# Karyostat™ + Service : An Efficient and High Throughput Alternative to Karyotyping Stem Cells using Chromosomal Microarray

Kevin Velez, Mir Rateb, Heather Showers, Ryan Cunningham, David Bayer, Juliana Falero Perez, Chris Hanna, Rebecca Sachtleben , Kat Smith, Omar Farah, Thomas Forbes, Rob Schleifer, Tayato Thew, and Meera Kumar

ThermoFisher Scientific, Verona Road Madison, WI, USA, 53711.

## Abstract

Genetic abnormalities, such as copy number and single nucleotide variants (SNVs), are common in several disease states. Population screening and cytogenetics studies are performed to identify these structural changes and point modifications in DNA. It is even more important in the context of hiPSC cells to confirm that the clonal population derived from the donors have the same genetic background and has not acquired any novel duplications or deletions in the derived hiPSCs. Chromosomal microarray analysis (CMA), also known as comparative genomic hybridization (CGH), and SNP arrays are clinical tools that are currently used to evaluate these stem cells. The Applied Biosystems™ CytoScan™ HT-CMA Arrays are optimized solutions for the detection of chromosomal abnormalities, such as deletions and duplications, in a high throughput automated format. We offer a high throughput Karyotyping service that provides much better resolution than the traditional G-Banding and provides customers with a fast turn around of data, which includes any chromosomal aberrations. The HT-CMA array provides a genome wide copy number analysis using an easy, interpretable Chromosome Analysis Suite software. In this poster, we describe in detail the Karyostat+™ service for cell characterization that produces reliable, low-noise image data with a high signal to noise ratio that is fast, efficient, high in throughput, and is offered at competitive pricing.

## CytoScan™ HT –CMA Arrays

**CytoScan HT-CMA Array specifications**

| | |
|---|---|
| **Markers for copy number analysis** | |
| Total number of copy number markers | 1,162,042 |
| Number of nonpolymorphic markers | 133,823 |
| Number of SNP markers | 1,028,219 |
| Genome build | hg19, hg38 |
| **Median marker spacing (base pairs)** | |
| Intragenic | 953 |
| Intergenic | 4,114 |
| Overall | 1,337 |
| **Percentage of genes covered (50 markers/100 kb)** | |
| DECIPHER (1,778) | 100% |
| Morbid OMIM (2,220) | 99% |
| **Percentage of genes covered (50 markers/400 kb)** | |
| RefSeq (37,580) | 96% |

**Table 1.** Provides an overview of the CytoScan HT-CMA array attributes. The array is in a 96 format with a 5-micron feature size that has 400 Kb resolution of Whole genome backbone and Refseq genes; 100Kb resolution of Morbid OMIM and decipher genes with 178 sequence and targeted variants.

## WorkFlow of Karyostat+ Service

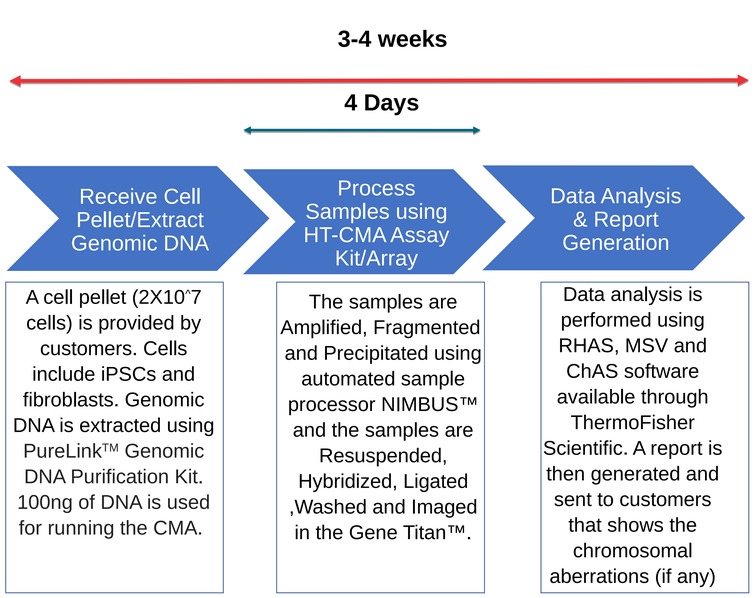



**Figure 1.** Shows the overview of the workflow of Karyostat + services. The boxes underneath shows more detailed breakdown of each step in the workflow.

## Comparison of Various Commercial Microarrays

| Feature | Agilent aCGH (60K) | Illumina microarray | CytoScan 750K | CytoScan Optima | CytoScan HT-CMA | CNV-Seq | Karyotyping |
|---|---|---|---|---|---|---|---|
| Copy number resolution (genome-wide) | Variable, Mb | Variable (100–500 kb) | 100–500 kb | 500 kb–2 Mb | 50–400 kb | 2–5 Mb | >5 Mb |
| LOH detection | No | 5 Mb | 3 Mb | 5 Mb | 3 Mb | No | No |
| Triploidy detection | No | Yes | Yes | Yes | Yes | No | Yes |
| Maternal contamination detection | No | Yes | Yes | Yes | Yes | No | No |
| Input gDNA | 100–200 ng | 100 ng | 50–250 ng | 50–250 ng | 50–100 ng | Low | NA |
| Sequence variant detection capability (e.g., hearing loss) | No | ? | No | No | Yes | No | No |
| Throughput | Low | Flexible | Low | Low | High | High | Low |

Table 2. Comparison of CytoScan HT-CMA Arrays to other microarrays and Karyotyping

## Principle Behind CytoScan HT CMA

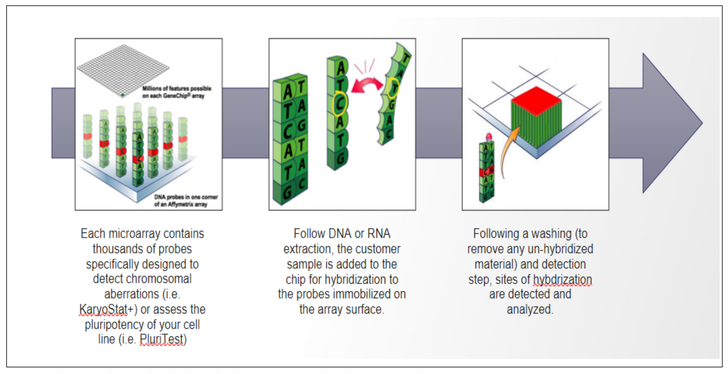


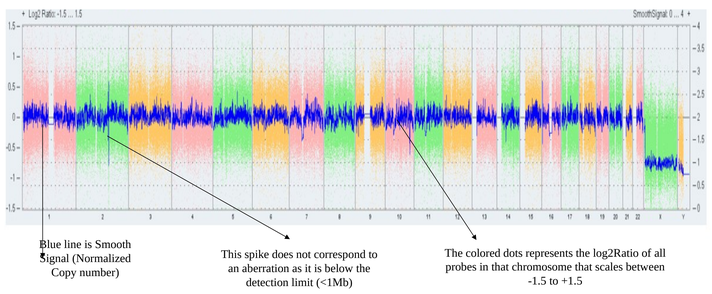

**Figure 2.** Shows the principle of chromosomal microarray in CytoScan HT assay.

## Statistical parameters for Sample QC

| In Process QC | Passing Criteria |
|---|---|
| Genomic DNA quantitation/purity | 260/280:1.8 to 2.0; 260/230:>1.5 |
| DNA conc after resuspension | >1000 ng |
| Fragmentation gel QC | 25-125bp |
| **Data Analysis QC** | |
| MAPD | ≤0.29 |
| SNPQC | ≥ 10.0 |

Table 3. Provides the parameters for a sample to be considered passed. If these stringent conditions are not met, we fail the sample and reprocess the sample one more time as per our terms and conditions. If the sample fails a second time, we notify our customers.

## Whole Genome View of Normal sample




## Whole Genome View of sample showing deletion

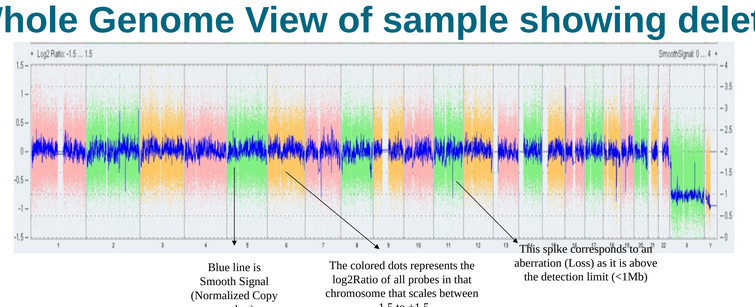



**Figure 3.** Shows the whole genome view showing log2 ratios and copy numbers for normal and aberrated sample.

## Side by side comparison of copy numbers, smooth signal, BAF, and AD in samples with aberrations

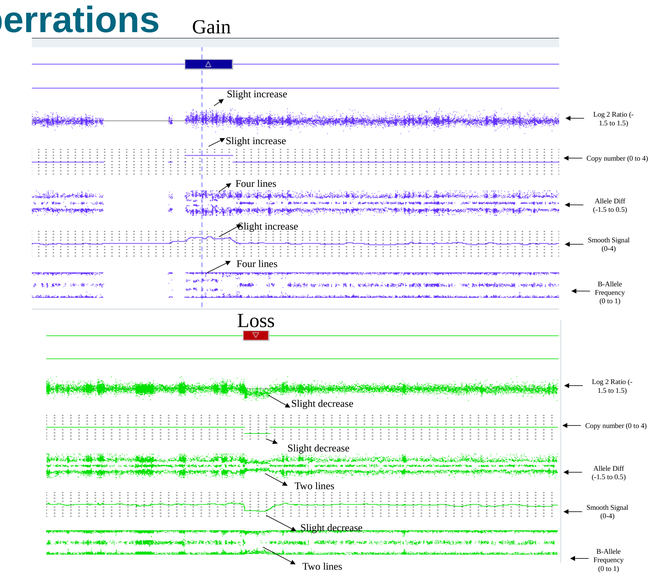



**Figure 4.** To decide if a spike is "True" aberration versus "False", we look at copy number and log2 Ratios in the context of allele difference and B-allele frequency. If we see duplication, we see a trace of four lines in AD and BAF data, and in cases of deletion, we observe a tracer of two lines in BAF and AD data.

## Karyoview of samples

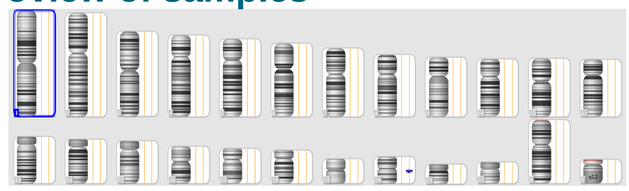


**Figure 5.** A Karyoview of sample is provided with the report only when samples have aberrations. In the above example chromosome 20 has a gain with a copy number =3

## Segment table showing the summary of aberrations

**File:** KS-11789_1_KSR-10051-real.rhchp
**CN State:** 3.00
**Type:** Gain
**Chromosome:** 20
**Materially Modified By:**
**Cytoband Start:** q11.21
**Cytoband End:** q11.22
**Size (kbp):** 4,561
**Marker Count:** 2,315
**Gene Count:** 106
**Genes:** DEFB115, DEFB116, DEFB118, DEFB119, DEFB121, DEFB122, DEFB123, DEFB124, REM1, LINC00028, HM13, HM13-AS1, ID1, MIR3193, COX4I2, BCL2L1, ABALON, TPX2, MYLK2, FOXS1, DUSP15, TTLL9, PDRG1, XKR7, CCM2L, HCK, TM9SF4, TSPY26P, PLAGL2, POFUT1, MIR1825, KIF3B, ASXL1, NOL4L, LOC101929698, NOL4L-DT, C20orf203, COMMD7, DNMT3B, MAPRE1, EFCAB8, SUN5, BPIFB2, BPIFB6, BPIFB3, BPIFB4, BPIFA2, BPIFA4P, BPIFA3, BPIFA1, BPIFB1, CDK5RAP1, SNTA1, CBFA2T2, NECAB3, C20orf144, ACTL10, E2F1, PXMP4, ZNF341, ZNF341-AS1, CHMP4B, RALY-AS1, RALY, MIR4755, EIF2S2, ASIP, AHCY, ITCH, MIR644A, DYNLRB1, MAP1LC3A, PIGU, TP53INP2, NCOA6, HMGB3P1, GGT7, ACSS2, GSS, MYH7B, MIR499A, MIR499B, TRPC4AP, EDEM2, MMP24-AS1-EDEM2, PROCR, MMP24, MMP24OS, EIF6, FAM83C-AS1, FAM83C, UQCC1, GDF5-AS1, GDF5, MIR1289-1, CEP250, C20orf173, ERGIC3, FER1L4, SPAG4, CPNE1, RBM12, NFS1, ROMO1, RBM39, PHF20
**OMIM ® Genes Count:** 71
**OMIM ® Genes:** DEFB118 (607650), DEFB119 (615997), DEFB121 (616075), DEFB122 (616077), DEFB123 (616076), REM1 (610388), HM13 (607106), ID1 (600349), COX4I2 (607976), BCL2L1 (600039), TPX2 (605917), MYLK2 (606566), FOXS1 (602939), DUSP15 (616776), PDRG1 (610789), HCK (142370), TM9SF4 (617727), PLAGL2 (604866), POFUT1 (607491), KIF3B (603754), ASXL1 (612990), NOL4L (618893), COMMD7 (616703), DNMT3B (602900), MAPRE1 (603108), SUN5 (613942), BPIFB2 (614108), BPIFB6 (614110), BPIFB3 (615717), BPIFB4 (615718), BPIFA4P (607627), BPIFA1 (607412), CDK5RAP1 (608200), SNTA1 (601017), CBFA2T2 (603672), NECAB3 (612478), E2F1 (189971), PXMP4 (616397), ZNF341 (618269), CHMP4B (610897), RALY (614663), EIF2S2 (603908), ASIP (600201), AHCY (180960), ITCH (606409), DYNLRB1 (607167), MAP1LC3A (601242), PIGU (608528), TP53INP2 (617549), NCOA6 (605299), GGT7 (612342), ACSS2 (605832), GSS (601002), MYH7B (609928), MIR499A (613614), TRPC4AP (608430), EDEM2 (610302), PROCR (600646), MMP24 (604871), EIF6 (602912), UQCC1 (611797), GDF5 (601146), CEP250 (609689), ERGIC3 (616971), SPAG4 (603038), CPNE1 (604205), RBM12 (607179), NFS1 (603485), ROMO1 (618894), RBM39 (604739), PHF20 (610335)
**Microarray Nomenclature:** arr[GRCh38] 20q11.21q11.22(31,227,615_35,788,564)x3

Table 4. Provides the list of all the genes that are duplicated on chromosome 20. The segment table also provides the size of the duplication along with the location on the chromosome.

## Conclusions

- **The KaryoStat™ assay allows for digital visualization of chromosome aberrations with a resolution similar to g-banding karyotyping.**
- **The size of structural aberration that can be detected is > 2 Mb for chromosomal gains and > 1 Mb for chromosomal losses (the resolution depends on the location of the aberration in the chromosome. Due to a lower probe density on the telomere ends and centromeres, the resolution in those locations may be closer to 5Mb).**
- **The KaryoStat array is optimized for balanced whole-genome coverage with a low-resolution DNA copy number analysis. The assay covers all 36,000 RefSeq genes, including 14,000 OMIM® targets, and enables the detection of aneuploidies, submicroscopic aberrations, and mosaic events.**