Introduction to Copy Number Variation

Paul Streng
Staff Field Applications Scientist
Importance of Copy Number

- Redon et. al. defined copy number variation (CNV) as a deletion or duplication event involving >1 kb of DNA
  - An important polymorphism
    - ~20,000 identified CNVs
    - Corresponding to >6,000 unique regions/locus in human genome assembly
  - Associated with diseases or genomic disorders such as cancer, immune diseases, and neurological disorders, etc.

- Gene dosage effects can be phenotypic
  - CYP2D6 is associated with drug metabolizing phenotype
  - CCL3L1 affects the susceptibility to HIV/AIDS

Copy Number Variation: Timeline

Timeline | Landmarks in the study of human genetic variation

(1960–1980) Analysis of protein sequences from several individuals revealed an extensive and largely unexpected level of variation.


(1980) Description of copy number variation of the ß-globin genes by Kan and co-workers.


(1981) Independent reports described the widespread existence of short sequence repeat (SSR) variants, also called microsatellites or short tandem repeats (STR) and their application as genetic markers.

(1989–1996) Microsatellites became the gold-standard DNA markers for genetic studies and thousands of microsatellite markers were used to create linkage maps of all human chromosomes.

The HapMap consortium genotyped 1 million single SNPs.

2004

2005

2007

Botstein et al. proposed to use RFLPs to generate linkage maps of the human genome.

Wyman and White described a highly variable restriction fragment length polymorphism (RFLP).

Nakamura et al. described the use of variable number tandem repeat (VNTR) markers for human gene mapping.

Genomic rearrangements are identified as the mutational mechanism that leads to Charcot-Marie-Tooth disease type 1A (REF. 53).

Identification that a large subset of SNPs are paralogous sequence variants (PSVs) and define regions of structural variability.

Interrogation of genomic variability by array hybridization methods demonstrated the existence of copy number variants (CNVs).

Redon et al. identified 1,447 copy number variable regions showing that at least 12% of the human genome contains CNVs.
Two main approaches for CNV detection

- **Microscopic structural variation**
  > Microscopic views of chromosomal variations
  >  - Unbanded karyotypes: aneuploidies, marker chromosomes, gross rearrangements, variation in Y-chromosome size
  >  - Solid stained, unbanded chromosomes → heteromorphisms
  >  - Chromosome banding → translocations, deletions, duplications, insertions, inversions

- **Submicroscopic structural variation**
  > Genome-wide analysis
    > Array based approaches
      > Array-CGH (comparative genome hybridization)
      > ROMA (representational oligonucleotide microarray analysis)
  > Targeted, PCR based approaches
    > Real-Time PCR (single target)
Techniques

- Microscopic structural variation techniques
  
  Giemsa banding (G banding)
  Centromere banding (C banding)

Spectral karyotyping → t(7;13)
Techniques

- Microscopic structural variation techniques

  3 color FISH showing a rare inversion in interphase nuclei

  ![Image of 3 color FISH](image1)

  Rare inversion

  Common

  ![Image of 2 color FISH](image2)

  Duplication

  2 color FISH showing a duplication in interphase nuclei
Techniques

• Submicroscopic structural variations
  — Array CGH
    > Labeled fragments from a genome of interest competitively hybridized with a second differentially labeled genome to arrays
    > Spotted with cloned DNA fragments
    > Compare CN differences between the two genomes
Are there any public databases available?
Database of Genomic Variants

A curated catalogue of structural variation in the human genome

About The Project | Genome Browser | Download | Links | Data Submissions | Email us

Please select genome assembly: Build 36 (Mar. 2008)

View Data by Chromosome

View Data by Genome

Keyword Search

Search

Exact Match? □ Yes □ No
Examples: gene name, accession number, cytoband, gene

BLAT Search

Enter sequence in FASTA format here:

BLAT Search

Summary Statistics

Total entries: 31615 (hg18)
CNVs: 19792
Inversions: 487
Indels (100bp-1Kb): 11336
Total CNV loci: 6225
Articles cited: 28
Last updated: Nov 10, 2008
Join our mailing list
DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources

DECIPHER v4.1

Newsletters
December 2008
November 2008
October 2008
September 2008

Important Date for Your 2009 Diary
DECIPHER Symposium 2008-22nd May 2009

Syndrome of the month
12q14 microdeletion

A mechanism of non-homologous end-joining (NHEJ)
may be responsible for the occurrence of the microdeletions in the three published cases...

DECIPHER in the World
Applied Biosystems Technique
Workflow

TaqMan® Copy Number Assay
Reporter = FAM
Quencher = NFQ

TaqMan® Copy Number Reference Assay
Reporter = VIC
Quencher = TAMRA

PCR Master Mix

gDNA
2 – 20 ng/μL

Real-time PCR

Data analysis with CopyCaller™

4 replicates per gDNA sample
TaqMan® Gene Copy Number Assays

Pre-developed CNV assays in a single tube
- Forward Primer
- Reverse Primer
- TaqMan® MGB Probe

➤ 1.6 Million assays to choose from

<table>
<thead>
<tr>
<th>Scale</th>
<th>96-well 20uL</th>
<th>384-well 10uL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small</td>
<td>360</td>
<td>720</td>
</tr>
<tr>
<td>Medium</td>
<td>750</td>
<td>1,500</td>
</tr>
<tr>
<td>Large</td>
<td>2,900</td>
<td>5,800</td>
</tr>
</tbody>
</table>
# TaqMan® Copy Number Assays in Detail

<table>
<thead>
<tr>
<th>Product</th>
<th>Part Number</th>
<th>Number of Reactions in 96-Well Format</th>
<th>Number of Reactions in 384-Well Format</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>TaqMan® Copy Number Assays, Small</td>
<td>4400291</td>
<td>360</td>
<td>720</td>
<td>20X</td>
</tr>
<tr>
<td>TaqMan® Copy Number Assays, Medium</td>
<td>4400292</td>
<td>750</td>
<td>1,500</td>
<td>20X</td>
</tr>
<tr>
<td>TaqMan® Copy Number Assays, Large</td>
<td>4400293</td>
<td>2,900</td>
<td>5,000</td>
<td>60X</td>
</tr>
<tr>
<td>TaqMan® Copy Number Reference Assay RNase P, 750 Reactions</td>
<td>4403226</td>
<td>750</td>
<td>1,500</td>
<td>1 tube, 20X</td>
</tr>
<tr>
<td>TaqMan® Copy Number Reference Assay RNase P, 3800 Reactions</td>
<td>4403228</td>
<td>3000</td>
<td>4,000</td>
<td>4 tubes, 20X</td>
</tr>
<tr>
<td>TaqMan® Copy Number Reference Assay TERT, 750 Reactions</td>
<td>4403216</td>
<td>750</td>
<td>1,500</td>
<td>1 tube, 20X</td>
</tr>
<tr>
<td>TaqMan® Copy Number Reference Assay TERT, 3800 Reactions</td>
<td>4403215</td>
<td>3000</td>
<td>6,000</td>
<td>4 tubes, 20X</td>
</tr>
<tr>
<td>Custom TaqMan® Copy Number Assays, Small</td>
<td>4400294</td>
<td>360</td>
<td>720</td>
<td>20X</td>
</tr>
<tr>
<td>Custom TaqMan® Copy Number Assays, Medium</td>
<td>4400295</td>
<td>750</td>
<td>1,500</td>
<td>20X</td>
</tr>
<tr>
<td>Custom TaqMan® Copy Number Assays, Large</td>
<td>4400296</td>
<td>2,900</td>
<td>5,000</td>
<td>60X</td>
</tr>
</tbody>
</table>
# Searching for Assays

**TaqMan® Copy Number Assays**

*Ordering Information*

- **GeneAssist™ Copy Number Assay Workflow Builder**: Browse the GeneAssist™ Copy Number Assay Workflow Builder. Navigate to the genomic region of interest to select pre-designed assays or submit sequences for custom assay design.

- **Full Assay Search**: Click the Assay Search tab at the top of this page for a more comprehensive search tool to find assays that match your gene of interest and species.

- **Quick Assay Search**: Use the orange Start Here box to the right and then search for assays that match your gene of interest.

### How to Order Assays

*Note*: See user’s manual or package insert for limited label license, and trademark information. For Research Use Only. Not for use in diagnostics procedures.

**Trademark**
GeneAssist™ Copy Number Assay Workflow Builder

**TaqMan® Copy Number Assays (Pre-designed)**

Search for pre-designed TaqMan® Copy Number Assays by specifying a location (chromosome location, cytoband), gene of interest (gene name, gene symbol, etc.), transcript, LGV variation ID or assay ID. Once found, simply add the desired assays to the Shopping List to start the ordering process.

**Custom TaqMan® Copy Number Assays**

**Generate Custom Copy Number Assays from Sequence**

Order Custom TaqMan® Copy Number Assays by entering one or more sequences, selecting the desired target sites, and submitting that list of target sites for assay design. You will be notified when the design job is complete, at which time you can proceed with the ordering process.

**View Submitted Custom Copy Number Assay Design Jobs**

Select this option if you have previously entered a sequence and submitted a list of target sites for custom copy number assay design and wish to see the status/results of the design job.
GeneAssist™ Alignment Map
Pre-designed Selection

Available pre-designed Copy Number Assays

<table>
<thead>
<tr>
<th>Assay ID</th>
<th>Assay Location</th>
<th>Cytoband</th>
<th>Gene Symbol</th>
<th>Size</th>
<th>Quantity</th>
<th>Add</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hs07139372_cn</td>
<td>chr19:50602619</td>
<td>19q13.32a</td>
<td>N/A</td>
<td>4400291 Small Scale</td>
<td>1</td>
<td>Add</td>
</tr>
<tr>
<td>Hs01239420_cn</td>
<td>chr19:50603134</td>
<td>19q13.32a</td>
<td>N/A</td>
<td>4400291 Small Scale</td>
<td>1</td>
<td>Add</td>
</tr>
<tr>
<td>Hs00489592_cn</td>
<td>chr19:50603321</td>
<td>19q13.32a</td>
<td>N/A</td>
<td>4400291 Small Scale</td>
<td>1</td>
<td>Add</td>
</tr>
</tbody>
</table>
Supported PCR Master Mixes

• TaqMan® Genotyping Master Mix
  — Recommended

• TaqMan® Gene Expression Master Mix

• TaqMan® Universal PCR Master Mix, No AmpErase® UNG

• TaqMan® Universal PCR Master Mix
DNA, Sample Types & Real-Time PCR

- **gDNA quantification & normalization**
  - TaqMan® RNase P Detection Reagents (PN 4316831)

- **gDNA input & reaction volume**
  - 2 - 20 ng in 10 µL (384-well format)
  - 4 - 40 ng in 20 µL (96-well format)

- **Sample types**
  - 4 replicates per sample
  - Calibrator sample or positive control
  - NTC

- **Real-Time PCR**
  - Absolute Quantitation
  - 2 Detectors/Targets per well
  - Analysis settings of Automatic Baseline & Threshold = 0.2
  - Export Results (Table) for copy number analysis by CopyCaller™
Supported Real-Time PCR Instruments

- 7900HT Fast Real-Time PCR System
  - Includes all legacy 7900 configurations

- 7500 Fast Real-Time PCR System

- 7500 Real-Time PCR System

- 7300 Real-Time PCR System

- StepOnePlus Real-Time PCR System
  - StepOne Real-Time PCR System is not used because it does not detect TAMRA
Supported Copy Number Analysis Software (free download)
Test assay (FAM, YIPF6)

Reference assay (VIC, RNase P)

Calibrator Sample (♀)
- Ct (FAM) = 27.0
- Ct (VIC) = 27.5

Test Sample (♂)
- Ct (FAM) = 28.5
- Ct (VIC) = 27.0

ΔCt = 1.5

ΔCt = 0.5

ΔΔCt = 1

Copy number

2^{-\Delta\Delta Ct} \times 2 = 1
CopyCaller™
CopyCaller™

- 2 Copies 62 Samples (68%)
- 3 Copies 14 Samples (15%)
- 4 Copies 2 Samples (2%)
- 1 Copy 13 Samples (14%)
## ΔCt and CN

<table>
<thead>
<tr>
<th>Copy Number Predicted</th>
<th>Confidence</th>
<th>Z-Score</th>
<th>Replicate Count</th>
<th>Replicates Analyzed</th>
<th>FAM Ct Mean</th>
<th>VIC Ct Mean</th>
<th>ΔCt Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>&gt; 0.99</td>
<td>0.65</td>
<td>16</td>
<td>16</td>
<td>25.2425</td>
<td>24.6407</td>
<td>0.6018</td>
</tr>
<tr>
<td>2</td>
<td>&gt; 0.99</td>
<td>0.05</td>
<td>4</td>
<td>1</td>
<td>24.0838</td>
<td>23.4153</td>
<td>0.668</td>
</tr>
<tr>
<td>1</td>
<td>&gt; 0.99</td>
<td>0.09</td>
<td>4</td>
<td>4</td>
<td>26.3948</td>
<td>24.6885</td>
<td>1.7063</td>
</tr>
<tr>
<td>2</td>
<td>&gt; 0.99</td>
<td>0.44</td>
<td>4</td>
<td>1</td>
<td>25.1546</td>
<td>24.5559</td>
<td>0.5987</td>
</tr>
<tr>
<td>2</td>
<td>&gt; 0.99</td>
<td>0.15</td>
<td>4</td>
<td>4</td>
<td>24.8105</td>
<td>24.1993</td>
<td>0.6113</td>
</tr>
<tr>
<td>2</td>
<td>&gt; 0.99</td>
<td>0.05</td>
<td>4</td>
<td>4</td>
<td>25.2992</td>
<td>24.6435</td>
<td>0.6558</td>
</tr>
<tr>
<td>2</td>
<td>&gt; 0.99</td>
<td>0.04</td>
<td>4</td>
<td>4</td>
<td>24.6171</td>
<td>23.9499</td>
<td>0.6672</td>
</tr>
<tr>
<td>2</td>
<td>&gt; 0.99</td>
<td>0.05</td>
<td>4</td>
<td>4</td>
<td>25.0925</td>
<td>24.3613</td>
<td>0.7312</td>
</tr>
<tr>
<td>2</td>
<td>&gt; 0.99</td>
<td>0.05</td>
<td>4</td>
<td>4</td>
<td>24.2106</td>
<td>23.5475</td>
<td>0.663</td>
</tr>
<tr>
<td>2</td>
<td>&gt; 0.99</td>
<td>0.07</td>
<td>4</td>
<td>4</td>
<td>25.1367</td>
<td>24.4553</td>
<td>0.6808</td>
</tr>
<tr>
<td>2</td>
<td>&gt; 0.99</td>
<td>0.08</td>
<td>4</td>
<td>4</td>
<td>25.1901</td>
<td>24.5037</td>
<td>0.6847</td>
</tr>
<tr>
<td>2</td>
<td>&gt; 0.99</td>
<td>0.35</td>
<td>4</td>
<td>4</td>
<td>25.3351</td>
<td>24.6043</td>
<td>0.7302</td>
</tr>
<tr>
<td>3</td>
<td>&gt; 0.99</td>
<td>0.35</td>
<td>4</td>
<td>4</td>
<td>24.6598</td>
<td>24.5384</td>
<td>0.1214</td>
</tr>
</tbody>
</table>
Why High CN are not Easily Achieved

Separation in Mean $\Delta C_T$ of consecutive Copy Numbers

Approaching a limit of detection, given a finite variability of measured $\Delta C_T$ values about the mean $\Delta C_T$ value expected for that Copy Number Group.

Mean $\Delta C_T$ for Copy Number N

- CN = 4
- CN = 3
- CN = 2
- CN = 1

$\Delta C_T$
Results Table Guidelines

Confidence

• Samples that have low copy number (1, 2 or 3) commonly have confidence values greater than 95%

• As copy number increases, confidence progressively decreases due to the separation of ΔC_T subdistribution values of different copy numbers

• REVIEW SAMPLES THAT HAVE CONFIDENCE VALUES GREATER THAN 95%

|Z-Score|

• Samples with high confidence values can sometimes deviate significantly from the mean copy number for the copy number subdistribution

• For samples with confidence values greater than 95%, look at the z-score and accept or reject based on the following:

<table>
<thead>
<tr>
<th>Z-Score</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 1.75</td>
<td>Pass</td>
</tr>
<tr>
<td>2.65 &gt; z ≥ 1.75</td>
<td>Pass with caution</td>
</tr>
<tr>
<td>≥ 2.65</td>
<td>Fall</td>
</tr>
</tbody>
</table>

Confidence | Z-Score
---|---
> 0.99 | 0.65
> 0.99 | 0.05
> 0.99 | 0.09
> 0.99 | 0.44
> 0.99 | 0.15
> 0.99 | 0.05
> 0.99 | 0.04
> 0.99 | 0.5
> 0.99 | 0.05
> 0.99 | 0.23
> 0.99 | 0.05
Where can I find more information?

www.appliedbiosystems.com/cnv