

nCounter[®]

Data Analysis Guidelines for Copy Number Variation (CNV)

NanoString Technologies, Inc.

530 Fairview Ave N
Suite 2000
Seattle, Washington 98109

www.nanostring.com

Tel: 206.378.6266
888.358.6266

E-mail: info@nanostring.com

Molecules That Count[®]

Translational Research • Gene Expression • miRNA Expression • Copy Number Variation

MAN-C0014-02

FOR RESEARCH USE ONLY. Not for use in diagnostic procedures.**Intellectual Property Rights**

This nCounter® Analysis System manual and its contents are the property of NanoString Technologies, Inc. (“NanoString”), and is intended solely for the use of NanoString customers, for the purpose of operating the nCounter Analysis System. The nCounter Analysis System (including both its software and hardware components) and this User Guide and any other documentation provided to you by NanoString in connection therewith, are subject to patents, copyright, trade secret rights and other intellectual property rights owned by, or licensed to, NanoString. No part of the software or hardware, may be reproduced, transmitted, transcribed, stored in a retrieval system, or translated into other languages without the prior written consent of NanoString.

Limited License

Subject to the terms and conditions of the nCounter Analysis System contained in the product quotation, NanoString grants you a limited, non-exclusive, non-transferable, non-sublicensable, research use only license to use the proprietary nCounter Analysis System only in accordance with the manual and other written instructions provided by NanoString. Except as expressly set forth in the terms and conditions, no right or license, whether express, implied or statutory, is granted by NanoString under any intellectual property right owned by, or licensed to, NanoString by virtue of the supply of the proprietary nCounter Analysis System. Without limiting the foregoing, no right or license, whether express, implied or statutory, is granted by NanoString, to use the nCounter Analysis System with any third party product not supplied or licensed to you by NanoString, or recommended for use by NanoString in a manual or other written instruction provided by NanoString.

Trademarks

NanoString Technologies, NanoString, nCounter and Molecules That Count are registered trademarks or trademarks of NanoString Technologies, Inc., in the United States and/or other countries. All other trademarks and/or service marks not owned by NanoString that appear in this manual are the property of their respective owners.

Copyright

© 2011 NanoString Technologies, Inc. All rights reserved.

CONTENTS

Preface	4
Conventions Used.....	4
Contact Information	4
CHAPTER 1: Introduction	5-15
nCounter Custom CNV Data Output	5
Raw Data.....	5
Code Classes.....	6
Basic CNV Data Analysis Workflow	6
Assay Quality Control Check.....	6
Normalization to Invariant Probes	8
Calculating Copy Number Estimates.....	10
X and Y Chromosome Copy Number Calculation.....	11
Averaging Copy Number Estimates by Genomic Region	12
Generating Integer Copy Number Calls.....	12
Reference Sample Adjustments	14
Adjusting the Copy Number Estimate Calculation.....	14
Reference Documents and Support	15

PREFACE

Conventions Used

The following conventions are used throughout this manual and are described below for your reference:

Note Types

Special font formatting is used in this manual. Such formatting conventions are used in specific instances as described below:

TIP Information contained in a Tip may offer helpful suggestions, alternative procedures, methods and/or shortcuts.

NOTE This note type emphasizes general information.

IMPORTANT  This note type presents essential content indicating that the potential exists for assay failure, diminished data quality, and/or a loss of data if the information presented is ignored.

WARNING  This note type indicates that a potential hazard to your personal safety, or the potential for equipment damage exists.

BOLD When appearing in text or in a procedure, the bold text serves to highlight a specific button, key stroke, or menu option available.

- **Bold** text may appear elsewhere to highlight important text or terms.
- **Green** text is used to help the reader identify active hyperlinks.

ITALICS Used to emphasize an important word or expression within the text.

- Formatting of a book title, journal, or other documentation.
- Used to indicate the special or unusual meaning of a word or phrase.

Contact Information

NanoString Technologies, Inc.

530 Fairview Ave N
Suite 2000
Seattle, Washington 98109 USA

Tel: 206.378.6266
888.358.NANO (6266)

Fax: 206.378.6288

E-mail: support@nanosting.com

1

Introduction

The nCounter® Custom Copy Number Variation (CNV) Assay utilizes NanoString's unique direct and multiplexed detection of nucleic acids in solution to generate estimates of copy number variation for hundreds of loci in a single reaction. Each NanoString Reporter and Capture Probe pair is complementary to ~100 nt of contiguous genomic DNA sequence at a user-specified locus. Genomic DNA is fragmented into small pieces (200-800 bp) and denatured to produce single strands. The Custom CNV CodeSet is then hybridized to the fragmented, denatured DNA sample in a single multiplexed reaction (up to 800 genomic loci per CodeSet). Hybridized DNA-CodeSet complexes are purified by the fully automated nCounter® Prep Station, and Reporters are counted by the nCounter® Digital Analyzer.

The results of an nCounter® Custom CNV Assay experiment are compiled and displayed using the CNV Collector Tool software included with the assay kit. Detailed instructions for using this tool are in the [CNV Collector Tool User Manual](#) provided with the software.

The following Data Analysis Guidelines for CNV are intended as a supplement to the [nCounter® CNV Collector Tool User Manual](#). They provide instructions and additional information for those who wish to do further QC and/or data manipulations with the data output from the CNV Collector Tool, such as:

- Modify the normalization method
- Apply additional assay quality control metrics
- Change reference samples within a single data set
- Use reference samples with known copy numbers that differ from 2

nCounter Custom CNV Data Output

Raw Data

The basic output of the Custom CNV Assay is a spreadsheet containing the CodeSet probe identifiers, sample identifiers, and the digital 'counts' recorded for each probe in each sample. This file, referred to as a Reporter Code Count (RCC) file, can be uploaded to the nCounter CNV Collector Tool for automated normalization and analysis.

FIGURE 1.1: Report Code Count (RCC) file

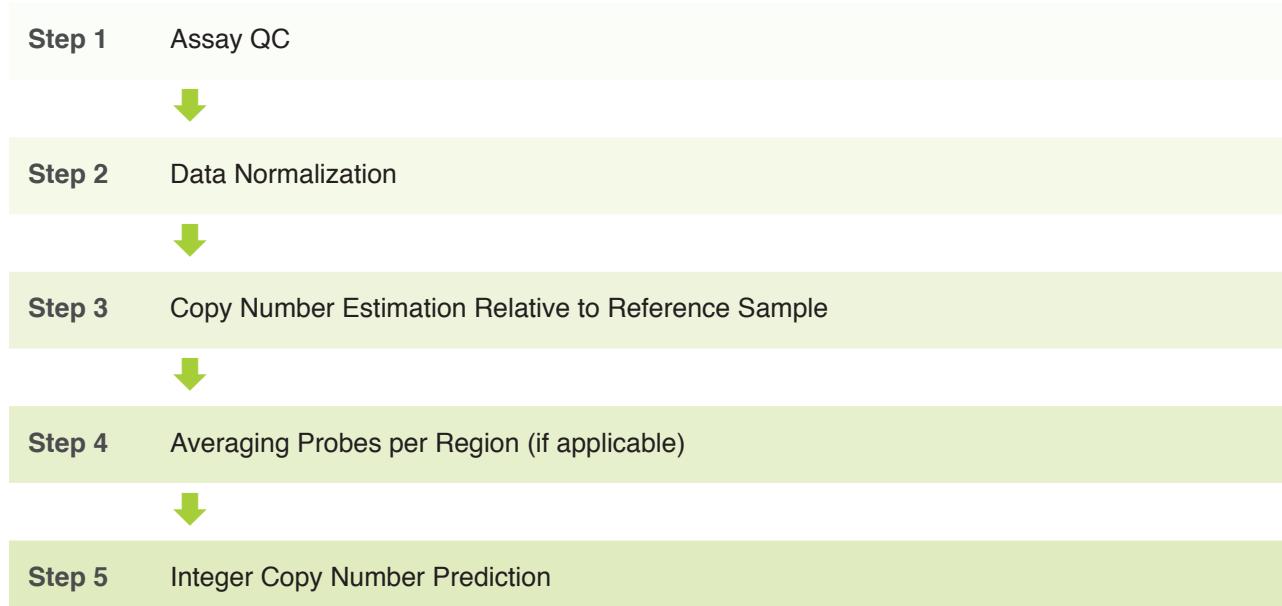
	A	B	C	D	E	F	G	H	I
19	Reporter Counts			Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6
20	Code Class	Name	Accession						
21	Positive	POS_A(128)	ERCC_00117.1	21918	25182	20426	26194	25293	22480
22	Positive	POS_B(32)	ERCC_00112.1	6608	7521	4640	6107	6070	5367
23	Positive	POS_C(8)	ERCC_00002.1	1405	1627	1351	1623	1637	1464
24	Positive	POS_D(2)	ERCC_00092.1	366	429	284	406	402	331
25	Positive	POS_E(0.5)	ERCC_00035.1	139	161	242	261	270	241
26	Positive	POS_F(0.125)	ERCC_00034.1	65	72	87	80	79	80
27	Negative	NEG_A(0)	ERCC_00096.1	0	1	6	8	14	7
28	Negative	NEG_B(0)	ERCC_00041.1	3	7	9	11	13	15
29	Negative	NEG_C(0)	ERCC_00019.1	8	6	9	10	9	10
30	Negative	NEG_D(0)	FRCC_00076.1	3	2	16	20	15	12

Code Classes

1. POSITIVE	nCounter Custom CNV CodeSets contains 6 positive dsDNA control probes, each targeting a unique DNA sequence present in every assay. The concentrations of DNA target range from 0.125 fM to 128 fM in the hybridization reaction.
2. NEGATIVE	nCounter Custom CNV CodeSets contain 8 negative control probes, for which there is no DNA target present in the hybridization reaction. These probes monitor the nonspecific, or background, counts for every assay.
3. INVARIANT	Each CodeSet contains a set of 10 probes (INV) designed to autosomal genomic regions predicted not to contain common CNVs.
4. RESTRICTION SITE	Custom CNV CodeSets contain four control probes to monitor the efficiency of the DNA fragmentation and denaturation steps of the CNV Assay Protocol. Probes A and B are designed to a DNA sequence containing an Alul restriction site, and will return low count when the Alul fragmentation is working correctly. Probes C and D are designed to sequences that lack an Alul restriction site, and serve as controls for the presence of target DNA in the sample preparation step. When used according to the assay manual, these controls help identify problems in restriction enzyme fragmentation and denaturation steps of the assay.
5. ENDOGENOUS	The Custom CNV Assay probes specified by the user, designed to specific regions of the genome.

Basic CNV Data Analysis Workflow

FIGURE 1.2: CNV Data Analysis Workflow



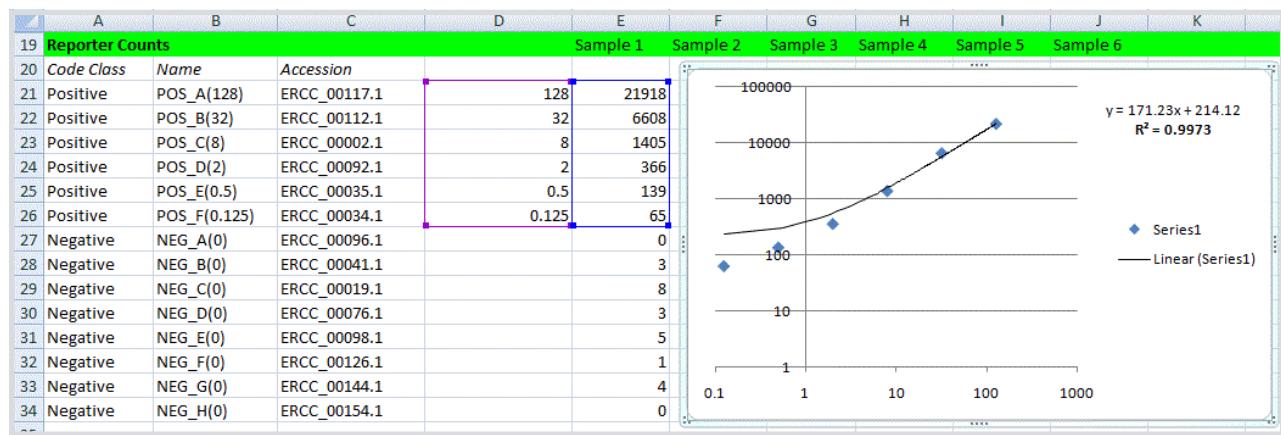
Assay Quality Control Check

When used according to the assay manual, each Custom CNV Assay contains controls which monitor hybridization efficiency, sample DNA fragmentation denaturation, and sample DNA input amount. Before continuing with CNV analysis, first look at the Raw Data output file from the CNV Collector to gauge the performance of the assay.

Positive Controls

The positive control (POS) DNA targets are added in a linear titration to each codeset to generate a standard curve. The final concentration in the hybridization of each target (in fM) is indicated in parentheses next to the POS probe identifier. To check the linearity of the POS control standard curve, insert a new column in the Raw Data spreadsheet to the left of the first assay data column. Add the POS target concentrations to the first 6 rows of this new column. Then create a scatter plot of the target concentration vs. raw data for each target ([Figure 1.3](#)).

FIGURE 1.3: Positive Controls



In the scatter plot shown in Figure 1.3, the axes are represented in logarithmic scale and the linear regression (R^2) is shown in the inset. The R^2 value for the POS control probes should be > 0.95 .

Restriction Fragmentation Controls

The Custom CNV Assay Kit comes with a set of four DNA controls that, when added to your genomic DNA sample prior to fragmentation, will monitor the efficiency of enzymatic digestion and heat denaturation.

The DNA targets for probes labeled RESTRICTIONBSITE+A and RESTRICTIONBSITE+B contain an Alul restriction site such that, after complete digestion, the target site will be cleaved by the enzyme and low probe count will be observed. The DNA targets for probes labeled RESTRICTIONBSITE-C and RESTRICTIONBSITE-D do not contain Alul sites, and will generate probe counts even in the absence of fragmentation. These targets will serve as controls for proper addition of the control DNA to the sample and proper heat denaturation. If the DNA sample is not denatured prior to hybridization, you will observe low counts (generally < 200) for RESTRICTIONBSITE-C and RESTRICTIONBSITE-D probes.

When the genomic DNA sample is completely digested with Alul enzyme and denatured, you should observe at least a 10-fold difference in counts between RESTRICTIONBSITE+ (Probes A and B) and RESTRICTIONBSITE- probes (Probes C and D):

FIGURE 1.4: Restriction Fragmentation Controls

50	RestrictionSite_RESTRICTIONBSITE+A + CTL Alul+A	RESTRICTIONBSITE	12	15	25	43	41	47
51	RestrictionSite_RESTRICTIONBSITE+B + CTL Alul+B	RESTRICTIONBSITE	10	15	54	71	83	47
52	RestrictionSite_RESTRICTIONBSITE-C - CTL Alul-C	RESTRICTIONBSITE	1255	1399	1144	1331	1295	1120
53	RestrictionSite_RESTRICTIONBSITE-D - CTL Alul-D	RESTRICTIONBSITE	975	1115	1336	1548	1541	1305

Normalization to Invariant Probes

Each nCounter Custom CNV CodeSet contains probes designed to invariant regions of 10 autosomes. It is assumed that these 10 regions will represent 2 chromosomal copies in a vast majority of samples analyzed. Therefore, normalizing data to the counts obtained from these 10 invariant probes should correct for any differences in sample to sample genomic DNA input arising from pipetting error or inaccuracies in DNA quantitation. Normalization is performed automatically by the nCounter CNV Collector Tool, but can be done manually as follows:

1. To normalize data to the Invariant probes, calculate the average count value for the 10 INV probes (AVE INV) in the first lane (sample) as shown in **Figure 1.5**:

FIGURE 1.5: Calculating the average count value.

	A	B	C	D	E	F	G	H	I	
19	Reporter Counts			Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6	
20	Code Class	Name	Accession							
21	Positive	POS_A(128)	ERCC_00117.1	21918	25182	20426	26194	25293	22480	
22	Positive	POS_B(32)	ERCC_00112.1	6608	7521	4640	6107	6070	5367	
23	Positive	POS_C(8)	ERCC_00002.1	1405	1627	1351	1623	1637	1464	
24	Positive	POS_D(2)	ERCC_00092.1	366	429	284	406	402	331	
25	Positive	POS_E(0.5)	ERCC_00035.1	139	161	242	261	270	241	
26	Positive	POS_F(0.125)	ERCC_00034.1	65	72	87	80	79	80	
27	Negative	NEG_A(0)	ERCC_00096.1	0	1	6	8	14	7	
28	Negative	NEG_B(0)	ERCC_00041.1	3	7	9	11	13	15	
29	Negative	NEG_C(0)	ERCC_00019.1	8	6	9	10	9	10	
30	Negative	NEG_D(0)	ERCC_00076.1	3	2	16	20	15	12	
31	Negative	NEG_E(0)	ERCC_00098.1	5	11	11	7	9	11	
32	Negative	NEG_F(0)	ERCC_00126.1	1	2	15	13	15	12	
33	Negative	NEG_G(0)	ERCC_00144.1	4	7	10	8	13	15	
34	Negative	NEG_H(0)	ERCC_00154.1	0	2	13	14	19	13	
35				=AVERAGE(D4:D9)	485	696	674	758	754	
36				average	AVERAGE(number1, [number2], ...)					
37				Norm.factor:	0.97	1.40	0.97	1.00	0.89	0.90
38										
39										
40	Invariant	_INVCONTROL-1 + Hg18 chr1:97239055-97239154	_INVCONTROL	985	632	906	856	880	995	
41	Invariant	_INVCONTROL-10 + Hg18 chr10:76411622-76411718	_INVCONTROL	382	226	414	388	435	440	
42	Invariant	_INVCONTROL-16 + Hg18 chr16:11026537-11026636	_INVCONTROL	433	356	495	500	575	571	
43	Invariant	_INVCONTROL-17 + Hg18 chr17:60617472-60617571	_INVCONTROL	792	530	732	761	813	832	
44	Invariant	_INVCONTROL-19 + Hg18 chr19:38818850-38818949	_INVCONTROL	656	464	715	653	702	694	
45	Invariant	_INVCONTROL-2 + Hg18 chr2:137221998-137222092	_INVCONTROL	605	414	636	586	739	663	
46	Invariant	_INVCONTROL-20 + Hg18 chr20:10504898-10504992	_INVCONTROL	488	320	511	527	566	538	
47	Invariant	_INVCONTROL-22 + Hg18 chr22:42509372-42509471	_INVCONTROL	895	637	828	847	991	936	
48	Invariant	_INVCONTROL-5 + Hg18 chr5:71261236-71261335	_INVCONTROL	913	690	887	837	956	991	
49	Invariant	_INVCONTROL-7 + Hg18 chr7:41541153-41541252	_INVCONTROL	800	577	838	789	919	878	
50	RestrictionSite	RESTRICTION SITE+A + CTL AluI+A	RESTRICTION SITE	12	15	25	43	41	47	

2. Calculate a normalization factor for each assay by first calculating the mean of the average INV count values (mean AVE INV) across all lanes you wish to analyze (**Figure 1.6**):

FIGURE 1.6: Calculating the mean of the average INV count values.

	A	B	C	D	E	F	G	H	I
19	Reporter Counts			Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6
20	Code Class	Name	Accession						
21	Positive	POS_A(128)	ERCC_00117.1	21918	25182	20426	26194	25293	22480
22	Positive	POS_B(32)	ERCC_00112.1	6608	7521	4640	6107	6070	5367
23	Positive	POS_C(8)	ERCC_00002.1	1405	1627	1351	1623	1637	1464
24	Positive	POS_D(2)	ERCC_00092.1	366	429	284	406	402	331
25	Positive	POS_E(0.5)	ERCC_00035.1	139	161	242	261	270	241
26	Positive	POS_F(0.125)	ERCC_00034.1	65	72	87	80	79	80
27	Negative	NEG_A(0)	ERCC_00096.1	0	1	6	8	14	7
28	Negative	NEG_B(0)	ERCC_00041.1	3	7	9	11	13	15
29	Negative	NEG_C(0)	ERCC_00019.1	8	6	9	10	9	10
30	Negative	NEG_D(0)	ERCC_00076.1	3	2	16	20	15	12
31	Negative	NEG_E(0)	ERCC_00098.1	5	11	11	7	9	11
32	Negative	NEG_F(0)	ERCC_00126.1	1	2	15	13	15	12
33	Negative	NEG_G(0)	ERCC_00144.1	4	7	10	8	13	15
34	Negative	NEG_H(0)	ERCC_00154.1	0	2	13	14	19	13
35									
36			mean INV counts	695	485	696	674	758	754
37			=AVERAGE(D35:I36)						
38			Norm.factor:	AVERAGE(number1, [number2], ...)	0	0.97	1.00	0.89	0.90

To generate a normalization factor, divide the mean average value (mean AVE INV from Step 2) by the average count value (AVE INV from Step 1) for each lane:

FIGURE 1.7: Generating a Normalization Factor

	A	B	C	D	E	F
19	Reporter Counts			Sample 1	Sample 2	Sample 3
20	Code Class	Name	Accession			
21	Positive	POS_A(128)	ERCC_00117.1	21918	25182	20426
22	Positive	POS_B(32)	ERCC_00112.1	6608	7521	4640
23	Positive	POS_C(8)	ERCC_00002.1	1405	1627	1351
24	Positive	POS_D(2)	ERCC_00092.1	366	429	284
25	Positive	POS_E(0.5)	ERCC_00035.1	139	161	242
26	Positive	POS_F(0.125)	ERCC_00034.1	65	72	87
27	Negative	NEG_A(0)	ERCC_00096.1	0	1	6
28	Negative	NEG_B(0)	ERCC_00041.1	3	7	9
29	Negative	NEG_C(0)	ERCC_00019.1	8	6	9
30	Negative	NEG_D(0)	ERCC_00076.1	3	2	16
31	Negative	NEG_E(0)	ERCC_00098.1	5	11	11
32	Negative	NEG_F(0)	ERCC_00126.1	1	2	15
33	Negative	NEG_G(0)	ERCC_00144.1	4	7	10
34	Negative	NEG_H(0)	ERCC_00154.1	0	2	13
35						
36			mean INV counts	695	485	696
37			average	677		
38			Norm.factor:	=D\$37/D36	1.40	0.97

3. Calculate INV-normalized counts for each probe. On a new sheet, generate normalized counts for each probe in the CodeSet by multiplying the RAW counts for each probe by the normalization factor for the lane as follows:

FIGURE 1.8

A	B	C	D	E	F	G	H	I
21	Reporter Counts		Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6
22	Code Class	Name	Accession					
23	Invariant	_INVCONTROL-1 + Hg18 chr1:97239055-97239154	_INVCONTROL	=`Raw data example`!D40*`Raw data example`!D\$38				
24	Invariant	_INVCONTROL-10 + Hg18 chr10:76411622-76411718	_INVCONTROL	372	316	403	389	389
25	Invariant	_INVCONTROL-16 + Hg18 chr16:11026537-11026636	_INVCONTROL	422	497	481	502	514
26	Invariant	_INVCONTROL-17 + Hg18 chr17:60617472-60617571	_INVCONTROL	772	740	712	764	726
27	Invariant	_INVCONTROL-19 + Hg18 chr19:38818850-38818949	_INVCONTROL	639	648	695	655	627
28	Invariant	_INVCONTROL-2 + Hg18 chr2:137221998-137222092	_INVCONTROL	589	578	618	588	660
29	Invariant	_INVCONTROL-20 + Hg18 chr20:10504898-10504992	_INVCONTROL	475	447	497	529	506
30	Invariant	_INVCONTROL-22 + Hg18 chr22:42509372-42509471	_INVCONTROL	872	890	805	850	885
31	Invariant	_INVCONTROL-5 + Hg18 chr5:71261236-71261335	_INVCONTROL	889	964	862	840	854
32	Invariant	_INVCONTROL-7 + Hg18 chr7:41541153-41541252	_INVCONTROL	779	806	815	792	821
33	Endogenous	probe1 + Hg18 chr1: coordinates	CNV region 1	204	330	225	269	303
34	Endogenous	probe2 + Hg18 chr8: coordinates	CNV region 2	456	711	155	225	237
35	Endogenous	probe3 + Hg18 chr10: coordinates	CNV region 3	399	609	284	377	415
36	Endogenous	probe4 + Hg18 chr12: coordinates	CNV region 4	480	779	232	317	320
37	Endogenous	probe5 + Hg18 chrX:111441962-111442061	CNV region 5	444	265	312	207	469
38	Endogenous	probe6 + Hg18 chrY:2715200-2715207	CNV region 6	0	214	229	229	0

For copy number analysis, it is not necessary to normalize the data for POS, NEG or Restriction Site controls. For simplicity, these probes can be left out of the INV normalization spreadsheet. The INV-normalization procedure should be carried out on the Invariant control and Endogenous code classes.



CAUTION: Copy number data generated by the Custom CNV Assay can be negatively affected when DNA input amounts are too low, at which point sampling error can introduce unacceptably high levels of variation in the data. NanoString recommends a minimum of 100 counts for the average of the 10 Invariant control probes in the INV-normalized data set to ensure reliable copy number estimation. This is particularly important for the reference sample(s), since accurate copy number calculations depend upon high quality reference sample data. A poor quality reference sample will adversely affect copy number calls for all samples. A poor quality test sample will result in unreliable copy number call for that sample alone. In general for purified genomic DNA, free of RNA contamination, 100 INV-normalized counts will correspond to ≥ 100 ng genomic DNA.

Calculating Copy Number Estimates

Reference Sample Selection

The basic data analysis strategy for determining copy numbers with the nCounter Custom CNV Assay is to calculate a copy number estimate for each probe relative to a reference sample (or samples). Each probe in the Custom CNV CodeSet is a unique sequence and bar code, and as a result small variations in probe efficiency can result in count variation between probes even when targeting genomic regions of equal copy number. However, this difference in counting efficiency will be constant for a given probe over all samples analyzed. Therefore, highly accurate copy number estimates can be generated by simply taking the ratio of counts from test samples to the counts of a fixed reference sample(s), and calculating copy numbers relative to that reference sample.

The example analysis that follows will assume a copy number of 2 for all genomic loci being assayed in the reference sample. In a later section ([Reference Sample Adjustments on page 10](#)), we will consider alternative analysis strategies for genomic loci that differ from 2 copies in the reference sample(s).

NOTE: If the copy number of your reference sample is not known, the analysis method presented here will only estimate the copy number relative to the unknown reference; an absolute integer copy number prediction will not be possible.

First, start a new sheet for copy number calculations and select your reference sample. In the example below, the reference sample will be Sample 1 (Column D). Next, divide each test samples probe value (INV-normalized counts) by the corresponding probes in the reference samples' INV-normalized counts. For autosomal chromosomes (1-22), multiply this quotient by 2 to account for the presence of two chromosomal copies in the diploid reference sample. (**Hint:** Use the "\$" shortcut command to hold the reference sample column constant in the formula.)

FIGURE 1.9

	A	B	C	D	E	F	G	H	I
21	Reporter Counts			Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6
22	Code Class	Name	Accession						
23	Invariant	_INVCONTROL-1 + Hg18 chr1:97239055-97239154	_INVCONTROL	=INV norm example!\$D23/'INV norm example'!D23*2					2.1
24	Invariant	_INVCONTROL-10 + Hg18 chr10:76411622-76411718	_INVCONTROL	2.0	2.4	1.8	1.9	1.9	1.9
25	Invariant	_INVCONTROL-16 + Hg18 chr16:11026537-11026636	_INVCONTROL	2.0	1.7	1.8	1.7	1.6	1.6
26	Invariant	_INVCONTROL-17 + Hg18 chr17:60617472-60617571	_INVCONTROL	2.0	2.1	2.2	2.0	2.1	2.1
27	Invariant	_INVCONTROL-19 + Hg18 chr19:38818850-38818949	_INVCONTROL	2.0	2.0	1.8	1.9	2.0	2.1
28	Invariant	_INVCONTROL-2 + Hg18 chr2:137221998-137222092	_INVCONTROL	2.0	2.0	1.9	2.0	1.8	2.0
29	Invariant	_INVCONTROL-20 + Hg18 chr20:10504898-10504992	_INVCONTROL	2.0	2.1	1.9	1.8	1.9	2.0
30	Invariant	_INVCONTROL-22 + Hg18 chr22:42509372-42509471	_INVCONTROL	2.0	2.0	2.2	2.1	2.0	2.1
31	Invariant	_INVCONTROL-5 + Hg18 chr5:71261236-71261335	_INVCONTROL	2.0	1.8	2.1	2.1	2.1	2.0
32	Invariant	_INVCONTROL-7 + Hg18 chr7:41541153-41541252	_INVCONTROL	2.0	1.9	1.9	2.0	1.9	2.0

X and Y Chromosome Copy Number Calculation

The formula for determining copy number of the X and Y chromosomes must be adjusted depending on the gender of the reference sample. If the reference sample is female, the formulas should reflect two copies of the X chromosome and 0 copies of the Y chromosome as shown on [Figure 1.10](#).

FIGURE 1.10

	A	B	C	D	E	F	G	H	I
21	Reporter Counts			Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6
22	Code Class	Name	Accession						
23	Invariant	_INVCONTROL-1 + Hg18 chr1:97239055-97239154	_INVCONTROL	2.0	2.2	2.2	2.2	2.4	2.1
24	Invariant	_INVCONTROL-10 + Hg18 chr10:76411622-76411718	_INVCONTROL	2.0	2.4	1.8	1.9	1.9	1.9
25	Invariant	_INVCONTROL-16 + Hg18 chr16:11026537-11026636	_INVCONTROL	2.0	1.7	1.8	1.7	1.6	1.6
26	Invariant	_INVCONTROL-17 + Hg18 chr17:60617472-60617571	_INVCONTROL	2.0	2.1	2.2	2.0	2.1	2.1
27	Invariant	_INVCONTROL-19 + Hg18 chr19:38818850-38818949	_INVCONTROL	2.0	2.0	1.8	1.9	2.0	2.1
28	Invariant	_INVCONTROL-2 + Hg18 chr2:137221998-137222092	_INVCONTROL	2.0	2.0	1.9	2.0	1.8	2.0
29	Invariant	_INVCONTROL-20 + Hg18 chr20:10504898-10504992	_INVCONTROL	2.0	2.1	1.9	1.8	1.9	2.0
30	Invariant	_INVCONTROL-22 + Hg18 chr22:42509372-42509471	_INVCONTROL	2.0	2.0	2.2	2.1	2.0	2.1
31	Invariant	_INVCONTROL-5 + Hg18 chr5:71261236-71261335	_INVCONTROL	2.0	1.8	2.1	2.1	2.1	2.0
32	Invariant	_INVCONTROL-7 + Hg18 chr7:41541153-41541252	_INVCONTROL	2.0	1.9	1.9	2.0	1.9	2.0
33	Endogenous	probe1 + Hg18 chr1: coordinates	CNV region 1	0.6	3.2	2.2	2.6	3.0	3.9
34	Endogenous	probe2 + Hg18 chr8: coordinates	CNV region 2	2.0	3.1	0.7	1.0	1.0	1.4
35	Endogenous	probe3 + Hg18 chr10: coordinates	CNV region 3	2.0	3.0	1.4	1.9	2.1	2.7
36	Endogenous	probe4 + Hg18 chr12: coordinates	CNV region 4	2.0	3.2	1.0	1.3	1.3	1.8
37	Endogenous	probe5 + Hg18 chrX:111441962-111442061	CNV region 5	=INV norm example!D37/'INV norm example'!D37*2					2.1
38	Endogenous	probe6 + Hg18 chrY:2715200-2715287	CNV region 6	0.0	1.0	1.0	1.1	0.0	0.0

In order to generate meaningful copy number estimates for the Y chromosome probe, it is best to use a reference sample that contains a Y chromosome (male). The formula for Y chromosome probes can then be adjusted to use this male reference to calculate copy numbers for only the Y chromosome probe(s). In the example below, the reference for the Y chromosome probe has been switched to Sample 2 (Column E):

FIGURE 1.11

A	B	C	D	E					
				Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6
21	Reporter Counts								
22	Code Class	Name	Accession						
23	Invariant	_INVCONTROL-1 + Hg18 chr1:97239055-97239154	_INVCONTROL	2.0	2.2	2.2	2.2	2.4	2.1
24	Invariant	_INVCONTROL-10 + Hg18 chr10:76411622-76411718	_INVCONTROL	2.0	2.4	1.8	1.9	1.9	1.9
25	Invariant	_INVCONTROL-16 + Hg18 chr16:11026537-11026636	_INVCONTROL	2.0	1.7	1.8	1.7	1.6	1.6
26	Invariant	_INVCONTROL-17 + Hg18 chr17:60617472-60617571	_INVCONTROL	2.0	2.1	2.2	2.0	2.1	2.1
27	Invariant	_INVCONTROL-19 + Hg18 chr19:38818850-38818949	_INVCONTROL	2.0	2.0	1.8	1.9	2.0	2.1
28	Invariant	_INVCONTROL-2 + Hg18 chr2:137221998-137222092	_INVCONTROL	2.0	2.0	1.9	2.0	1.8	2.0
29	Invariant	_INVCONTROL-20 + Hg18 chr20:10504898-10504992	_INVCONTROL	2.0	2.1	1.9	1.8	1.9	2.0
30	Invariant	_INVCONTROL-22 + Hg18 chr22:42509372-42509471	_INVCONTROL	2.0	2.0	2.2	2.1	2.0	2.1
31	Invariant	_INVCONTROL-5 + Hg18 chr5:71261236-71261335	_INVCONTROL	2.0	1.8	2.1	2.1	2.1	2.0
32	Invariant	_INVCONTROL-7 + Hg18 chr7:41541153-41541252	_INVCONTROL	2.0	1.9	1.9	2.0	1.9	2.0
33	Endogenous	probe1 + Hg18 chr1: coordinates	CNV region 1	0.6	3.2	2.2	2.6	3.0	3.9
34	Endogenous	probe2 + Hg18 chr8: coordinates	CNV region 2	2.0	3.1	0.7	1.0	1.0	1.4
35	Endogenous	probe3 + Hg18 chr10: coordinates	CNV region 3	2.0	3.0	1.4	1.9	2.1	2.7
36	Endogenous	probe4 + Hg18 chr12: coordinates	CNV region 4	2.0	3.2	1.0	1.3	1.3	1.8
37	Endogenous	probe5 + Hg18 chrX:111441962-111442061	CNV region 5	2.0	1.2	1.4	0.9	2.1	2.1
38	Endogenous	probe6 + Hg18 chrY:2715200-2715287	CNV ref	=INV norm example!E38/'INV norm example'!\$E38				0.0	0.0

Averaging Copy Number Estimates by Genomic Region

If your nCounter Custom CNV CodeSet contains multiple probes for a single genomic locus, it may be desirable to generate an average copy number estimate value based on all probes for that particular locus. This average value can then be used to generate an integer copy number assignment. If you are using a single probe per region, proceed to **Generating Integer Copy Number Calls**, below, for instructions on how to round the estimated copy number.

To average probes for each locus, create a new spreadsheet. For each genomic region (locus), calculate the average of copy number estimate values using the “=AVERAGE” function as described in **Normalization to Invariant Probes on page 5**.

Generating Integer Copy Number Calls

To convert copy number estimates to integer copy number predictions, the copy number estimate values must be rounded. The simplest method is to round each estimated value to the nearest integer, using an “IF” function in Excel. In the following example we demonstrate an integer copy number +/- 0.4, although each investigator should determine the appropriate rounding criteria based on their own specific data analysis requirements.

NOTE: Copy number estimates that are half-integer values (e.g., 1.5, 2.5, 3.5) will require further interpretation by the investigator to determine the integer copy number value. In the following example, such values are returned as the decimal copy number estimate value and are not rounded to the nearest integer.

On a new sheet, enter an “IF” formula in the cell corresponding to the first probe and first sample following the example shown below. The values entered in the formula can be adjusted to alter the rounding criteria. For example, values between 0.8 and 1.2 (rather than 0.6 and 1.4) can be rounded to 1. Once the formula has been entered, click on the lower-right corner of the cell and drag the formula down (to apply to all Invariant and Endogenous probes), then across to apply to all samples.

FIGURE 1.12

A	B	C	D	E	F	G	H	I
21	Reporter Counts		Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6
22	Code Class	Name	Accession					
23	Invariant	_INVCONTROL-1 + Hg18 chr1:97239055-97239154	_INVCONTROL	=IF('CN estimate example'!D23<=0.4,"0",IF(AND('CN estimate example'!D23>=0.6,				
24	Invariant	_INVCONTROL-10 + Hg18 chr10:76411622-76411718	_INVCONTROL	'CN estimate example'!D23<=1.4),"1",IF(AND('CN estimate example'!D23>=1.6,'CN				
25	Invariant	_INVCONTROL-16 + Hg18 chr16:11026537-11026636	_INVCONTROL	estimate example'!D23<=2.4),"2",IF(AND('CN estimate example'!D23>=2.6,'CN				
26	Invariant	_INVCONTROL-17 + Hg18 chr17:60617472-60617571	_INVCONTROL	estimate example'!D23<=3.4),"3",IF('CN estimate example'!D23>=3.6,"4 ",'CN				
27	Invariant	_INVCONTROL-19 + Hg18 chr19:38818850-38818949	_INVCONTROL	estimate example'!D23)))))				
28	Invariant	_INVCONTROL-2 + Hg18 chr2:137221998-137222092	_INVCONTROL	2 [IF(logical_test, [value_if_true], [value_if_false])]	2	2		
29	Invariant	_INVCONTROL-20 + Hg18 chr20:10504898-10504992	_INVCONTROL	2	2	2	2	
30	Invariant	_INVCONTROL-22 + Hg18 chr22:42509372-42509471	_INVCONTROL	2	2	2	2	
31	Invariant	_INVCONTROL-5 + Hg18 chr5:71261236-71261335	_INVCONTROL	2	2	2	2	
32	Invariant	_INVCONTROL-7 + Hg18 chr7:41541153-41541252	_INVCONTROL	2	2	2	2	
33	Endogenous	probe1 + Hg18 chr1: coordinates	CNV region 1	1	3	2	3	4
34	Endogenous	probe2 + Hg18 chr8: coordinates	CNV region 2	2	3	1	1	1.4418876
35	Endogenous	probe3 + Hg18 chr10: coordinates	CNV region 3	2	3	1.4217305	2	3
36	Endogenous	probe4 + Hg18 chr12: coordinates	CNV region 4	2	3	1	1	2
37	Endogenous	probe5 + Hg18 chrX:111441962-111442061	CNV region 5	2	1	1.4052658	1	2
38	Endogenous	probe6 + Hg18 chrY:2715200-2715287	CNV region 6	0	1	1	0	0

Here we have presented a simple three-step method for generating integer copy number calls from the nCounter CNV Assay raw data: normalization to INV probe counts, copy number estimation relative to a reference sample, and conversion to integer copy number calls.

In the next section (**Reference Sample Adjustments**) we will consider situations when the copy number of the reference sample differs from 2, and when it is desirable to use multiple reference samples.

Reference Sample Adjustments

In some cases, the reference sample selected may contain genomic regions with a copy number of 0 (deletion), 1 (single-copy), or greater than 2 copies. Since it is not possible to calculate copy number estimates using a reference sample with 0 copies at a particular locus, it is necessary to change the reference to a sample that has counts registered for that specific region, and that has a known copy number.

Adjusting the Copy Number Estimate Calculation

After the appropriate reference sample has been identified, it may be necessary to adjust the formula for calculating the copy number estimate. If the reference sample copy number is 1 for a particular genomic region, the formula for the corresponding probes can be altered by deleting the multiplication factor of "2". In the example below, the reference sample for probe1 (Chr1) is changed to Sample2 (Column E) and the copy number formula is adjusted for a copy number of "1":

NOTE: It may not be necessary to change the reference sample, only to alter the known copy number of the original reference sample following the same method outlined here.

FIGURE 1.13

A	B	C	D						E	F	G	H	I
			Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6					
21	Reporter Counts												
22	Code Class	Name	Accession										
23	Invariant	_INVCONTROL-1 + Hg18 chr1:97239055-97239154	_INVCONTROL	2.0	2.2	2.2	2.2	2.4	2.1				
24	Invariant	_INVCONTROL-10 + Hg18 chr10:76411622-76411718	_INVCONTROL	2.0	2.4	1.8	1.9	1.9	1.9	1.9	1.9		
25	Invariant	_INVCONTROL-16 + Hg18 chr16:11026537-11026636	_INVCONTROL	2.0	1.7	1.8	1.7	1.6	1.6	1.6	1.6		
26	Invariant	_INVCONTROL-17 + Hg18 chr17:60617472-60617571	_INVCONTROL	2.0	2.1	2.2	2.0	2.1	2.1	2.1	2.1		
27	Invariant	_INVCONTROL-19 + Hg18 chr19:38818850-38818949	_INVCONTROL	2.0	2.0	1.8	1.9	2.0	2.0	2.1	2.1		
28	Invariant	_INVCONTROL-21 + Hg18 chr21:137221998-137222092	_INVCONTROL	2.0	2.0	1.9	2.0	1.8	2.0	2.0	2.0		
29	Invariant	_INVCONTROL-20 + Hg18 chr20:10504898-10504992	_INVCONTROL	2.0	2.1	1.9	1.8	1.9	2.0	2.0	2.0		
30	Invariant	_INVCONTROL-22 + Hg18 chr22:42509372-42509471	_INVCONTROL	2.0	2.0	2.2	2.1	2.1	2.0	2.1	2.1		
31	Invariant	_INVCONTROL-5 + Hg18 chr5:71261236-71261335	_INVCONTROL	2.0	1.8	2.1	2.1	2.1	2.1	2.0	2.0		
32	Invariant	_INVCONTROL-7 + Hg18 chr7:41541153-41541252	_INVCONTROL	2.0	1.9	1.9	2.0	1.9	2.0	2.0	2.0		
33	Endogenous	probe1 + Hg18 chr1: coordinates	CNV region 1	=INV norm example!D33/INV norm example!\$E33									3.9
34	Endogenous	probe2 + Hg18 chr8: coordinates	CNV region 2	2.0	3.1	0.7	1.0	1.0	1.0	1.0	1.4		
35	Endogenous	probe3 + Hg18 chr10: coordinates	CNV region 3	2.0	3.0	1.4	1.9	2.1	2.1	2.7			
36	Endogenous	probe4 + Hg18 chr12: coordinates	CNV region 4	2.0	3.2	1.0	1.3	1.3	1.3	1.8			
37	Endogenous	probe5 + Hg18 chrX:111441962-111442061	CNV region 5	2.0	1.2	1.4	0.9	2.1	2.1	2.1			
38	Endogenous	probe6 + Hg18 chrY:2715200-2715287	CNV region 6	0.0	1.0	1.0	1.1	0.0	0.0	0.0			

If the copy number of the new reference is 3, the formula can be adjusted by changing the multiplication factor to "3" to generate the correct copy number. In the following example, the reference sample is Sample 1, but the formula is adjusted by adding a multiplication factor of 3 to reflect 3 copies of probe region1 in the reference sample:

FIGURE 1.14

A	B	C	D	E	F	G	H
21	Reporter Counts		Sample 1	Sample 2	Sample 3	Sample 4	Sample 5
22	Code Class	Name	Accession				
23	Invariant	_INVCONTROL-1 + Hg18 chr1:97239055-97239154	_INVCONTROL	2.0	2.2	2.2	2.2
24	Invariant	_INVCONTROL-10 + Hg18 chr10:76411622-76411718	_INVCONTROL	2.0	2.4	1.8	1.9
25	Invariant	_INVCONTROL-16 + Hg18 chr16:11026537-11026636	_INVCONTROL	2.0	1.7	1.8	1.7
26	Invariant	_INVCONTROL-17 + Hg18 chr17:60617472-60617571	_INVCONTROL	2.0	2.1	2.2	2.0
27	Invariant	_INVCONTROL-19 + Hg18 chr19:38818850-38818949	_INVCONTROL	2.0	2.0	1.8	1.9
28	Invariant	_INVCONTROL-2 + Hg18 chr2:137221998-137222092	_INVCONTROL	2.0	2.0	1.9	2.0
29	Invariant	_INVCONTROL-20 + Hg18 chr20:10504898-10504992	_INVCONTROL	2.0	2.1	1.9	1.8
30	Invariant	_INVCONTROL-22 + Hg18 chr22:42509372-42509471	_INVCONTROL	2.0	2.0	2.2	2.1
31	Invariant	_INVCONTROL-5 + Hg18 chr5:71261236-71261335	_INVCONTROL	2.0	1.8	2.1	2.1
32	Invariant	_INVCONTROL-7 + Hg18 chr7:41541153-41541252	_INVCONTROL	2.0	1.9	1.9	2.0
33	Endogenous	probe1 + Hg18 chr1: coordinates	CNV region 1	=INV norm example!D33/'INV norm example'!\$D33*3			
34	Endogenous	probe2 + Hg18 chr8: coordinates	CNV region 2	2.0	3.1	0.7	1.0
35	Endogenous	probe3 + Hg18 chr10: coordinates	CNV region 3	2.0	3.0	1.4	1.9
36	Endogenous	probe4 + Hg18 chr12: coordinates	CNV region 4	2.0	3.2	1.0	1.3
37	Endogenous	probe5 + Hg18 chrX:111441962-111442061	CNV region 5	2.0	1.2	1.4	0.9
38	Endogenous	probe6 + Hg18 chrY:2715200-2715287	CNV region 6	0.0	1.0	1.0	1.1

The adjusted formula can then be copied to the appropriate cells for each probe by clicking on the lower right corner of the highlighted cell and dragging over the cells you would like to change.

Reference Documents and Support

For additional information on the nCounter CNV Assay and data analysis, please refer to these documents available for download at www.nanostring.com:

- nCounter® CNV Collector Tool User Manual
- nCounter® Custom CNV Assay Manual

Contacting Support

For questions about the nCounter® CNV Assay, nCounter® CNV Collector Tool or data analysis, please contact:

support@nanostring.com

Phone: 206-378-NANO

or contact your NanoString Field Applications Scientist.



NanoString Technologies, Inc.

530 Fairview Ave N
Suite 2000
Seattle, Washington 98109

CONTACT US

info@nanosting.com
Tel: (888) 358-6266
Fax: (206) 378-6288
www.nanosting.com

SALES CONTACTS

United States: us.sales@nanosting.com
Europe: europe.sales@nanosting.com
Japan: japan.sales@nanosting.com
Other Regions: info@nanosting.com

© 2011 NanoString Technologies, Inc. All rights reserved. NanoString®, NanoString Technologies®, nCounter®, and Molecules That Count® are registered trademarks of NanoString Technologies, Inc., ("NanoString") in the United States and/or other countries. All other trademarks and/or service marks not owned by NanoString that appear in this document are the property of their respective owners. The manufacture, use and/or sale of NanoString product(s) may be subject to one or more patents or pending patent applications owned by NanoString or licensed to NanoString from Life Technologies Corporation and other third parties.

FOR RESEARCH USE ONLY. Not for use in diagnostic procedures.

MAN-C0014-02