Discriminating Copy Number Variation from Heterogeneous Samples Using the S3™ Cell Sorter and the QX200™ Droplet Digital™ PCR System

Eli Hefner, Carol Oxford, Teresa Rubio, Carolyn Reifsnyder,
Brandon Williams, Melissa Ma
Bio-Rad Laboratories, Inc., 2000 Alfred Nobel Drive, Hercules, CA 94547

Introduction

Many biological samples are composed of mixtures of physiologically different cell types. These include stem cell cultures, tumor samples, and blood. Understanding the genetic makeup of a subpopulation within these samples may provide valuable clues about the characteristics of the sample and, by extension, the organism. For example, identifying and characterizing rapidly growing subpopulations within a tumor may reveal the metastatic potential of the cancer.

Measuring potentially interesting genetic properties, such as copy number, mutation status, or transcript abundance, is confounded in a mixed sample by competing signals. In stem cell cultures, the pluripotent cells grow alongside differentiated cells, feeder cells, and stromal cells. In some stem cell cultures, the pluripotent cells represent a minority of the total cells within the culture. Accurately measuring transcript abundance within a subpopulation of cells requires separation from the heterogeneous population, or the use of a calibrator, a pair of genes that can be used to determine the relative abundance of each cell type. However, applying a calibrator to mixed cell measurements comes with a host of assumptions that may or may not be valid for that particular culture. Therefore, the most effective way to obtain reliable data from the mixed culture is to sort the cells of interest prior to performing the measurement.

Powerful Solutions by Pairing Key Technologies

Using Bio-Rad's S3 or S3e™ Cell Sorter, heterogeneous populations can be easily sorted into enriched or pure subpopulations based on cell morphology or fluorescently labeled markers (Gilsbach et al. 2014). Sorting with the S3 or S3e Cell Sorter can add tremendous value to the sample preparation procedure upstream to digital PCR. Once a purified sample is collected, Bio-Rad's QX200 Droplet Digital PCR (ddPCR™) System can be used to gather highly precise answers about the genetic makeup of the sample. Absolute quantification using digital PCR is achieved through partitioning the sample into nanoliter-sized droplets, which allows for thousands of discrete measurements per sample.

Copy Number Variation

Copy number variation (CNV), the gain or loss of genes within a genome, is known to play a role in cancer, neurological disease, drug metabolism, and infectivity (Stankiewicz and Lupski 2010). Roughly 12% of the human genome is copy number variable and the contribution of these variations to human phenotypes is being intensely studied (Stankiewicz and Lupski 2010). Characterizing stable, heritable copy number variations is relatively straightforward given the homogeneous nature of the CNVs in the sample. The difficulty arises when genes are amplified or lost within the somatic tissue (Nadauld et al. 2012). For example, a multitude of genes are known to be copy number variable in cancerous cells while cells in surrounding, normal tissue often do not display

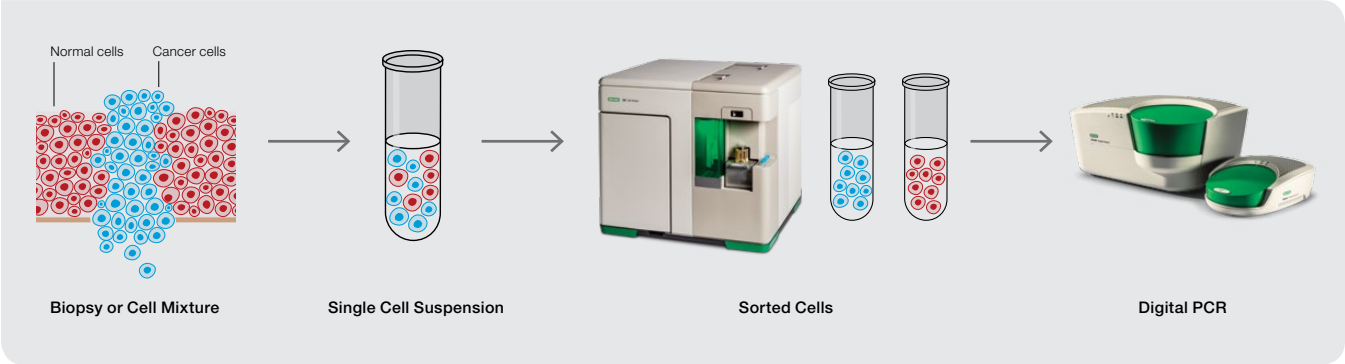
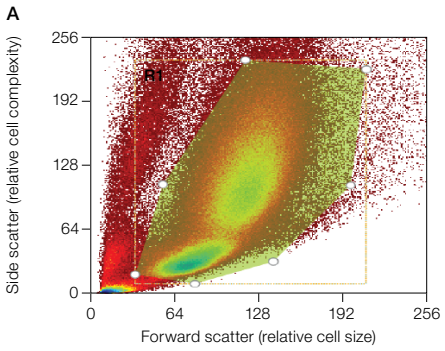
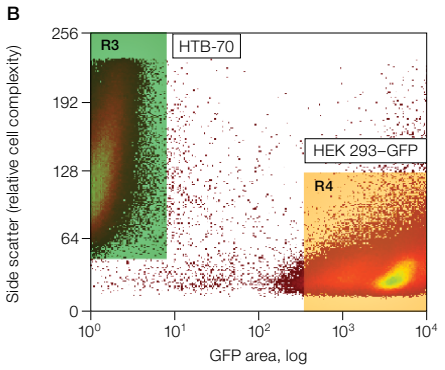


Fig. 1. Workflow for a biopsy or cell mixture.



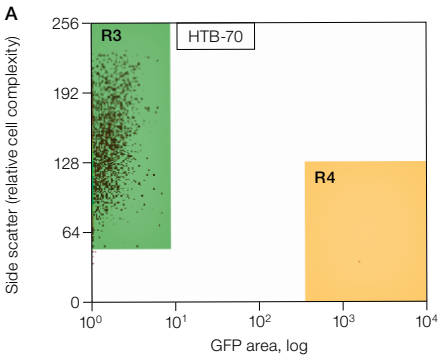


Label	Count	Plot, %	Total, %
Total	307,334	100	79.48
R3	199,195	64.81	51.52
R4	103,714	33.75	26.82

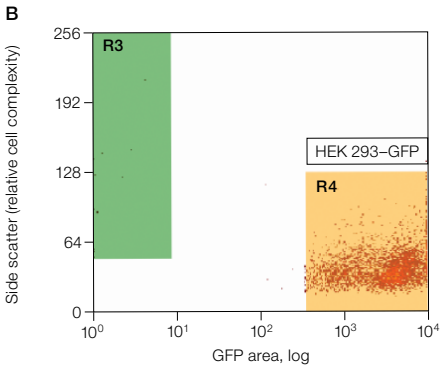
Fig. 2. Presort flow cytometry analysis of HEK 293–GFP vs. HTB-70 cells. **A**, cells were first identified using forward scatter or relative cell size on the x-axis against side scatter or relative cell complexity on the y-axis. A region was drawn and applied to subsequent plot. **B**, HEK 293–GFP cells were identified by plotting HEK 293–GFP against side scatter or relative cell complexity. Regions were drawn and populations were sorted. GFP, green fluorescent protein.

the variations. This can be further confounded by the copy number heterogeneity that can arise within a tumor cell mass and this can make it difficult to identify whether a tumor contains CNVs in genes of interest. Even when a copy number variation is detectable in the heterogeneous tissue, it is not possible to measure the specific copy number state of the amplified gene.

Integrating the S3 or S3e Cell Sorter as part of the sample preparation upstream of the QX200 System can make the results from the cells of interest more specific and informative. The experiment shown in Figure 1 outlines sorting homogeneous populations from a heterogeneous mixture to dissect out key genetic alterations, such as CNVs, that can lead to better identification and quantification of the sample composition.



Label	Count	Plot, %	Total, %
Total	4,021	100	78.44
R3	4,016	99.88	78.35
R4	1	0.02	0.02



Label	Count	Plot, %	Total, %
Total	4,077	100	80.38
R3	11	0.27	0.22
R4	4,051	99.36	79.87

Fig. 3. Postsort purity analysis of HEK 293–GFP vs. HTB-70 cells. **A**, HTB-70 cells were sorted to achieve 99.88% purity; **B**, HEK 293–GFP cells were sorted to achieve 99.36% purity.

Results

The copy number state of the *CCND1* gene was measured in a nonsorted mixed sample and in a sorted enriched sample. Presorted mixed cultured cells indicated a 60:40 ratio of HEK 293–GFP to HTB-70 cells after 24 hr of growth (Figure 2). After sorting, enriched HEK 293–GFP cells were found to be 99.36% pure while enriched HTB-70 cells were 99.88% pure (Figure 3).

Next, digital PCR was used to examine the copy number state of *CCND1* in both individually cultured HEK 293–GFP and HTB-70 cell lines, as well as in the sorted and unsorted HEK 293–GFP/HTB-70 mixed cell co-cultures (Figure 4). Using this method, two copies of *CCND1* were found in the individually cultured

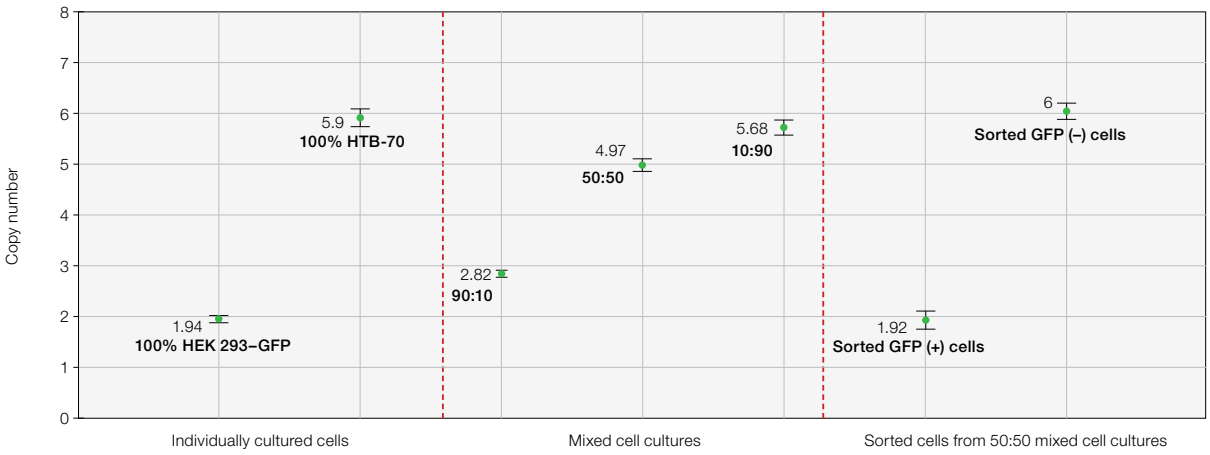


Fig. 4. ddPCR results reveal differences in mixed cultures and sorted cells. Sorted cells from the mixed culture displayed a similar copy number as the individually cultured cells. Mixed cultured cells resulted in an averaging of copy number based on the ratio of cells in the heterogeneous sample.

HEK 293–GFP cell line while six copies of *CCND1* were in the individually cultured HTB-70 cell line. Similarly, two copies of *CCND1* were found in sorted HEK 293–GFP cells while six copies of *CCND1* were also found in sorted HTB-70 cells. However, unlike both the individually cultured or sorted HEK 293–GFP and HTB-70 cell lines, *CCND1* copy number values in unsorted HEK 293–GFP/HTB-70 mixed co-cultures were found to be averaged, ultimately yielding noninteger values of 2.82 copies of *CCND1* in the 90:10 mixture, 4.97 copies in the 50:50 mixture, and 5.68 copies in the 10:90 mixture. These findings demonstrate that when working with samples that may display copy number heterogeneity, such as tumor or stem cell cultures, cell sorting can be used to distinguish subpopulations within a heterogeneous cell mixture, and can also be used to significantly enhance digital PCR findings.

Discussion

An additional layer of information can be gleaned by combining flow cytometry and sorting with ddPCR. By incorporating a DNA binding dye, researchers can estimate the absolute copy number for a particular class of cells. Copy number variation is usually reported as a ratio of a gene of interest relative to a reference gene, where a ratio of 1:1 indicates that the reference gene is present at the same abundance as the gene of interest. Within the G1 state of the cell cycle in a well-karyotyped and normal diploid organism, this would mean that there are two copies of the reference gene and two copies of the gene of interest. In many instances, the cells of interest are polyploid to some degree. This is often true in situations involving analysis of cancer samples. A recent report estimated that 37% of cancers exhibit some type of polyploidy (Zack et al. 2013). The extent of the polyploidy can be estimated by staining cells with a DNA binding dye and comparing the relative fluorescence of the G1 and G2 peaks between the cell types. For example, if the mean fluorescence amplitude for the G1 peak in cell type A is similar and twice that of cell type B, it can be reasonably assumed that cell type A has two times the genome complement. If cell type B is a well-characterized normal diploid, then it is a reasonable assumption that cell type A is tetraploid. In this example, if copy number analysis were performed and the ratio of reference gene to gene of interest were 1:1, then the total number of genes in the G1 phase of the cell cycle would be four for cell type A, assuming the reference is a single locus gene.

Materials and Methods

Cell Culture

HEK 293–GFP (Cell Biolabs, Inc.), a stably expressing green fluorescent protein (GFP) cell line that contains one copy of the cyclin D1 gene (*CCND1*) per genome (two copies of *CCND1* are present in the G1 phase of the cell cycle), and ATCC HTB-70 (American Type Culture Collection), a melanoma cell line that does not express GFP and contains three copies of *CCND1* per genome (six copies of *CCND1* are present in G1 phase of the cell cycle) were used. Both cell types were either individually cultured, or co-cultured at 90:10, 50:50, and 10:90 ratios of HEK 293–GFP to HTB-70 cells for 24 hr prior to trypsinization and detachment from the tissue culture plate.

Cell Sorting

A mixed culture containing an equal ratio of HEK 293–GFP and HTB-70 cells was sorted using the S3 Cell Sorter. Cells were first identified by their relative size and complexity using forward and side scatter analysis. Cell regions were then selected based on GFP expression in fluorescence channel 1 (FL1). Sorting was conducted in Purity mode and approximately 100,000 cells from each population were sorted and collected for analysis.

Droplet Digital PCR

Prior to amplification, cellular DNA was extracted from individual, co-cultured, and enriched sorted cell populations using a DNA extraction kit (QIAGEN). DNA was subsequently quantified using a NanoDrop Spectrophotometer (Thermo Fisher Scientific Inc.). Cellular DNA was then amplified using the QX200 System and assayed using Bio-Rad's PrimePCR™ ddPCR™ Copy Number Assay: Cyclin D1 (*CCND1*) Human Kit. This kit produced an exonic amplicon of 104 nucleotides.

Conclusions

In this study, we show that combining technologies such as flow cytometry, cell sorting, and digital PCR can enable highly accurate and more specific results attributed to cell populations of interest vs. the bulk heterogeneous populations. By pairing tools like Bio-Rad's S3 Cell Sorter and the QX200 Droplet Digital PCR System, targeted populations can be further characterized to uncover correlations in cancers and diseases.

References

Gilsbach R et al. (2014). Dynamic DNA methylation orchestrates cardiomyocyte development, maturation and disease. Nat Commun 5, 5288.

Nadauld L et al. (2012). Quantitative and sensitive detection of cancer genome amplifications from formalin fixed paraffin embedded tumors with droplet digital PCR. Transl Med (Sunnyvale) 2, 1000107.

Stankiewicz P and Lupski JR (2010). Structural variation in the human genome and its role in disease. Ann Rev Med 61, 437–455.

Zack TI et al. (2013). Pan-cancer patterns of somatic copy number alteration. Nat Genet 45, 1134–1140.

Visit bio-rad.com/web/S3ddPCR for more information.

Bulletin 6681 Rev A 14-2213 0115