Copy number variation analysis using the QuantStudio 3D Digital PCR System

Introduction

Copy number variation (CNV) is an imbalance in the genome that increases or decreases the wild type copy number of a locus in comparison to a reference genome. These genomic alterations can range from small (<10 kb) insertions or deletions to large (>1 Mb), complex, multiallelic duplications. CNVs are one of the most common genetic variations in the human genome and have been implicated in many diseases, including cancer or inherited disease susceptibility [1,2]. As a result, simple and reliable methods are needed to quantify CNVs as potential biomarkers and for understanding the molecular mechanisms of tumor formation.

Current methods to assess CNVs are summarized in Table 1. These methods include *in situ* hybridization, but this method is lengthy, labor-intensive, and interpretation of results can be subjective [3]. Higher precision is offered by array comparative genome hybridization (aCGH), although this method may require significant hands-on time and resolution is dependent on the type of array chosen [4]. More recently, advances in next-generation sequencing technologies have enabled cost-effective methods to detect multiple types of genomic variations in a single run [5]. Lastly, of all current methods, real-time and digital PCR–based technologies offer the simplest workflows that enable accurate copy number results with minimal hands-on time and fast turnaround time.

Applied Biosystems™ TaqMan™ Copy Number Assays are widely used to evaluate CNVs using traditional real-time PCR instruments and software. With over 1.6 million predesigned assays and custom design tools available, they offer simple workflows along with specific and highly reproducible copy number results. Despite these significant workflow benefits, a limitation of copy number assays run using traditional real-time PCR is the reduction in measurement precision the further the measurement is from a simple doubling or loss. In many cases, measurements are not sufficiently precise for determining high absolute copy number or for determining very small copy number differences.

The Applied Biosystems™ QuantStudio™ 3D Digital PCR System uses digital PCR (dPCR), a technology capable of highly precise measurements, to differentiate subtle changes in copy number. In this application note, we demonstrate the precise measurement of genes at both low and high copy numbers. High reproducibility at a resolution superior to that achievable with real-time PCR is demonstrated, even using formalin-fixed, paraffin-embedded (FFPE) tissue samples. We also provide general guidance on measuring the copy number of genes that are closely linked together in the same genomic region. Our results demonstrate that the QuantStudio 3D system offers a sensitive, accurate, and robust method with unparalleled precision for CNV analysis in cancer research and other fields.

ThermoFisher SCIENTIFIC

Table 1. Common molecular methods for assessment of CNVs.

	In situ hybridization	aCGH	Next-generation sequencing	Real-time PCR	Digital PCR
Description	DNA sequence detection within individual fixed cells using a labeled probe	Independent labeling of single-stranded test and reference DNA in a 1:1 ratio with subsequent hybridization to an oligo-spotted or BAC array	High-throughput sequencing followed by mapping and counting the sequence of interest to determine absolute or relative copy number	Relative quantification of target DNA sequences through the real-time monitoring of the PCR amplification process	Absolute quantification of target DNA molecules by the separation of a PCR reaction sample at limiting dilution into a large number of partitions; concentration is then calculated using standard Poisson statistical methods
Sample preparation requirements	Cell preservation and sectioning	Standard genomic DNA isolation methods	Standard genomic DNA isolation methods	Standard genomic DNA isolation methods	Standard genomic DNA isolation methods
Hands-on time	Extensive	Significant	Significant	Minimal	Minimal
Time-to-results	Days	Days	Days	Hours	Hours
Interpretation of results	Subjective	Objective	Objective	Objective	Objective
Quantitative precision	•	••	••••	••••	•••••
Resolution	•	•••	•••••	•••••	•••••

Materials and methods

The general workflow for the QuantStudio 3D Digital PCR System is shown in Figure 1. The system uses a high-density nanofluidic chip containing 20,000 reaction wells to partition a sample into thousands of independent PCR reactions. The procedures for sample preparation can vary depending on the specific experiment being performed, but these procedures are typically no different than those currently being used for standard PCR approaches. For detailed instructions on performing the workflow, and details of the underlying digital approach, please refer to the QuantStudio 3D Digital PCR System user guide [6].

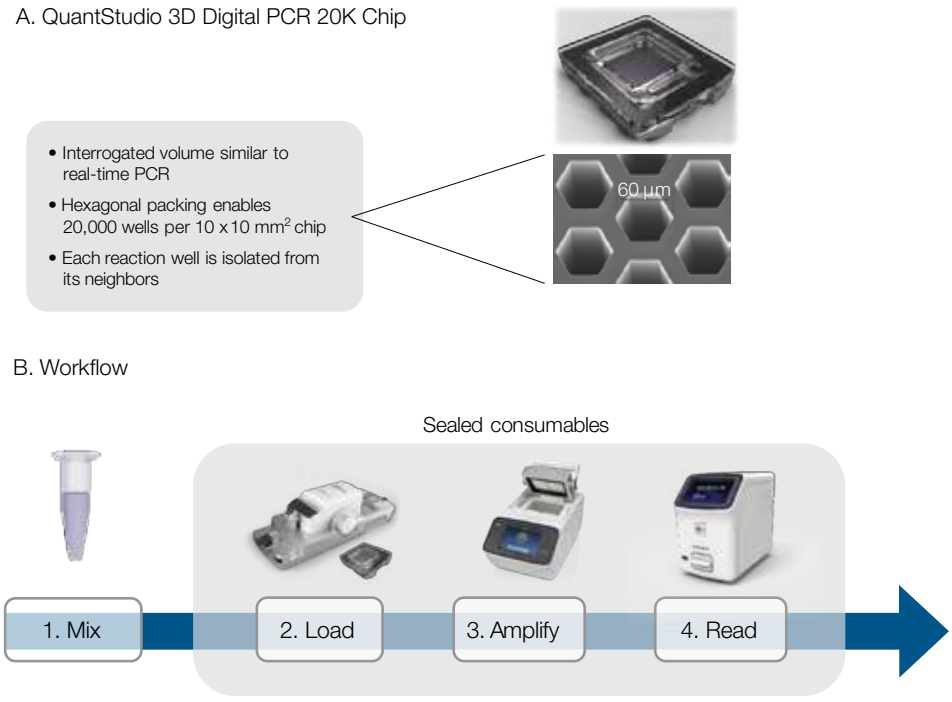


Figure 1. The Applied Biosystems™ QuantStudio™ 3D Digital PCR System 20K Chip and workflow. (A) The QuantStudio 3D Digital PCR Chip consists of an array of 20,000 independent reaction wells. Samples and amplification products are completely contained in the chip throughout the workflow. **(B)** A PCR reaction mix composed of sample, assay(s), and master mix is loaded onto the QuantStudio 3D Digital PCR Chip, amplified on a thermal cycler, and the target concentration in copies/µL is read on the QuantStudio 3D Digital PCR Instrument. Secondary analysis is performed with QuantStudio 3D AnalysisSuite Cloud Software. The workflow features sealed consumables, limited hands-on time, and minimal sample loss.

Genomic DNA samples and predigestion

Genomic DNA (gDNA) samples for CNV analysis were purchased from the Coriell Repository. The genes analyzed in this study were *CCL3L1*, *ERBB2 (HER2)*, and *C4A/C4B*. For higher copy numbers, or where genes are in close proximity to each other, it might be necessary to separate copies by restriction enzyme digestion. For copy number analysis of *C4A/C4B*, DpnII (New England Biolabs) was used in a 25 µL digestion reaction in its recommended digestion buffer at 37°C for 1 hour and then followed by heat inactivation at 65°C for 20 minutes. For *CCL3L1* and *HER2* copy number analysis, samples were demonstrated not to require restriction enzyme digestion before loading onto the chip. Please see the Appendix for more detailed information on choosing a restriction enzyme.

TaqMan Copy Number Assays

Existing TaqMan Copy Number Assays are compatible with dPCR. Table 2 shows the Assay IDs for the copy number and reference assays used in this study. The Applied Biosystems™ FAM™ dye–labeled assay for the target of interest was duplexed with the Applied Biosystems™ VIC™ dye–labeled TaqMan Copy Number Reference Assay for RNase P (Cat. No. 4403326). Alternatively, the copy number reference assay for *TERT* (Cat. No. 4403316) can be used as a substitute. While these reference assays are commonly suitable for samples in most studies, it should be empirically confirmed that the reference genes have been maintained at 2 copies across all samples being tested.

Table 2. Human test and reference assays used in this study.

Gene name	Assay ID or Cat. No.
CNV assay	
CCL3L1	Hs03198166_cn
C4A/C4B (C4L)	Hs07226352_cn
ERBB2 (HER2)	Hs00817646_cn
CNV reference assay	
RNase P *(RPPH1)*	Cat. No. 4403326
TERT	Cat. No. 4403316

Digital PCR reaction setup

Reaction mixes were set up containing the Applied Biosystems™ QuantStudio™ 3D Digital PCR Master Mix, TaqMan Copy Number Assay for the target of interest, RNase P Reference Assay, and Coriell gDNA sample (see Figures 2 and 3 for specific sample IDs). A total volume of 16 µL PCR reaction mix was prepared for each sample, and 14.5 µL was loaded onto the chip. The amount of gDNA to be loaded on the chip usually contains 200–2,000 copies/µL in the final dPCR reaction mix so that each reaction well in the chip receives, on average, 0.6–1.6 copies of the target sequence. Adjustment for input gDNA may be required depending on the copy number of the target of interest. Since the dPCR reaction is a duplex reaction, it is recommended that the target of interest and reference target be within this range. If the target-to-reference ratio is very high (8–10 or more), a dilution series may be needed. PCR thermal cycling conditions and additional details can be found in the user guide [6].

Data analysis

After PCR amplification, chips were read on the QuantStudio 3D Digital PCR System. Absolute quantification data were exported from QuantStudio 3D AnalysisSuite Cloud Software. Copy number per diploid genome was calculated with Excel™ Software using the absolute quantification number of FAM dye–labeled target and VIC dye–labeled RNase P reference to determine the haploid gene copy number and then multiplied by 2 to convert to diploid genome copy number.

Results

Quantification of high copy number with high precision and accuracy

Precision limitations and the logarithmic algorithm of traditional real-time PCR prevent the resolution of subtle differences beyond a copy number of 4. To demonstrate the ability of dPCR to detect higher copy number, a panel of 9 gDNA samples was analyzed using the QuantStudio 3D Digital PCR System and a TaqMan Copy Number Assay for the *CCL3L1* gene, which is a chemokine gene of variable copy number among individuals [7]. Results indicated that the samples contained variations from 0 to 8 copies per genome (Figure 2). A statistically measurable difference between samples containing 7 and 8 copies was clearly discernable as a result of the high degree of precision achieved, confirming that dPCR can differentiate less than a 1.2-fold difference.

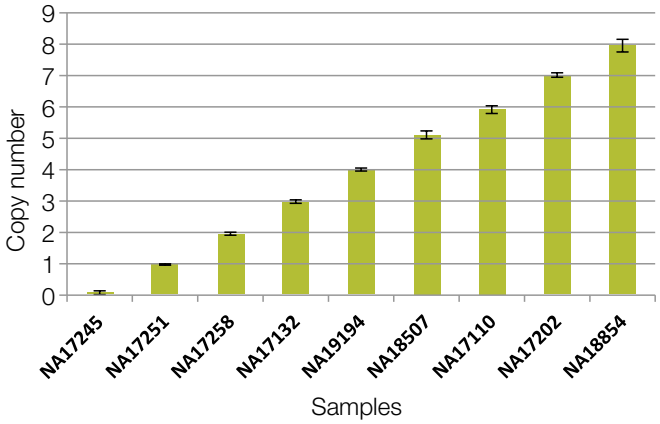
High-resolution copy number analysis in heterogeneous samples

To further demonstrate the capability of dPCR for highly precise measurements, two Coriell Repository samples that contained 2 copies and 3 copies of the *CCL3L1* genetic locus were mixed at different ratios to simulate samples with predicted copy numbers between 2 and 3 copies in 0.1 copy increments. While individual cells would not be expected to display fractional copy number differences, this simulation is highly relevant to heterogeneous cell populations where the detected copy number is an average of all cells present. Copy numbers for the mixed samples were measured using the QuantStudio 3D Digital PCR System and plotted against predicted copy number (Figure 3). The results indicate that the QuantStudio 3D Digital PCR System is capable of resolving very small differences (<5%) in copy number.

A

Samples	Number of replicates	Expected CN	Detected CN_Mean	Standard deviation	CV (%)
NA17245	6	0	0.08	0.060	NA
NA17251	6	1	0.98	0.022	2.21
NA17258	6	2	1.96	0.048	2.47
NA17132	6	3	2.98	0.055	1.85
NA19194	8	4	4.00	0.049	1.22
NA18507	8	5	5.11	0.128	2.50
NA17110	8	6	5.91	0.122	2.07
NA17202	8	7	7.02	0.072	1.02
NA18854	8	8	7.95	0.203	2.55

B



C

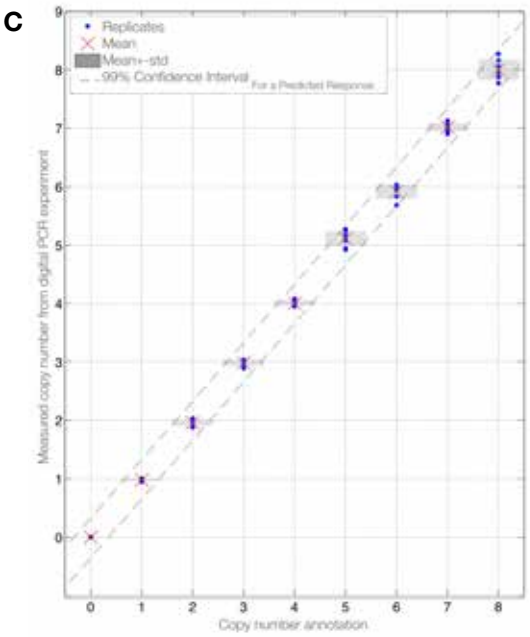


Figure 2. Precision and accuracy of copy number analysis of the *CCL3L1* genetic locus. (A) Copy number (CN) was measured across 9 gDNA samples. The coefficient of variation (CV) was below 2.6% for each set of technical replicates, demonstrating a high degree of measurement reproducibility within each replicate group. **(B)** As demonstrated by non-overlapping error bars, the precise measurements enable statistical discernment of samples containing 7 and 8 copies of *CCL3L1*. **(C)** Measured copy number plotted against expected copy number for each sample. The blue dots represent copy number from individual chips, and red crosses represent the mean of all measurements for the given sample. The grey shaded rectangular bars are their standard deviations, and the dotted line represents the 99% Confidence Interval for a Predicted Response. Data quality was confirmed in QuantStudio 3D AnalysisSuite Cloud Software, and copy number was calculated in Excel™ Software from absolute quantification data reported in QuantStudio 3D AnalysisSuite Cloud Software.

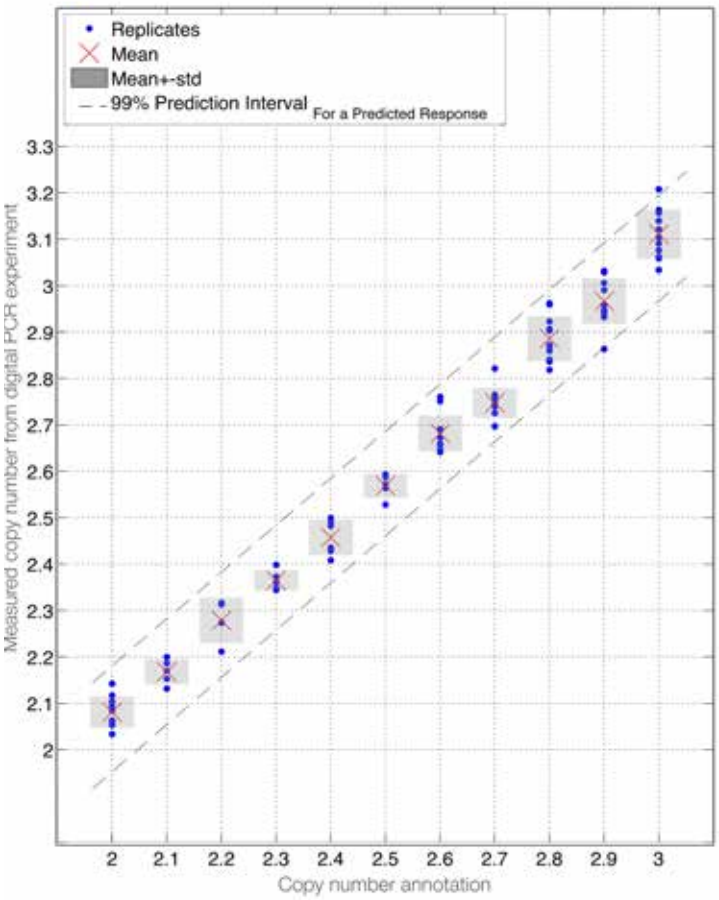


Figure 3. High-resolution copy number analysis of the *CCL3L1* genetic locus. Two samples from the Coriell Repository that contained 2 copies (NA17258) or 3 copies (NA17132) of the *CCL3L1* gene were mixed to generate samples with predicted copy numbers between 2 and 3 copies in 0.1 increments (2.0, 2.1, 2.2, etc.). Absolute quantification data from the QuantStudio 3D Instrument were exported from QuantStudio 3D AnalysisSuite Cloud Software and copy numbers for the mixed samples were calculated using Excel Software.

Copy number analysis of the HER2 gene using FFPE samples

FFPE samples are a widespread resource for molecular profiling and biomarker discovery of clinical samples. We analyzed 43 breast cancer FFPE samples for amplification of the *HER2* gene (also known as *ERBB2*), a driver of cellular proliferation in cancer cells and an important biomarker for targeted therapy [8,9]. A predesigned TaqMan Copy Number Assay for *HER2* and Applied Biosystems™ TaqMan™ RNase P Copy Number Reference Assay was used for analysis. *HER2* copy number was determined by using Excel Software to calculate the ratio of absolute quantification with data exported from QuantStudio 3D AnalysisSuite Cloud Software (Figure 4). The TaqMan Copy Number Reference Assay for *TERT* was also tested, and consistent results were observed (data not shown).

To validate the results, we compared the data generated by dPCR to data from the same samples that were previously analyzed by silver *in situ* hybridization (SISH). For SISH, *HER2* gene amplification was defined as positive (P) or negative (N) using American Society of Clinical Oncology (ASCO) and College of American Pathologists (CAP) 2007 guidelines [10]. A high concordance between dPCR and SISH results is shown in Figure 4. For more information on this study, please see reference 11.

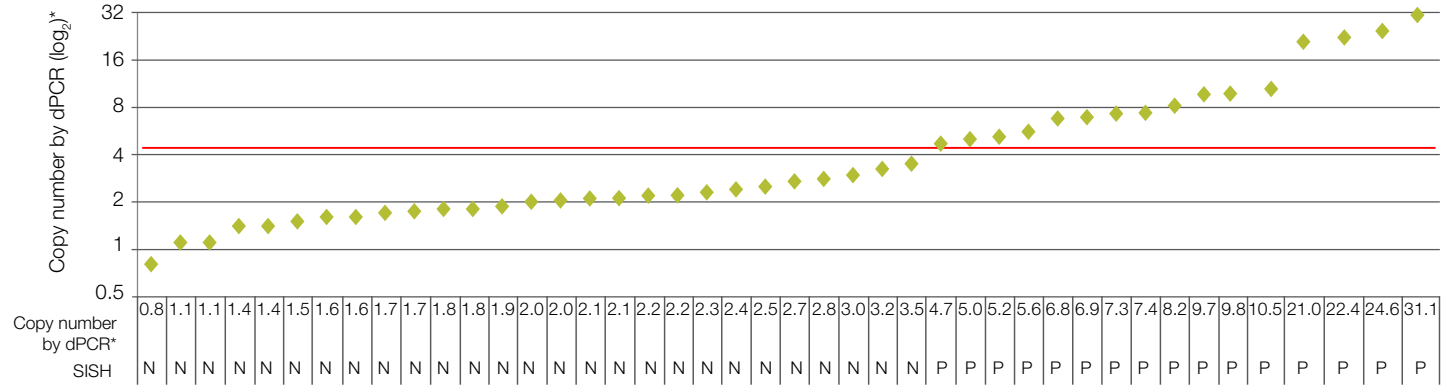


Figure 4. Copy number analysis of the *HER2* gene using dPCR and SISH. Forty-three FFPE samples were analyzed for *HER2*:RNase P copy number ratios as determined on the QuantStudio 3D system. The ratios were subsequently converted to diploid copy number. *HER2* amplification status is shown as positive (P) or negative (N) as previously determined by SISH. The red line is drawn at a copy number of 4.4, corresponding to the *HER2/CEP17* cutoff ratio of 2.2 for a positive result using ASCO/CAP guidelines (2007). *CEP17* is the reference sequence used in SISH, which targets the centromeric region of chromosome 17. *HER2* statuses determined by the two methodologies were in agreement across all but one sample that immediately bordered the threshold cutoff.

*Copy number ratios calculated using Excel Software

Conclusions

The QuantStudio 3D Digital PCR System provides a sensitive, accurate, and robust technology with a simple workflow for copy number analysis. We demonstrate highly accurate measurement of 0 to 8 copies with high precision (CV <2.6%) using samples with known copy number for the *CCL3L1* gene. Moreover, high reproducibility at a resolution superior to that achievable with real-time PCR is demonstrated, enabling characterization of very small detectable differences among simulated highly heterogeneous samples. Finally, we detect *HER2* gene amplification in FFPE samples and demonstrate high concordance with SISH results from the same set of samples.

A catalog of over 1.6 million predesigned TaqMan Copy Number Assays is available for use with the QuantStudio 3D Digital PCR System. Should a locus of interest not be represented in this large collection, custom design options are available. While real-time PCR remains highly relevant for CNV analysis due to its cost and throughput benefits, digital PCR is enabling studies that require detection of small and subtle differences between samples. This unparalleled precision enables CNV analysis of heterogeneous samples and analysis of targets with high copy that are commonly studied in cancer research and other fields.

Appendix

Choosing a restriction enzyme for copy number analysis

It might be necessary to separate closely linked copies by restriction enzyme digestion, depending on the distance between adjacent copies and sample type. Our data support that no digestion is required if the distance is over 100 kb (*CCL3L1* in our study), or if using FFPE samples, due to their inherent fragmented state (*HER2* in our study). We chose the long form of *C4* (*C4L* or *C4A/C4B*) for analysis because the distance between copies is shorter than the other genes in our study. These genes are components of the blood complement system and are encoded by two highly similar loci [12]. Digestion is recommended if the distance is less than 100 kb (*C4A/C4B* in our study) or if the distance is not known. Figure 5 illustrates the difference in copy number that can be observed between digested and non-digested samples. Copy numbers are underestimated in non-digested samples that contain more than 3 copies of *C4A/C4B*, but restriction enzyme digestion has no effect on results for *CCL3L1* because the distance between copies is over 100 kb.

Candidate restriction sites can be mapped using DNA sequence information downloaded from NCBI or EMBL and one of the many mapping tools available on the Web, or specialized sequence analysis software such as Invitrogen™ Vector NTI™ Software. The chosen restriction enzyme sites must be between the duplications but not within the assay's target sequence. We recommended starting with a restriction enzyme that cuts more than 150 bp from the assay position (available from the assay description on the Thermo Fisher Scientific website), and that is methylation-insensitive. For example, DpnII is a good candidate for *C4A/C4B* and *CCL3L1* gene targets, as it fulfills these criteria.

Determining the appropriate sample dilution

It is important to calculate gDNA concentration prior to the digestion reaction, since this will be carried through the workflow, and ensure that the optimal concentration of material is loaded on the QuantStudio 3D Digital PCR Chip. To minimize potential effects of the digestion buffer on downstream PCR, post-digestion dilution should be calculated and performed. If starting material is limited, as little as 100–500 ng undigested gDNA can be used. Reducing the restriction enzyme digestion volume might be needed if the gDNA input amount is low. For further details on sample dilution, please refer to the user guide [6].

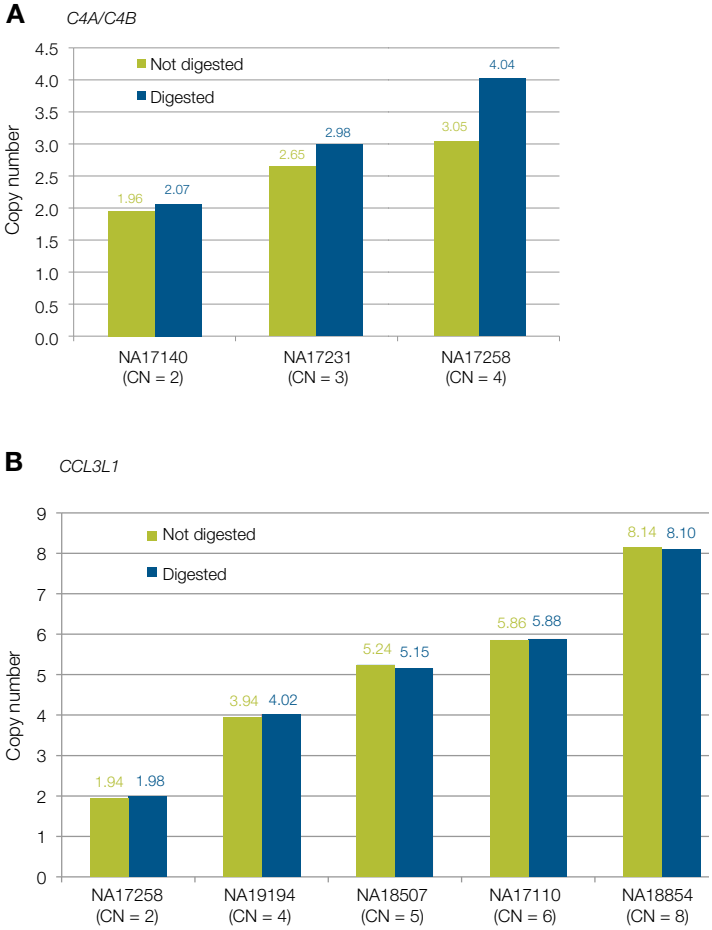


Figure 5. Effect of restriction enzyme digestion on measured copy number. Samples with the indicated copy number were either untreated or digested with DpnII and analyzed on the QuantStudio 3D system. **(A)** Without restriction enzyme digestion, copy numbers are underestimated in samples containing more than 3 copies of *C4A/C4B*. **(B)** Restriction enzyme treatment has no effect on CCL3L1 copy number results because the distance between copies is over 100 kb, even at up to 8 copies per genome.

References

1. Inaki K, Liu ET (2012) Structural mutations in cancer: mechanistic and functional insights. *Trends Genet* 28:550–559.
2. Almal SH, Padh H (2012) Implications of gene copy number variation in health and diseases. *J Hum Genet* 57:6–13.
3. Starczynski J, Atkey N, Connelly Y et al. (2012) *HER2* gene amplification in breast cancer: a rogues' gallery of challenging diagnostic cases. *Am J Clin Pathol* 137:595–605.
4. Li W, Olivier M (2013) Current analysis platforms and methods for detecting copy number variation. *Physiol Genomics* 45:1–16.
5. Fang LT, Lee S, Choi H et al. (2014) Comprehensive genomic analyses of a metastatic colon cancer to the lung by whole exome sequencing and gene expression analysis. *Int J Oncol* 44:211–221.
6. QuantStudio 3D Digital PCR System User Guide. Pub. No. MAN0007720, Rev. A.
7. Irving SG, Zipfel PF, Balke J et al. (1990) Two inflammatory mediator cytokine genes are closely linked and variably amplified on chromosome 17q. *Nucleic Acids Res* 18:3261–3270.
8. Slamon DJ, Clark GM, Wong SG et al. (1987) Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science* 235:177–182.
9. Mitri Z, Constantine T, O'Regan R (2012) The HER2 receptor in breast cancer: Pathophysiology, clinical use, and new advances in therapy. *Chemother Res Pract* 2012:743193.
10. Wolff AC, Hammond ME, Schwartz JN et al. (2007) American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. *J Clin Oncol* 25:118–145.
11. Copy number variation in breast cancer translational research; QuantStudio 3D Digital PCR System as a cost-effective and sensitive alternative for *HER2* gene amplification assessment. Pub. No. CO09771, available at thermofisher.com.
12. Blanchong CA, Chung EK, Rupert KL et al. (2001) Genetic, structural and functional diversities of human complement components C4A and C4B and their mouse homologues, Slp and C4. *Int Immunopharmacol* 1:365–392.

applied biosystems

Ordering information

Product	Cat. No.
QuantStudio 3D Digital PCR System Package v2—includes:*	A29154**
QuantStudio 3D Digital PCR Instrument with Power Cord	4489084
ProFlex 2 x Flat PCR System	4484078
QuantStudio 3D Digital PCR Chip Adapter Kit for Flat Block Thermal Cycler	4485513
QuantStudio 3D Digital PCR Master Mix v2	A26358
QuantStudio 3D Digital PCR Chip Loader	4482592
QuantStudio 3D Digital PCR 20K Chip v2	A26316

* Part numbers listed in bundle are for individual components.

**Cat. No. A29154 is for all regions except Europe, the Middle East, and Africa (EMEA). Please use Cat. No. A29737 or A29738 for customers residing in EMEA. Package components are slightly different. Please check with your regional sales representative for details.

Find out more at **thermofisher.com/quantstudio3d**

ThermoFisher SCIENTIFIC