single-cell libraries. In this experiment, SNPs have to be observed in at least two libraries of each cell line and they have to be different between the cell lines to be identified as cell line-specific variants. Uniformity of coverage across the genome was quantified by counting reads per 100 kb genomic window, calculating Lorenz curves (not shown) and deriving Gini coefficients, where Gini = 0 indicates perfect uniformity and Gini = 1 indicates complete lack of uniformity.

c) Validation of NGS results using Pyrosequencing

For validation of the NGS results, six Pyrosequencing assays were designed for 3 LoVo- and 3 SW48-specific SNPs using the PyroMark® Assay Design Software (based on the results from Experiment 1b). ▷

Six selected SNPs for Pyrosequencing analysis:

1) rs12879019 (SW48-specific SNP: C/C)

2) rs7077718 (SW48-specific SNP: C/C)

3) rs6443930 (SW48-specific SNP: C/C)

4) rs1539172 (LoVo-specific SNP: G/G)

5) rs7246856 (LoVo-specific SNP: G/G)

6) rs2229429 (LoVo-specific SNP: A/A)

All six Pyrosequencing assays were performed with genomic bulk DNA of each cell line and with 44 WGA samples from Experiment 1a. For Pyrosequencing analysis of SNPs, loci were amplified using the PyroMark PCR Kit. The amplicons were then analyzed by Pyrosequencing using the PyroMark Q48 Autoprep. For the PCR and Pyrosequencing reactions, refer to Protocols: PyroMark PCR Kit and *PyroMark Q48 Autoprep User Manual*.

Experiment 2: Selective isolation of cells from a subpopulation of a mixed cell population

From two co-cultivated cell lines (LoVo and SW48), one cell type (LoVo) was targeted for isolation representing a subpopulation of a mixed cell population.

The QIAscout array was prepared by loading 2 ml of cell culture media containing 6000 cells from the cell mixture (3000 each of LoVo and SW48 cells). Cells in the array were left overnight in a standard cell culture incubator. On the next day, the QIAscout was installed on a 10x objective of an inverted microscope (Axio Vert.A1, Zeiss®) and the array was screened using brightfield microscopy. A single cell was found to be growing on approximately 30% of the microrafts, nearly 5% of the microrafts had more than one cell and about 60% of the microrafts were empty. Sixteen single LoVo cells were selected based on their elongated morphology (Figure 2) and isolated using the efficient and reliable QIAscout method. Microrafts containing the single LoVo cells were dislodged and transferred to a secondary vessel containing 4 µl PBS (= starting volume for whole genome amplification).

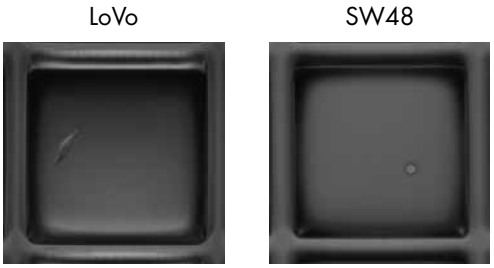


Figure 2. Elongated shape of LoVo and round shape of SW48 cells allow targeted selection of LoVo cells based on morphology.

All samples collected were further processed for amplification of genomic DNA from the single cells using the REPLI-g® Single Cell Kit (refer to *REPLI-g Single Cell Handbook* for further details, see Protocol: Amplification of Genomic DNA from Single Cells). The PyroMark PCR Kit was used to amplify the target region, which was then analyzed by Pyrosequencing using the PyroMark Q48 Autoprep. Six Pyrosequencing reactions were performed per sample in order to identify each single cell for the presence of LoVo-specific SNPs and absence of SW48-specific variants identified in Experiment 1 (refer to Material and Methods 1c).

Results and Discussion

Experiment 1: Identification of cell line-specific variants using low-pass sequencing

a) Sequencing results

Analysis of sequencing results is based on the method described by Zhang et al. [1]. In brief, variants have to be observed in at least two libraries of each cell line and they have to be different between the cell lines in order to be identified as cell line-specific mutation.

Quality of sequencing run

Figure 3 shows that 86.3% bases in all reads have a Q Score of 30 and above, i.e., for the vast majority of bases the Inferred Base Call Accuracy is 99.9% with the probability of an incorrect base call being 1/1000.

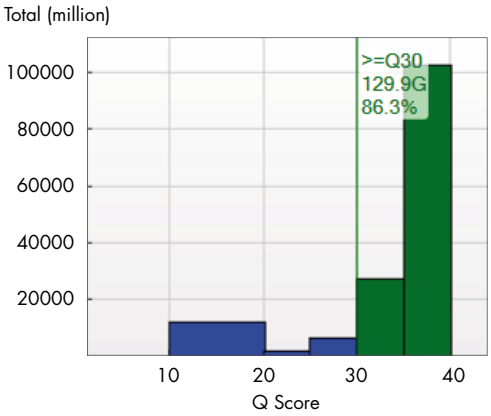


Figure 3. Number of reads per Q Score. Reads with a Q Score ≥ 30 have an Inferred Base Call Accuracy of 99.9%.

In addition, the similarity among all 44 samples is very high (Figures 4–5), not only in terms of identified reads per sample (ranging from 1.45–3.78%, mean: 2.14%, median: 1.98%; expected value if equally distributed: 100/44 = 2.27%), but also with regards to the coverage per bin within a sample (described by the Gini coefficient; smaller is more even).

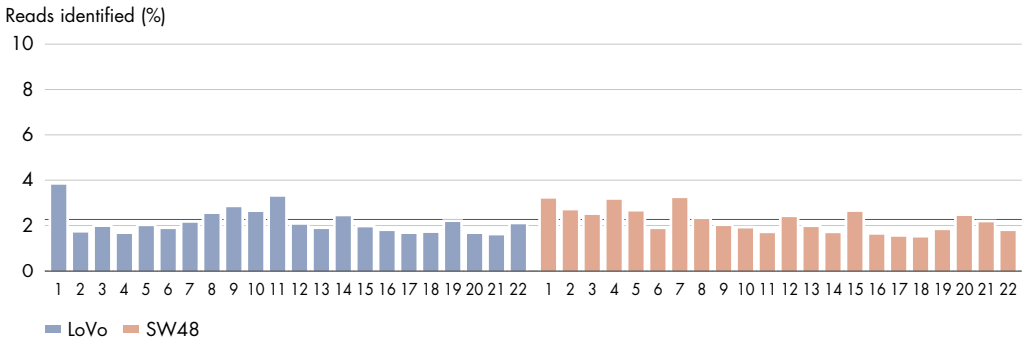


Figure 4. Reads identified per sample. The expected value (green line) if equally distributed is 2.27% (100%/44 samples).

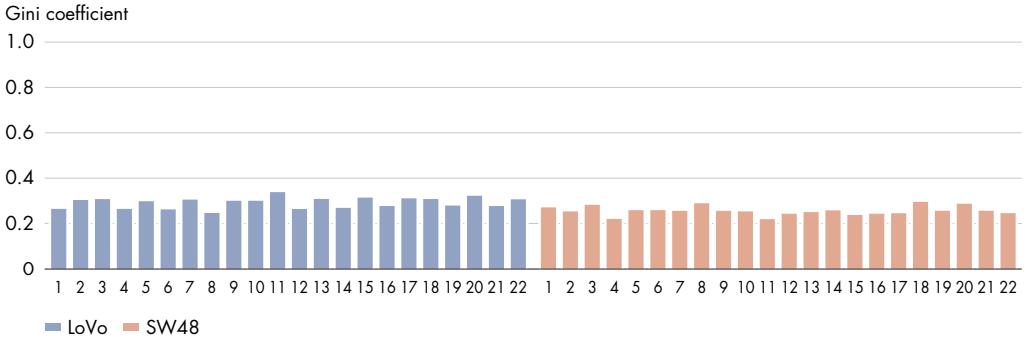


Figure 5. Uniformity of coverage across the genome (counting reads per 100 kb genomic window). Gini = 0 indicates perfect uniformity and Gini = 1 indicates complete lack of uniformity.

Analysis of sequencing results based on census-based variant calling

In total, 190,000 SNPs were identified per cell line including ~66,000 SNPs (~35%) that were specific for each cell line. When taking into account only SNPs that are nonsynonymous missense variations in a coding region, 279 SW48-specific SNPs and 255 LoVo-specific SNPs were detected. Out of these 534 SNPs, three SNPs per cell line were selected for validation using Pyrosequencing (Table 1).

Table 1. SNPs selected for Pyrosequencing.

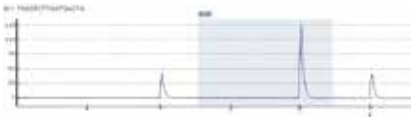
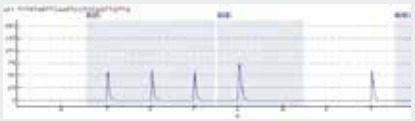
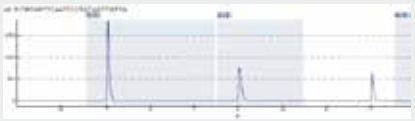
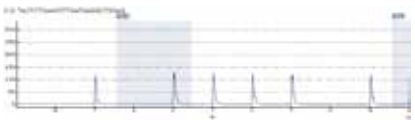
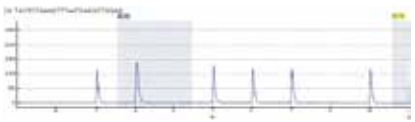
					Number of cells for which SNPs were detected	
Chromosome	**Position**	**Reference SNP ID number (RS)**	**Reference base**	**Altered base**	**LoVo**	**SW48**
chr14	94469564	rs12879019	G	C	0	18
chr10	99929621	rs7077718	G	C	0	17
chr3	184036506	rs6443930	G	C	0	17
chr9	15784633	rs1539172	A	G	17	0
chr19	57445460	rs7246856	T	G	15	0
chr19	7166377	rs2229429	G	A	15	0

b) Validation

Pyrosequencing analysis was used to validate the NGS results. In detail, the specificity of the identified SNPs was verified using the PyroMark Q48 Autoprep platform. This platform allows analysis of 48 samples in parallel and ensures fast and reliable identification of isolated single cells. An easy-to-use standard PyroMark Q48 Autoprep protocol facilitates fully automated template preparation and subsequent sequencing steps requiring minimal hands-on time.

As a first control, genomic bulk DNA from both cell lines was analyzed with all six Pyrosequencing assays (Table 2). The results confirmed the cell line specificity of the six SNPs as indicated by the NGS data.

Table 2. Representative Pyrograms of bulk samples of LoVo and SW48 cells.

Reference SNP ID number (RS)	LoVo	SW48
rs12879019		
rs7077718		
rs6443930		
rs1539172*		
rs7246856*		
rs2229429*		

* For SNPs rs1539172, rs7246856 and rs2229429 reverse Pyrosequencing assays were used. Therefore, cells having a G/G genotype show C/C in Pyrosequencing analysis and A/A genotype is represented as T/T in the Pyrogram®.

As a second step, 44 WGA samples that were used for library preparation were analyzed using Pyrosequencing (Table 3). Similar to the previous Pyrosequencing run, this experiment also verified the NGS results and enabled detection of the cell line-specific SNPs for all 44 samples.

Table 3. Pyrosequencing results of 44 single-cell samples used for low-pass sequencing (22 each of LoVo and SW48 cells).

Reference SNP ID number (RS)	rs12879019	rs7077718	rs6443930	rs1539172	rs7246856	rs2229429
Reference LoVo (NGS and bulk)	G/G	G/G	G/G	C/C	C/C	T/T
LoVo 1–22	G/G	G/G	G/G	C/C	C/C	T/T
Reference SW48 (NGS and bulk)	C/C	C/C	C/C	T/T	A/A	C/C
SW48 1–22	C/C	C/C	C/C	T/T	A/A	C/C

Experiment 2: Selective isolation of cells from a subpopulation of a mixed cell population

The aim of this experiment was to selectively isolate single cells of one cell type from a heterogenous co-culture (LoVo and SW48). Isolated cells were identified by six different Pyrosequencing assays covering cell line-specific SNPs (see Experiment 1).

Using the QIAscout, we could ensure that all the selected and isolated microrafts contained single LoVo cells (Table 4). This result clearly demonstrates the efficacy and reliability of targeted single-cell isolation using the QIAscout method.

Table 4. Pyrosequencing results for 16 targeted isolated single cells. The aim of this experiment was to isolate LoVo cells from a mixture of LoVo and SW48 cells.

Assay	rs12879019	rs7077718	rs6443930	rs1539172	rs7246856	rs2229429
LoVo 1–16	G/G	G/G	G/G	C/C	C/C	T/T
Reference LoVo	G/G	G/G	G/G	C/C	C/C	T/T
Reference SW48	C/C	C/C	C/C	T/T	A/A	C/C

Conclusions

- The QIAscout method allows efficient and reliable isolation of single targeted cells from a heterogenous cell population.
- QIAscout offers a gentle mechanism for the isolation of single cells without affecting cellular status, viability and purity. Single-cell isolation is performed under the visual control of an inverted microscope, thus ensuring high certainty of isolating only a single cell.
- Single cells isolated using the QIAscout are compatible with the QIAseq FX Single Cell DNA Library Kit and demonstrates good sequencing metrics in downstream NGS analysis.
- Single cells isolated with the QIAscout system can be subsequently used for the detection of subpopulation-specific SNPs in combination with the QIAseq FX Single Cell DNA Library Kit using low-pass sequencing and the census-based variant calling method.
- The complete whole genome sequencing workflow can be successfully applied to single-cell research with only minimum hands-on time when using the QIAscout for isolation, the QIAseq FX Single Cell DNA Library Kit for sequence analysis and the PyroMark Q48 Autoprep for NGS data validation.

Reference

1. Zhang C-Z, et al. Calibrating genomic and allelic coverage bias in single-cell sequencing. Nat Commun. 2015;6:6822.

Ordering Information

Product	Contents	Cat. no.
QIAscout	Includes instrument platform and starter pack of 5 arrays	9002733
QIAscout 12,000-Microraft Arrays	5 arrays	928031
REPLI-g Single Cell Kit (24)*	REPLI-g sc Polymerase, Buffers, and Reagents for 24 whole genome amplification reactions (yields up to 40 µg/reaction)	150343
QIAseq FX Single Cell DNA Library Kit (24)*	For 24 reactions: Buffers and reagents for cell lysis, whole genome amplification and library preparation including DNA fragmentation, end-repair and adapter ligation; includes a plate containing 24 barcoded adapters for use with Illumina instruments	180713
PyroMark PCR Kit (200)*	For 200 reactions: 2x PyroMark PCR Master Mix (includes HotStarTaq® DNA Polymerase and optimized PyroMark Reaction Buffer containing 3 mM $MgCl_2$ and dNTPs), 10x CoralLoad® Concentrate, 5x Q-Solution®, 25 mM $MgCl_2$ and RNase-Free Water	978703
PyroMark Q48 Autoprep System	PyroMark Q48 Instrument, multistep pipet, software, documentation and installation	9002470
PyroMark Q48 Advanced Reagents	Reagents for 4 x 48 PyroMark Q48 Autoprep standard reactions: PyroMark Advanced Enzyme Mix, PyroMark Advanced Substrate Mix, Denaturation Solution, Annealing Buffer, Binding Buffer, Nucleotides	974002

* Larger kit sizes available; see **www.qiagen.com**.

The QIAscout and PyroMark Q48 Autoprep System are intended for molecular biology applications. These products are not intended for the diagnosis, prevention, or treatment of a disease. The applications described here are not intended for diagnostic use. No claim or representation is intended to provide information for the diagnosis, prevention, or treatment of a disease.

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.