Pharmacogenomics Copy Number Experiments

Best Practices & Troubleshooting Guide

For use with:
QuantStudio® 12K Flex System
CopyCaller® Software

Revision August 14 2015

For Research Use Only. Not for use in diagnostic procedures.

The information in this guide is subject to change without notice.

DISCLAIMER

TO THE EXTENT ALLOWED BY LAW, LIFE TECHNOLOGIES AND/OR ITS AFFILIATE(S) WILL NOT BE LIABLE FOR SPECIAL, INCIDENTAL, INDIRECT, PUNITIVE, MULTIPLE, OR CONSEQUENTIAL DAMAGES IN CONNECTION WITH OR ARISING FROM THIS DOCUMENT, INCLUDING YOUR USE OF IT.

Important Licensing Information

These products may be covered by one or more Limited Use Label Licenses. By use of these products, you accept the terms and conditions of all applicable Limited Use Label Licenses.

Trademarks

All trademarks are the property of Thermo Fisher Scientific and its subsidiaries unless otherwise specified**.** TaqMan and AmpliTaq Gold are registered trademarks of Roche Molecular Systems, Inc., TaqMan used under permission and license. NanoDrop is a registered trademark of NanoDrop Technologies, LLC. Puritan and PurFlock are trademarks of the Puritan Medical Products Company, LLC.

© 2015 Thermo Fisher Scientific. All rights reserved.

Contents

CHAPTER 1: About TaqMan® Copy Number Assays

Introduction to copy number analysis

It is important to measure the copy number of drug metabolizing enzyme (DME) genes known to exhibit copy number variation (CNV), particularly CYP2D6, for accurate assessment of an individual's drug metabolization status. For full details on running TaqMan® Copy Number Assay experiments, refer to the TaqMan® Copy Number Assays Protocol (Pub. no. 4397425). The guide provides best practices and techniques for optimal performance of copy number variation quantitation experiments using TaqMan® Copy Number Assays and TaqMan® Copy Number Reference Assays for the QuantStudio® 12K Flex Real-Time PCR System.

How the assays work

TaqMan® Copy Number Assays are run simultaneously with a TaqMan® Copy Number Reference Assay in a duplex real-time PCR reaction. The Copy Number Assay detects the target gene or genomic sequence of interest, and the Reference Assay detects a sequence that is known to exist in two copies in a diploid genome (for example, the human RNase P H1 RNA gene).

The number of copies of the target sequence in each test sample is determined by relative quantitation (RQ) using the comparative C_T ($\Delta\Delta C_T$) method. This method measures the C_T difference (ΔC_T) between target and reference sequences, and then compares the ΔC_T values of test samples to a calibrator sample(s) that contain a known number of copies of the target sequence. The sample copy number of the target is calculated to be two times the relative quantity when the calibrator sample has two copies of the target (recommended). During the PCR, a single C_T difference indicates a doubling of the amount of DNA in a cycle. The amplification efficiency of each assay in a PCR must be close to 100% to ensure the C_T values collected for the samples within a dataset accurately measure the amount of sample DNA present. Perturbations to the PCRs (e.g. PCR inhibitors) can impact assay efficiency and data quality.

Sample Copy Number (CN) Calculation:

1. $\Delta C_T = C_T\_target - C_T\_reference$
2. Calculate the average for replicate runs
3. $\Delta\Delta C_T = \Delta C_T\_sample - \Delta C_T\_calibrator$
4. $RQ = 2^{-\Delta\Delta CT}$
5. CN_sample= RQ x CN_calibrator(2)

CYP2D6 CNV analysis

An important DME gene, CYP2D6, is copy number variant. Individuals may carry deletion alleles (known as the *5 allele) or extra copies of CYP2D6 (e.g. *1XN, *2XN, *4X2, etc.). In addition, hybrid alleles exist that formed by recombination between CYP2D6 and the highly homologous CYP2D7 pseudogene; e.g. *36 contains a gene conversion to CYP2D7 sequences in exon 9 and is associated with negligible CYP2D6 enzyme activity[1,2].

Three CYP2D6 copy number assays are available for detection of CYP2D6 CNV alleles and hybrid alleles.

Recommended: The CYP2D6 exon 9 assay [(Hs00010001_cn]) is the primary copy number assay for detection of CYP2D6 CNV. This assay detects normal full-length CYP2D6 alleles and will not amplify

CYP2D7 or CYP2D8 pseudogenes or CYP2D6/CYP2D7 hybrid alleles carrying CYP2D7 exon 9 sequences (e.g. CYP2D6*36). Hybrid alleles do not contribute to drug metabolizer status and it is not necessary to identify them in most samples.

Additional Assays: The CYP2D6 intron 2 assay (Hs04083572_cn]) and the CYP2D6 intron 6 assay (Hs04502391_cn) will not amplify pseudogenes, but they will amplify full length CYP2D6 and CYP2D6/CYP2D7 hybrid alleles containing CYP2D6 intron 2 and intron 6 sequences, respectively (e.g. CYP2D6*36).
If detection of both full length CYP2D6 and hybrid alleles is desired, then intron 2 and/or intron 6 assays can be used in addition to the exon 9 assays. In some cases, use of intron 2 and intron 6 assays aides proper sample genotyping. For example, if a sample that has one copy of exon 9 yet is heterozygous for the 100C>T SNP, it likely carries a *36 hybrid allele and has the diplotype: *1/*36. The intron 2 and intron 6 assays could be used to detect the *36 allele and confirm the sample diplotype.

IMPORTANT! CYP2D6 intron 2 or intron 6 copy number assays should not be used alone to detect CYP2D6 duplications as they will also detect nonfunctional hybrid alleles that do not represent duplications and that do not contribute to CYP2D6 metabolizer status. The CYP2D6 exon 9 copy number assay must be used to detect true gene duplication event.

For information on other DME genes that are known to exhibit CNV, and the TaqMan® Copy Number Assays used to detect them, please refer to the Pharmacogenomics Experiments User Guide (Publication Number MAN0009612).

CHAPTER 2: QuantStudio® 12K Flex Instrument Requirements

Instrument Clearance

Proper clearance for instrument ventilation is critical for optimal performance and reproducibility.

IMPORTANT! Clearance – At least 15.2 cm (6 in) of clearance for ventilation, service access, and cable routing at the back of the instrument. The total width of the required space with the instrument and a desktop computer will be 162.5 cm (64 in) and the total depth will be 81.3 cm (32 in). If a robot is included, the total width will be 213 cm (84 in).

Instrument Calibration

Ensure that the instrument is calibrated according to the maintenance task chart below:

IMPORTANT! Calibrate the QuantStudio™ 12K Flex System at the same ambient temperature at which you will run experiments. Extreme variations in ambient temperature can affect the heating and cooling of the QuantStudio™ 12K Flex System and, in extreme cases, influence experimental results.

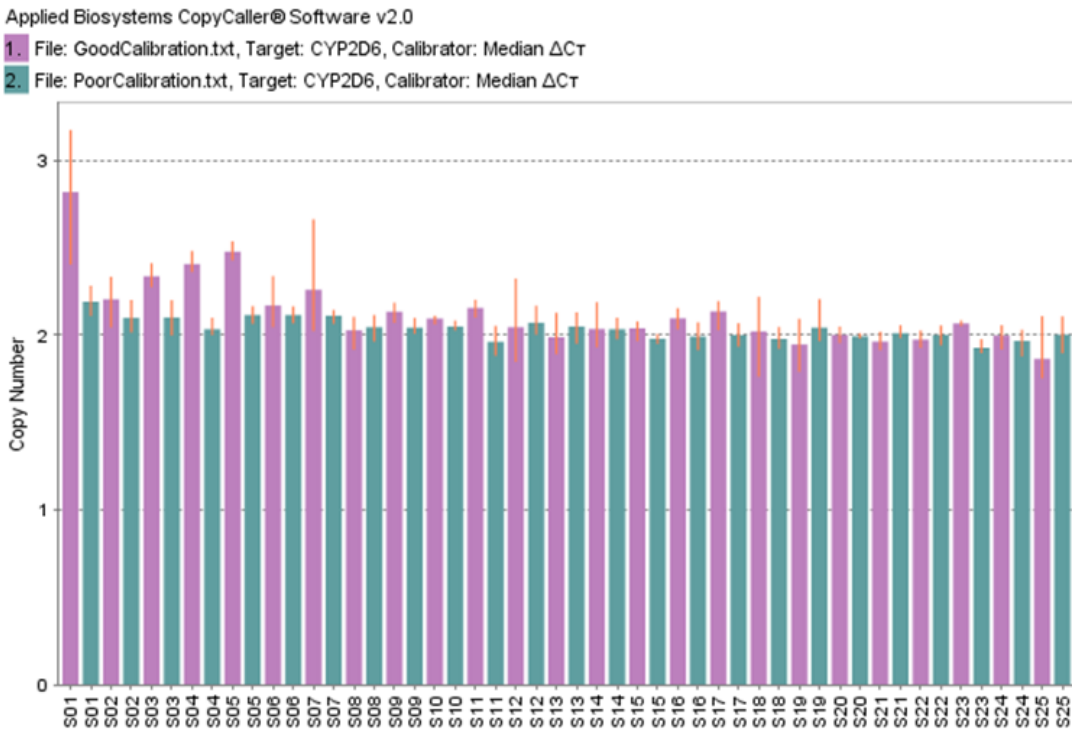
IMPORTANT! Do not use organic solvents to clean the QuantStudio™ 12K Flex System.

Table 1 Multi-well plate and array card sample block maintenance

Frequency	Maintenance task
Weekly	Check the computer disk space. If necessary, archive or back up your experiment files and instrument settings.
	Power off the computer that controls the QuantStudio™ 12K Flex System, then after 30 seconds, power on the computer.
	Clean the surface of the QuantStudio™ 12K Flex System with a lint-free cloth.
	Perform a QuantStudio™ 12K Flex Instrument self test.
Monthly	Perform a background calibration.‡
	Run disk cleanup and disk defragmentation.
Annually	Perform a regions of interest (ROI) calibration.
	Perform a background calibration.
	Perform a uniformity calibration.
	Perform a dye calibration.
	Perform a normalization calibration.
	Perform an instrument verification run.
As needed	Decontaminate the QuantStudio™ 12K Flex System.
	Replace the QuantStudio™ 12K Flex System fuses.
	Update the Windows® operating system.
	Update the QuantStudio™ 12K Flex Software and firmware.

‡ You can perform a background calibration to check for contamination. If any parts of the optics are replaced or moved, you must perform all calibrations, including an RNase P instrument verification run.

Expired or improper instrument calibration can lead to inaccurate copy number results. The figure below compares copy number calculations for a properly calibrated instrument (green bars) versus one that has poor calibration files (purple bars). Data for the experiment with poor calibration resulted in 4 samples with incorrect or ambiguous calls.



CHAPTER 3: Sample collection and preparation

Buccal Swab Types

High DNA yields from buccal swabs are not needed specifically for copy number testing, but are needed for the PGx workflow overall. For this reason, it is important to take steps to maximize DNA yield, including use of the recommended buccal swab types and swabbing protocol. The CNV analysis portion of the workflow requires using 5 ng/uL of high quality DNA sample stocks. The presence of PCR inhibitors, from food or other contaminants in the sample preparation, can negatively impact copy number call accuracy.

The following buccal swab types are recommended:

- 4N6FLOQSwabs™ (p/n 4473979, Life Technologies)
- PurFlock Ultra® Flocked Swab (p/n 22-025-192, Fisher Scientific)
- HydraFlock® Flocked Swab (p/n 25-3306-H, Puritan Medical Products)
- Sterile Foam Tipped Swabs (Puritan, Cat. no. 25-1506 1PF)

The following swab types are not supported:

- Generic Cotton
- Generic Polyester

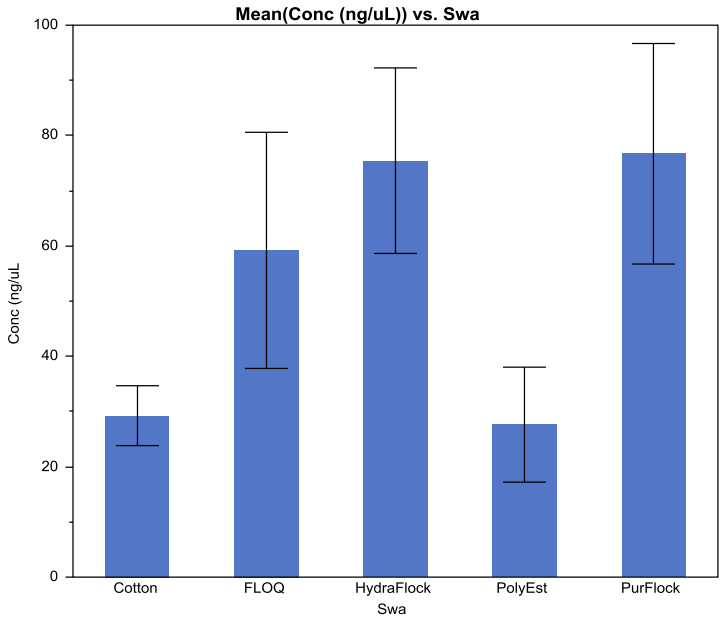


Figure 1: Cotton and polyester swab yields drop below 50 ng/uL minimum. Note: Each error bar is constructed using 1 standard error from the mean. N=7

Sample Collection Best Practices

Swab If using an external collection site, upon receipt of samples, ensure that the buccal swab type returned from

Type the collection site matches the type that was provided

Procedure Swabbing procedure is critical to avoid inclusion of PCR inhibitors and to ensure sufficient DNA yields

1. Perform a water rinse prior to collection to remove residual food, beverage, tobacco or other contaminants that can contain PCR inhibitors and can reduce DNA yields.
2. Ensure the proper collection techniques are performed:
    a. Wash hands and then carefully remove the swab from its packaging (figure 2a). Note: Do not fully separate the 2 sides of the wrapper. The packaging will be reused in step c.
    b. Vigorously scrape the inside of each cheek for 30 seconds (1 minute total).
    c. Let the swab air dry for 1 minute and then carefully place it back into its original packaging (figure 2c).

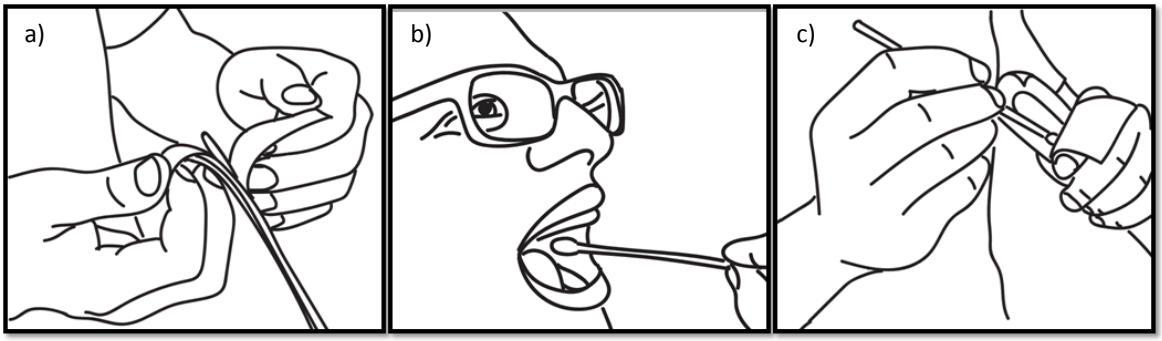


Figure 2: Swab collection instructions

3. Verify that collection technicians are properly trained in swabbing techniques

Storage Conditions Proper swab storage conditions

1. Buccal swab samples can be safely stored for up to 3 weeks at either room temperature or at -20°C before extraction with the MagMax™ DNA isolation system
2. Verify swabs are equilibrated to room temperature prior to processing if frozen for storage
3. Store Buccal swabs in wax paper pouches versus plastic tubes. Bacterial growth in sealed plastic tubes can cause DNA degradation
4. Record the storage time and temperature conditions for all samples

Effect of PCR inhibitors

PCR inhibitors are present in the food we eat and can be carried over during the sample preparation process. Furthermore, PCR inhibitors from the purification (such as alcohols, phenols, and salts) can make their way into the final DNA sample solution. During the PCR process, these inhibitors can delay amplification and/or distort C_T values, as well as reduce the dye fluorescence. Since copy number values are derived from an exponential function of the difference between the FAM Target C_T and the VIC Target C_T (ΔC_T), small deviations in the ΔC_T can lead to inaccurate or ambiguous copy number values. The figures below illustrate a drastic example where food carried over from swab sample B10 delayed the VIC target's C_T more than the FAM C_T (figure 4c). Compared to uninhibited sample B09, the VIC amplification curve is shallower for sample B10. This is indicative of PCR inhibition and resulted in an incorrect copy number of 7, whereas the sample is known to have 2 copies for this gene.

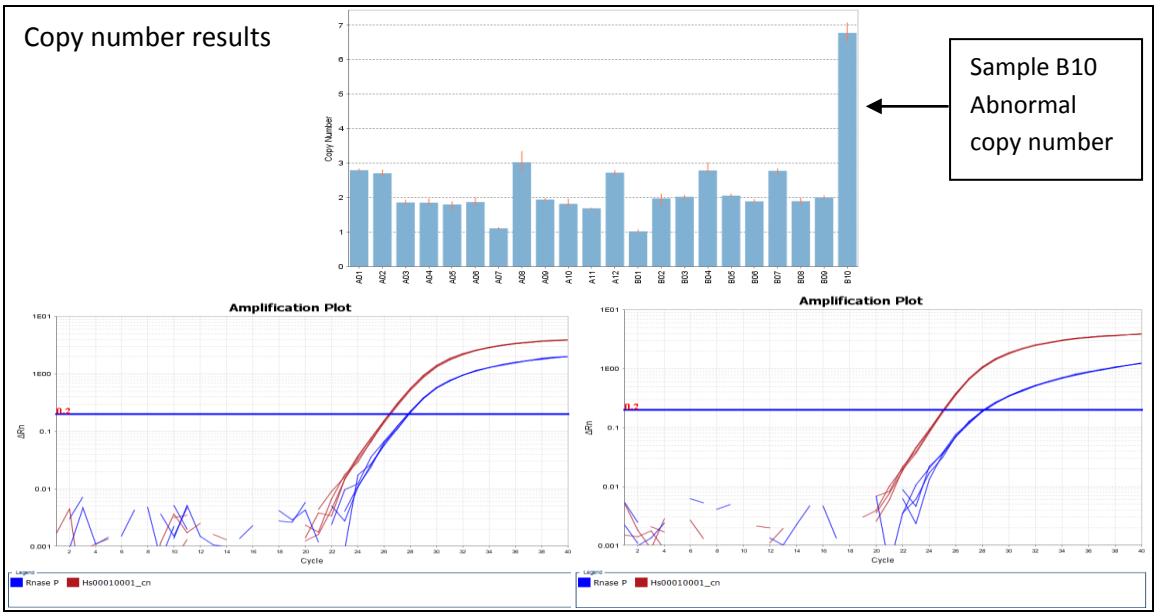


Figure 3: Sample B09 with similar amplification efficiency from VIC and FAM: ΔC_T= -1.42

Sample B10 with shallow VIC amplification (blue curve) caused by inhibition: ΔC_T = -3.18

Aberrant copy calls

CYP2D6 exon 9 copy number calls of 4 or higher are possible but are rare

If one or more test samples are providing aberrant copy numbers relative to other samples, or with multiple assays, the sample may contain PCR inhibitors.

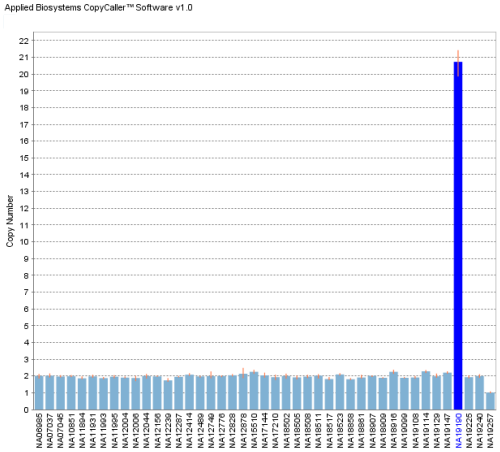


Figure 4: Sample NA19190 with copy number >3 is an aberrant copy number relative to other samples in this study

Presence of inhibition

Presence of inhibitors can be observed in the amplification curves for the copy number or reference assay which may look weak or different in shape compared to other samples.

Dilute samples to remove inhibitors

To check for the presence of inhibitors, and to potentially generate usable copy number data for the sample, dilute the sample preparation. Loading a 1:10 and/or 1:20 dilution of the normalized 5 ng/ul sample has worked for some cases of inhibition.

Sample Name	Average FAM	Average VIC	Δ C_T
NA19190_20ng	32.97	Undetermined	---
NA19190_10ng	27.21	29.68	-2.47
NA19190_5ng	27.06	27.96	-0.9
NA19190_2.5ng	27.96	28.63	-0.67

Figure 5: Sample dilution leads to recovery of VIC amplification in sample NA19190 by decreasing the concentration of inhibitors. Note that inhibitors can differentially impact target and reference assays impacting the ΔC_T value and copy number calculation.

DNA isolation using the MagMAX™ DNA Multi-Sample Ultra Kit

Reagent purity

Isopropanol and Ethanol should be of molecular grade or higher

Stability of Proteinase K

Prepare the Proteinase K (PK) Mix just before use. Do not place PK Buffer or PK Mix on ice to avoid precipitation.

Elution Buffers

A Elution Buffers 1 & 2 must be added at the appropriate steps in the protocol for maximum DNA recovery and proper final pH for downstream applications.

- If the order of addition is switched (Elution buffer 2 added first and Elution buffer 1 added second), DNA recovery and yields will be reduced
- If Elution buffer 1 is used twice, the final pH of the sample will be too basic and will not work in downstream qPCR

If Elution buffer 2 is used twice, DNA recovery and yields will be reduced.

Quantitation methods

Quantifying each DNA sample is an important step in quality control. Quantification allows the user to normalize the DNA and apply the same amount of DNA to each copy number reaction, improving consistency in the results. Also, quantification alerts users to problematic samples before downstream analysis

Multiple methods are available for DNA quantification:

Recommended Method: TaqMan® Copy Number Reference Assay RNase P

Use the TaqMan® RNase P Copy Number Reference Assay (Cat. No. 4403326) to specifically quantify human genomic DNA. The DNA quantification can be calculated from CT values.

Sample DNA Concentration Calculation:

1. Mean ΔC_T = Mean C_T_sample – Mean C_T_calibrator
2. RQ = $2^{-\text{Mean } \Delta CT}$
3. Concentration_sample = RQ x Concentration_calibrator

Recommended calibrator sample: Control DNA (from CEPH Individual 1347-02) (Cat. No. 403062)

Alternative 1: Fluorometric analysis can be used to quantify all double stranded DNA in a sample. This method is not species specific, so contamination from organisms can cause overestimation of the human DNA concentration.

Use a Qubit® dsDNA BR or HS Assay Kit.

Alternative 2: UV absorbance measurements can be used to assess both the concentration and the quality of the isolated DNA. Sample degradation and impurity (All DNA absorbs at 260nm, regardless of species) can cause overestimation of sample with UV absorbance concentrations methods.

Use a NanoDrop® or other comparable instrument. Pure genomic DNA should have an A260/A280 ratio of approximately 1.6–2.0.

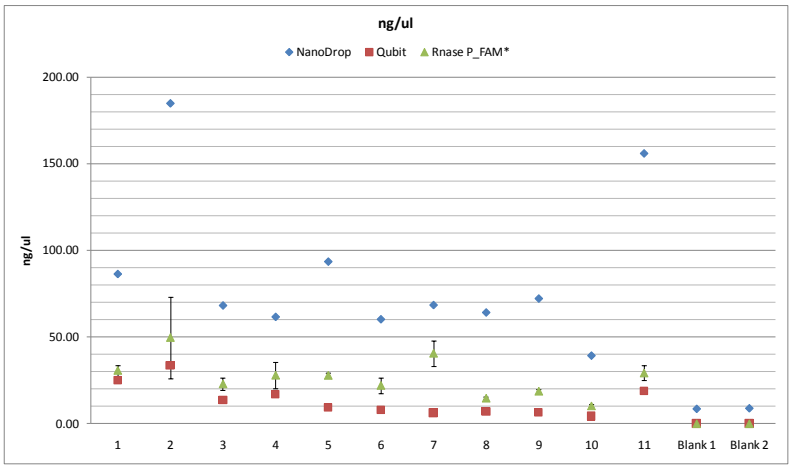


Figure 6: TaqMan RNase P and Qubit® dsDNA measurement report comparable results in the study of 11 samples while the Nanodrop overestimates sample concentration.

DNA Normalization

Normalize DNA to the recommended 5 ng/uL concentration; use 2 uL reaction for 10 ng total input

- The DNA amount in each technical replicate, and in each sample on a single plate, should be the same (i.e. 10 ng)
- Correct gDNA input may be verified by expected VIC C_T values & FAM C_T values (if FAM assay detects 2 copies)

IMPORTANT! A large C_T spread between samples (e.g. C_T range > 4) indicates that variable amounts of DNA were loaded per well; this could negatively impact copy number analysis.

CHAPTER 4: Experiment setup conditions

Sample Types

All gDNA samples that will be analyzed together should be run on the same plate and ideally should be of the same sample type and sample preparation method. Different sample types and samples prepared by different methods, if not sufficiently pure, can amplify with different efficiencies.
Note: a small number of control samples (e.g. a few Coriell samples on a 384-well plate) should not strongly impact analysis of buccal or blood derived samples)

Control Samples

DNA samples with a known copy number for the target of interest should be run, if possible, with test samples of unknown copy number. A sample of known copy number to be used as the calibrator, or reference, sample should be of the same sample type as the test samples.

Sample Number

CopyCaller assigns confidence & |z-score| metrics to predicted sample copy number calls when at least 7 samples of the same target copy number are available in the data set. In a typical sample set, we recommend using 12 unique samples.

Template Replicates

To generate reliable copy number calls, it is strongly recommended that four replicates are used for each gDNA sample.

Master Mix Selection

TaqMan® Genotyping Master Mix, Catalog no. 4371355, is the recommended master mix for optimal performance with TaqMan® SNP Genotyping Assays and TaqMan® Copy Number Assays run on 96- or 384-well plates. Alternatively, if samples contain PCR inhibitors resulting in poor quality or incorrect copy number calls, inhibitor-tolerant master mixes may perform better. Contact your local Field Application Scientist or call technical support to discuss alternative master mix options.

Protocol

Run experiments according to the validated TaqMan Copy Number Assays Protocol:

- 10 ng gDNA/10 uL reaction
- Universal Cycling Conditions
- Analyze the real time PCR data for both copy number and reference assays using a manual C_T threshold of 0.2 and auto baseline settings
- Note: a C_T threshold of 0.1 may be used if it crosses the linear phase of the amplification curves better than the 0.2 setting

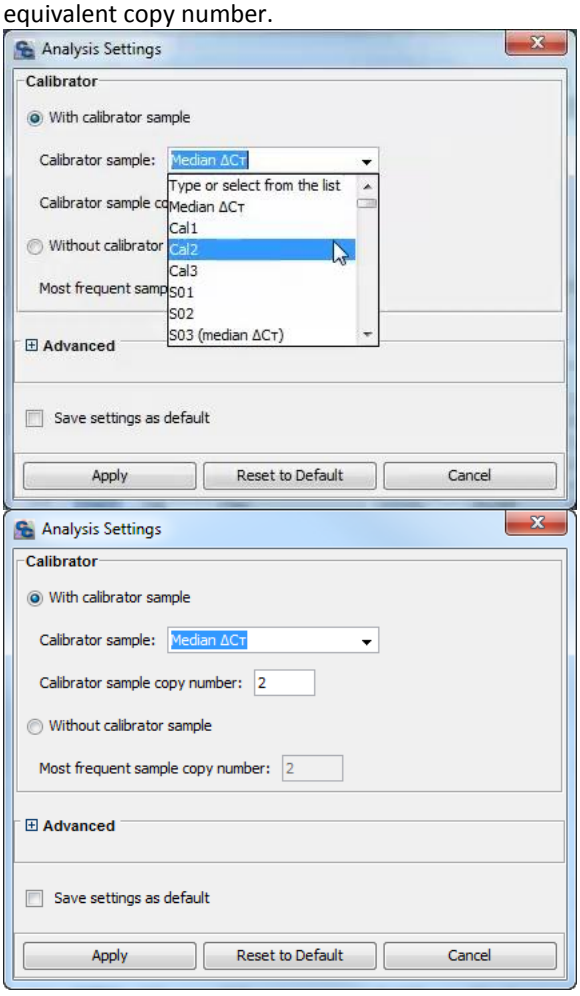
CHAPTER 5: Copy Number Analysis using CopyCaller® Software

General Guidelines

All samples and all technical replicates to be analyzed together must be run on the same plate

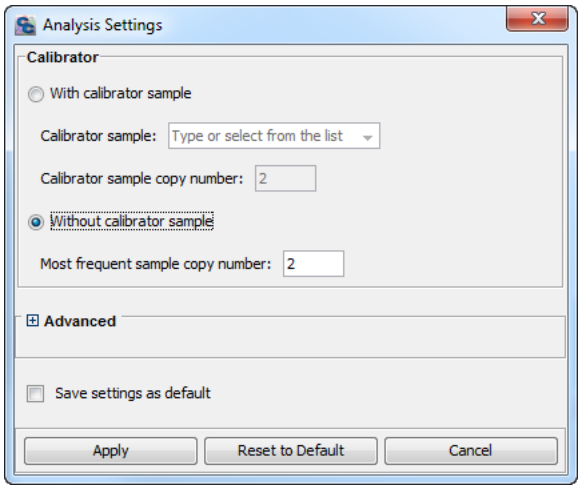
Calibration Method

1. With Calibrator Sample: two options are available for calibration samples
   - Median ΔC_T value: Best practice when the majority of samples in the analysis have an equivalent copy number.



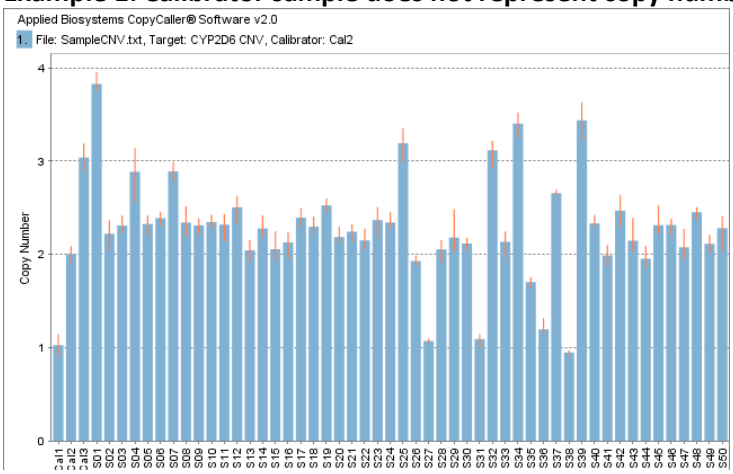
   - Characterized calibrator sample of known copy number
     IMPORTANT! a calibrator sample should be of the same sample type and preparation method as the test samples and accurately represent other samples in the same copy number group (i.e. have the same calculated copy number). Coriell samples should not be used as the calibrator for gDNA samples extracted from buccal or blood.

2. Without Calibrator Sample: A maximum likelihood algorithm fits the data to a model
   Enter into the 'Most Frequent Sample Copy Number' field the number of copies that the majority of samples in the analysis are expected to contain.

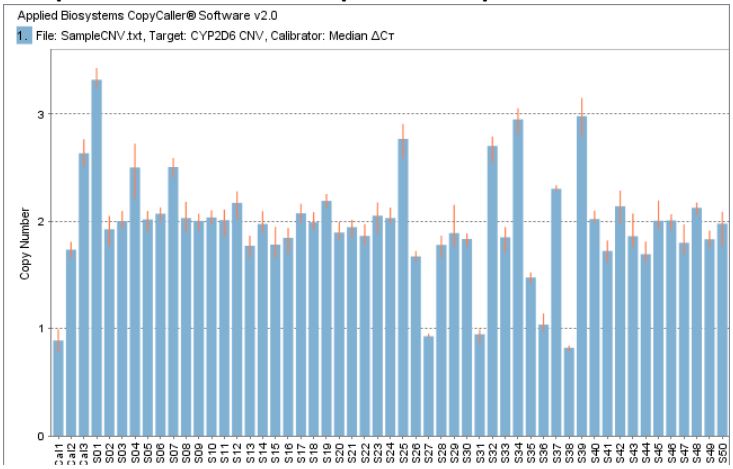


IMPORTANT! the data is fit to a model to determine the predicted copy number. Thus, the predicted copy number is not necessarily equal to the rounded calculated copy number as it is in the 'with calibrator method' (e.g. a sample with a between integer calculated copy number of 0.4 may be called as a low confidence 1-copy versus a 0-copy sample).

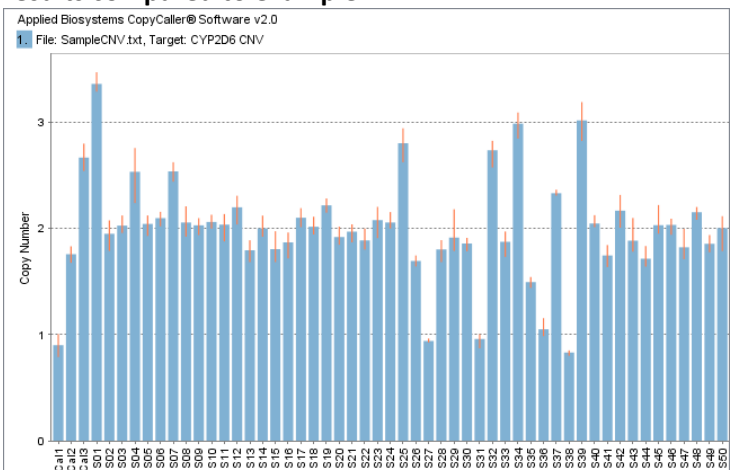
Example 1: Calibrator sample does not represent copy number 2 samples:



Example 2: Median ΔC_T calibrator improves results for copy number 2 samples compared to calibrator sample in example 1



Example 3: Without calibrator sample analysis method improved the quality of the results compared to example 1



Confirm Duplicate Calls

CYP2D6 duplication alleles occur in most studied populations at a frequency of roughly 1-7%[4,5]. If the number of samples with duplications is extraordinarily high and/or samples with calls >3 are present, we recommend reviewing the sample data quality to double check the accuracy of the calls by performing the following checks:

- Ensure that all amplification curves have normal geometric shapes. For guidance, refer to the 'Effect of PCR inhibitors' section.
- Confirm that samples of similar or identical concentrations have comparable VIC C_T values. Large deviations from this expected pattern suggest a systematic problem.
- To exclude samples with insufficient DNA from analysis, set the VIC C_T threshold to 32(default value) or below. Samples above the set threshold will be omitted from CopyCaller Analysis.
- Inspect the calculated copy number values to ensure that they are close to integer values.
- Review the confidence and absolute z-score values for the predicted copy number calls
- Confirm that samples cluster into well-defined, well-separated copy number groups in the ΔC_T plot view

Copy number call quality metrics

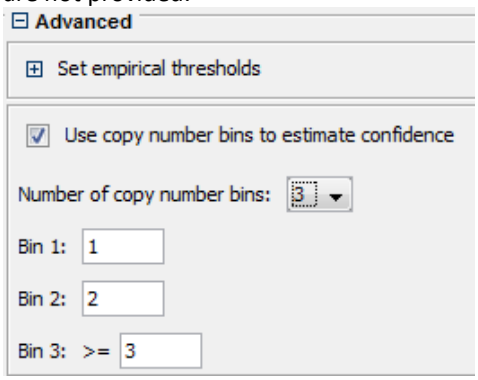
Confidence value evaluation

CopyCaller assigns confidence & |z-score| metrics to predicted sample copy number calls when at least 7 samples of the same target copy number are available in the data set. The confidence score evaluates whether a sample was assigned correctly to its predicted copy number group compared to any other group, whereas the |z-score| measures how close the sample ΔC_T value is to the mean of the median ΔC_T value for its copy number group. Passing the default 95% confidence threshold, plus a |z-score| of <1.75, is very achievable for good quality samples having 1-3 copies and is more difficult to achieve for lower quality samples carrying duplications, as well as for samples with >3 copies. The confidence values are calculated for >=3 copy samples based on their mathematically predicted ΔC_T values relative to the copy number 2 group ΔC_T median value. As copy numbers increase, confidence progressively decreases due to decreased separation of ΔC_T subdistributions (= copy number groups). In addition, if sample data is of lower quality, some sample ΔC_T values may not well fit the copy number group model and may have low confidence values. Consider manually passing samples that fall below the 95% confidence level if they have calculated copy numbers that are close to integer values and if they cluster with passing samples of the same copy number group, which can be evaluated by examining the ΔC_T plot (see figure below). Note that |z-score| threshold guidance only apply to high confidence calls (>95%). Consider samples with |z-

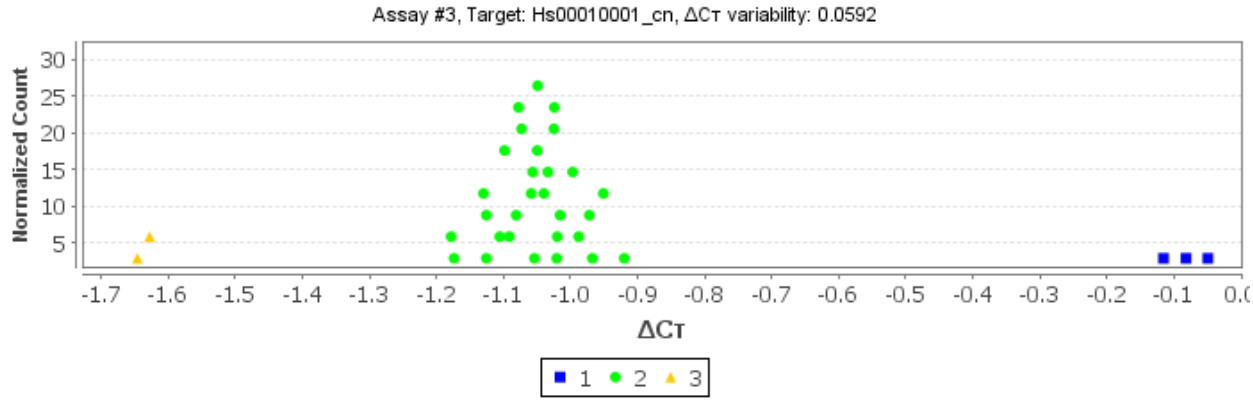
scores| <1.75 as passing, >= 2.65 as failing, and in between these values as 'passing with caution'. In the latter case, review the sample replicates for outliers, large variance in replicate values, or other anomalies.

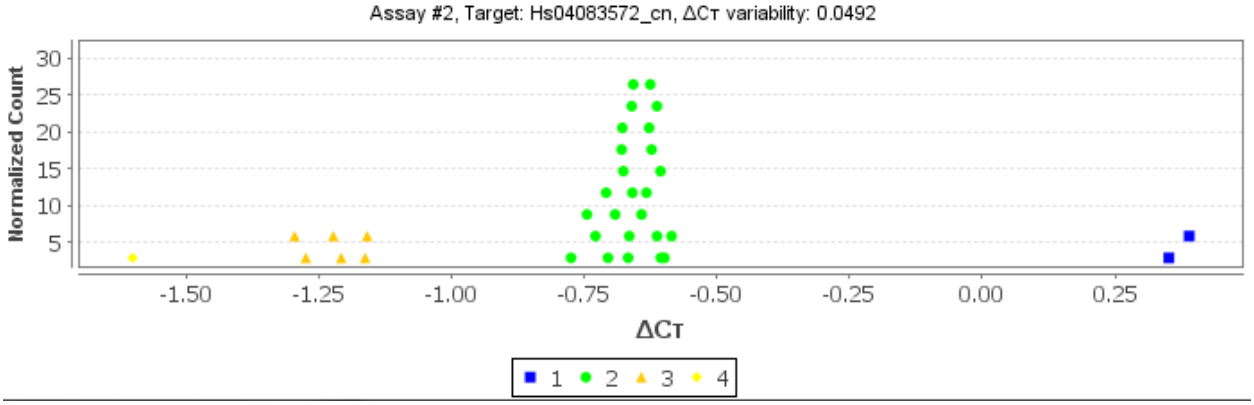
Copy number bins and confidence values

Samples with three or more copies of most CYP2D6 alleles may have the same metabolizer phenotype, thus it may be sufficient to report the presence of a duplication versus discreet copy numbers of 3 or 4 or more. In this case, confidence values for samples with >=3 copies may be improved by using the Advanced settings copy number binning option to estimate confidence. E.g. for CYP2D6 analysis, 3 bins of 1, 2, >=3 could be used. When this option is used, |z-score| values, which are a property of normal distributions, are not provided.

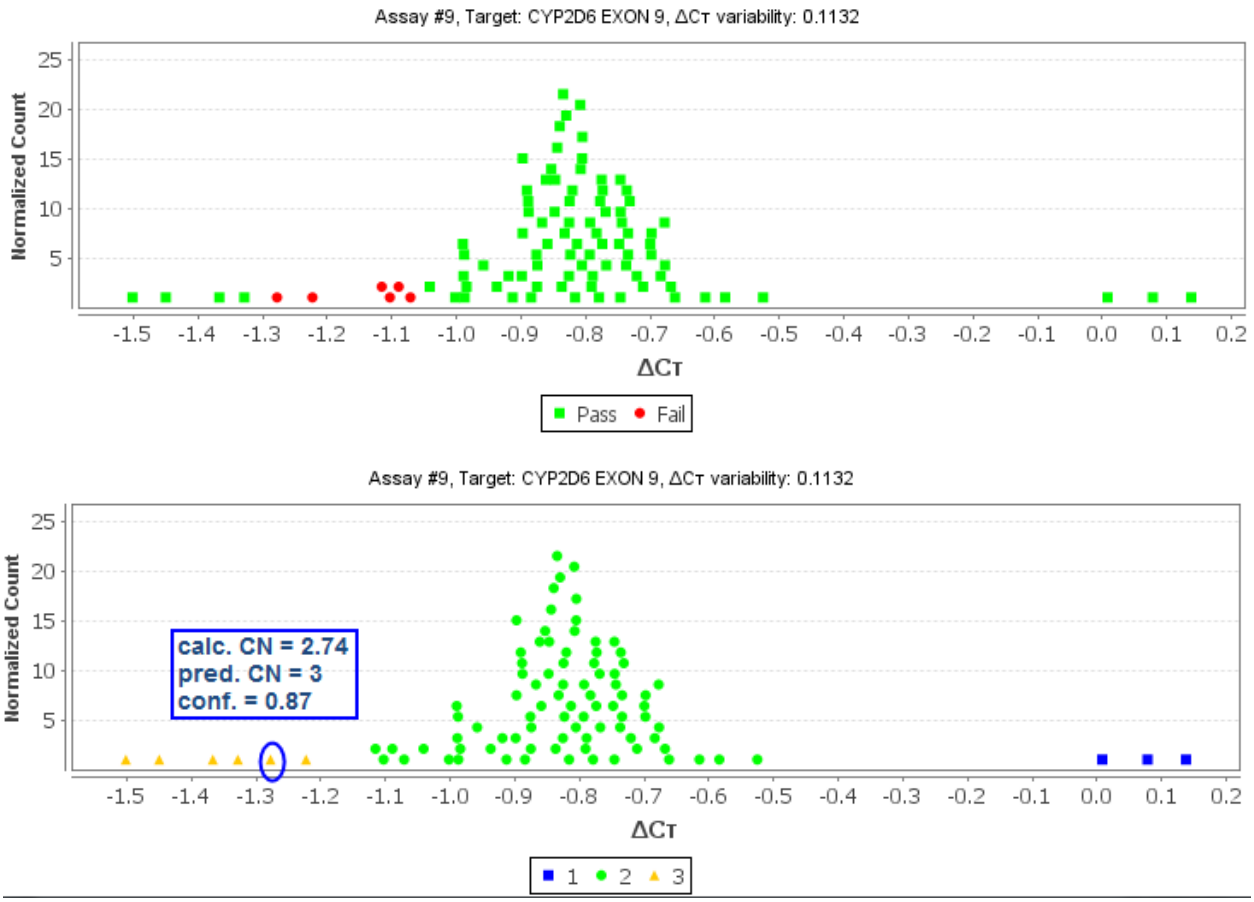


Shown are ΔC_T plots for buccal sample DNAs run with the CYP2D6 exon 9 and intron 2 assays, showing excellent separation between copy number 1-3 and 1-4 groups, respectively.





Shown is the ΔC_T plot for an experiment using buccal sample DNAs run with the CYP2D6 exon 9, wherein some samples did not pass the default quality metrics (top plot: Confidence view; bottom plot: Predicted Copy Number view). Note that the sample circled in the lower plot has a below-passing confidence value of 0.87 though it has a calculated copy number value of 2.74 and groups well with passing 3-copy samples; it is a good candidate for manual passing.



Other considerations:

Examine the reference assay (RNase P or TERT) VIC C_T values across the sample set to ensure that they are relatively close to one another and are within the expected range (i.e. ~25-28 C_T s depending on the instrument used). A large C_T spread between samples (e.g. C_T range > 4) indicates that variable amounts of DNA were loaded per well; this could negatively impact copy number analysis.

- Consider removing high and low end samples and reanalyze; this may improve the copy number analysis when the median d C_T sample or value is used as the calibrator, or the no calibrator method is used.
- The Advance setting: 'Exclude wells with VIC C_T greater than' value can be lowered from the default VIC threshold value of 32 to 30 or so to remove lower quality data points.

References

1. Gaedigk A, et. al. (2006) CYP2D6*36 gene arrangements within the cyp2d6 locus: association of CYP2D6*36 with poor metabolizer status. Drug Metab Dispos. 34:563-9.
2. Soyama A, et. al. (2006). Sequence-based analysis of the CYP2D6*36-CYP2D6*10 tandem-type arrangement, a major CYP2D6*10 haplotype in the Japanese population.
3. Ramamoorthy A, et. al. (2010) Differential quantification of CYP2D6 gene copy number by four different quantitative real-time PCR assays. Drug Metab Pharmacokinet. 21:208-16.
   (this study confirms use of the TaqMan CYP2D6 copy number assays for their intended purpose).
4. Bradford, L.D. (2002) CYP2D6 allele frequency in European Caucasians, Asians, Africans and their descendants. Pharmacogenomics 3:229-243.
5. Sistonen J, et. al. (2007) CYP2D6 worldwide genetic variation shows high frequency of altered activity variants and no continental structure. Pharmacogenet Genomics. 17:93-101.