

A Technical Guide to Aneuploidy Calling with VeriSeq PGS

Introduction	3
Typical Cases	4
Examples of Calling Challenging Profiles	7
Mosaic Samples	10
Segmental Copy Number Changes	11
Differences Between VeriSeq PGS and 24sure	12
Limitations	14
Technical Assistance	15



This document and its contents are proprietary to Illumina, Inc. and its affiliates ("Illumina"), and are intended solely for the contractual use of its customer in connection with the use of the product(s) described herein and for no other purpose. This document and its contents shall not be used or distributed for any other purpose and/or otherwise communicated, disclosed, or reproduced in any way whatsoever without the prior written consent of Illumina. Illumina does not convey any license under its patent, trademark, copyright, or common-law rights nor similar rights of any third parties by this document.

The instructions in this document must be strictly and explicitly followed by qualified and properly trained personnel in order to ensure the proper and safe use of the product(s) described herein. All of the contents of this document must be fully read and understood prior to using such product(s).

FAILURE TO COMPLETELY READ AND EXPLICITLY FOLLOW ALL OF THE INSTRUCTIONS CONTAINED HEREIN MAY RESULT IN DAMAGE TO THE PRODUCT(S), INJURY TO PERSONS, INCLUDING TO USERS OR OTHERS, AND DAMAGE TO OTHER PROPERTY.

ILLUMINA DOES NOT ASSUME ANY LIABILITY ARISING OUT OF THE IMPROPER USE OF THE PRODUCT(S) DESCRIBED HEREIN (INCLUDING PARTS THEREOF OR SOFTWARE) OR ANY USE OF SUCH PRODUCT(S) OUTSIDE THE SCOPE OF THE EXPRESS WRITTEN LICENSES OR PERMISSIONS GRANTED BY ILLUMINA IN CONNECTION WITH CUSTOMER'S ACQUISITION OF SUCH PRODUCT(S).

FOR RESEARCH USE ONLY

© 2014 Illumina, Inc. All rights reserved.

Illumina, 24sure, BaseSpace, BeadArray, BlueFish, BlueFuse, BlueGnome, cBot, CSPro, CytoChip, DesignStudio, Epicentre, GAllx, Genetic Energy, Genome Analyzer, GenomeStudio, GoldenGate, HiScan, HiSeq, HiSeq X, Infinium, iScan, iSelect, ForenSeq, MiSeq, MiSeqDx, MiSeq FGx, NeoPrep, Nextera, NextBio, NextSeq, Powered by Illumina, SeqMonitor, SureMDA, TruGenome, TruSeq, TruSight, Understand Your Genome, UYG, VeraCode, verifi, VeriSeq, the pumpkin orange color, and the streaming bases design are trademarks of Illumina, Inc. and/or its affiliate(s) in the U.S. and/or other countries. All other names, logos, and other trademarks are the property of their respective owners.

Introduction

Technology

VeriSeq PGS is a new technology for pre-implantation genetic screening (PGS) utilizing next-generation sequencing. VeriSeq PGS enables copy number analysis in single or a few cells from blastomere and trophectoderm biopsies comparable to the proven 24sure technology. In most cases, data interpretation and karyotype-calling are straightforward. However, as with other genomic screening technologies, the quality of DNA and precise experimental conditions can affect the quality of results. It is important to understand these potential effects. While the BlueFuse Multi software provides automated calling functionality, these calls are not definitive, particularly for more complex samples. Laboratories must understand the calling process in detail so that they can consistently obtain the most reliable results.

This document provides an introduction and guidelines for copy number calling for VeriSeq PGS samples with BlueFuse Multi. The focus is on making the best calls from the available data. Refer to *A Technical Guide to Aneuploidy Calling with 24sure Single Channel* (part # 15056973) for a comparison with calling on this platform. Refer to the *BlueFuse Multi Reference Guide* (part # 15053620) for guidance on how to use the software.

Copy number calling and sample assessment with BlueFuse Multi and VeriSeq PGS is clear and unambiguous for good quality samples. The **CNV charts** present results directly as copy numbers plotted against the genomic position of the region. The scale has a large dynamic range allowing for clear identification of whole chromosome gains or losses. The normalized read count for every combined region (bin) of the genome is plotted as a green circle and the copy number assigned to the whole chromosome is drawn as a thick green line.

The CNV charts can be a helpful tool for sample assessment. They can indicate whether an experiment was successful, what the karyotype of the sample is, and how reliable copy number calls might be.

The complete set of copy numbers assigned to all chromosomes of a sample are listed with corresponding confidence values in the **sample report**. Additionally Sample QC measures are listed to assist with passing or failing a sample. The VeriSeq PGS-MiSeq QC Assessment Guide can be used as guidance.

Cycle reports list chromosomal gains and losses, a karyotype summary, and an overall call of "Normal/Abnormal" for all processed samples within a cycle.

Terminology

This field presents some terminology challenges. Euploid is often used to mean a normal chromosome complement of 23 pairs. Technically, euploid indicates an exact multiple of the haploid chromosome complement (23 individual chromosomes or n), and therefore 46 chromosomes (23 pairs or $2n$) is euploid. However, euploid can equally indicate 23 chromosomes (n) or 69 chromosomes ($3n$). In general, when discussing the number of copies of an individual chromosome the terminology uses the "somy" ending: monosomy (one copy), disomy (two copies), trisomy (three copies), tetrasomy (four copies), polysomy (many copies). When referring to copies of a complete set or complement of chromosomes, the ending used is "ploid": diploid ($2n$), triploid ($3n$), tetraploid ($4n$), polyploid (multiples of n greater than $3n$), euploid (multiple of n), aneuploid (any state that is not euploid).

Typical Cases

Assumptions

Copy numbers are assigned based on the numbers of sequencing reads detected for a given region of the genome. BlueFuse Multi internally normalizes the data to remove bias and re-scales the counts so that they correspond to copy numbers. A minimal number of reads per sample must be provided in order for these calculations to work.

BlueFuse Multi works on the assumption that the normal copy number state for autosomes is two and that most chromosomes within a sample have a normal copy number state. If this assumption is not true, profiles can be manually recentered.

Biopsy Types

The objective of copy number calling is to determine which samples have a normal complement of chromosomes for the cell type being analyzed and which have an abnormal copy number. Each biopsy type is different and the differences must be considered when interpreting the data.

A blastomere biopsy typically has a single cell taken from a day three embryo, which has grown to 6–8 blastomeres following fertilization of the egg.

A normal copy number blastomere has two copies of each chromosome. A blastomere with abnormal copy number can have zero (nullisomy), one (monosomy), three (trisomy), four (tetrasomy), or more copies of one or more chromosomes.

A trophoctoderm biopsy is typically taken at day five from a blastocyst stage embryo. A biopsy of the outer cell layer of the blastocyst, called the trophoblast, is taken and typically has 3–5 cells, but varies from biopsy to biopsy and can range from 1–10 cells or more. Interpretation is the same as for a blastomere. However, mosaicism is possible because there are multiple cells.

Confidence Values

Every automatic call is listed in the sample report with a corresponding confidence value ranging from 0.0–1.0. A value above 0.9 can be considered a high confidence copy-number assignment. If bins within a chromosome have low and high values, the confidence will be lower and the call should be carefully examined. A high level of noise across the sample or a segmental copy number change in the chromosome can cause lower confidence values. A value of 0.0 indicates that the analysis failed for this experiment.

Euploid Samples

Figure 1 gives an example of a female euploid sample with little noise and Figure 2 shows a male euploid sample. Note that X and Y copy numbers have higher variability and are not always exactly at 0, 1 or 2, even in normal samples.

Figure 1 Euploid Female Sample

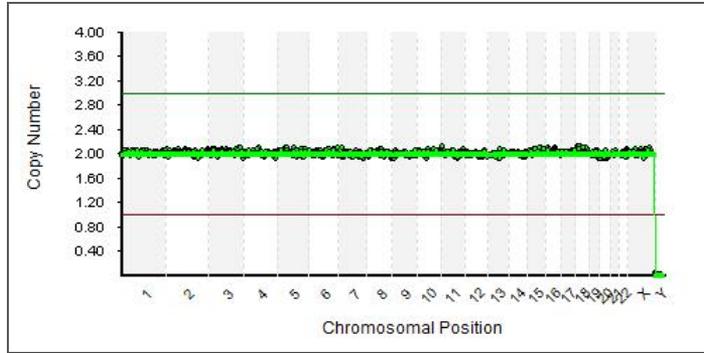
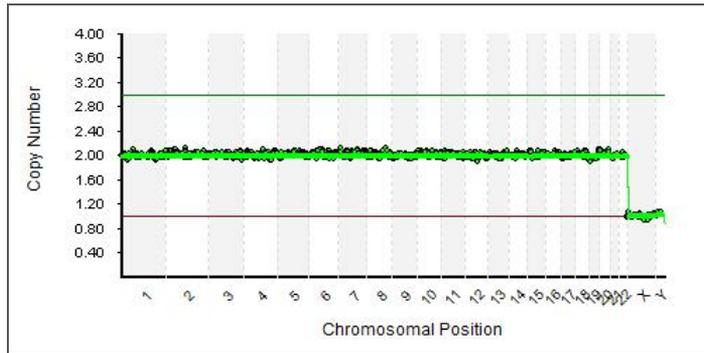


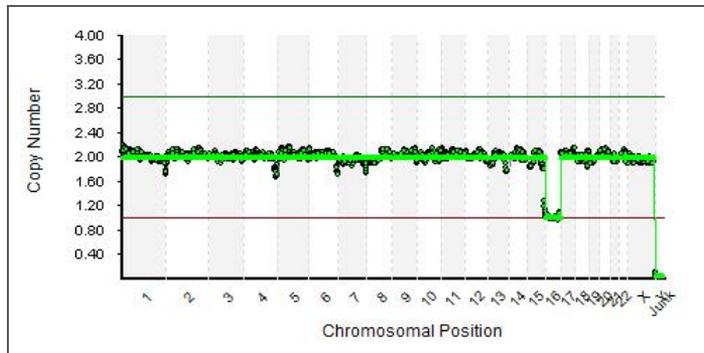
Figure 2 Euploid Male Sample



Aneuploid Samples with Monosomies

Losses of whole chromosomes should be clearly visible with the copy number line drawn around 1 (or 0 for nullisomies) and the individual bin values for this scattered closely around it.

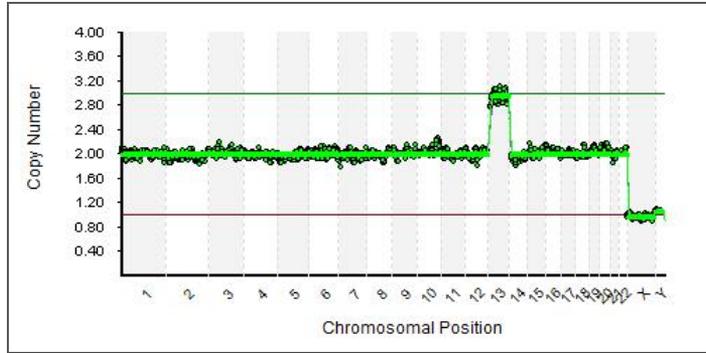
Figure 3 Female Sample with Loss of Chromosome 16



Aneuploid Samples with Trisomies

Single copy gains of whole chromosomes will display the copy number line around 3 and the individual bin values for this chromosome clouded around it.

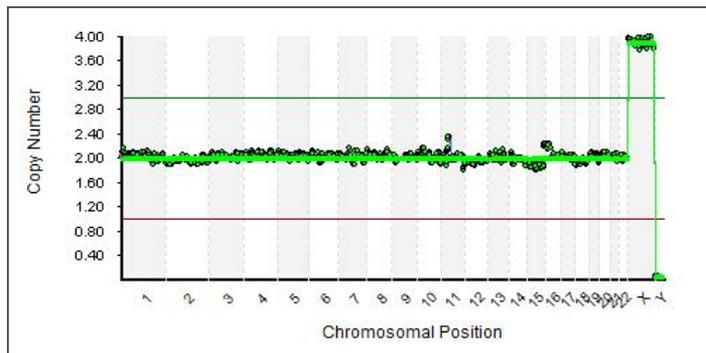
Figure 4 Male Sample with Trisomy 13



Aneuploid Samples with Tetrasomies

Figure 5 shows the chart for a single cell isolated from a cell line with four copies of chromosome X, highlighting the ability of VeriSeq PGS to detect precise copy number levels.

Figure 5 Cell Line Sample with the Green Copy Number Line Indicating the Four Copies of Chromosome X



Examples of Calling Challenging Profiles

Increased Noise

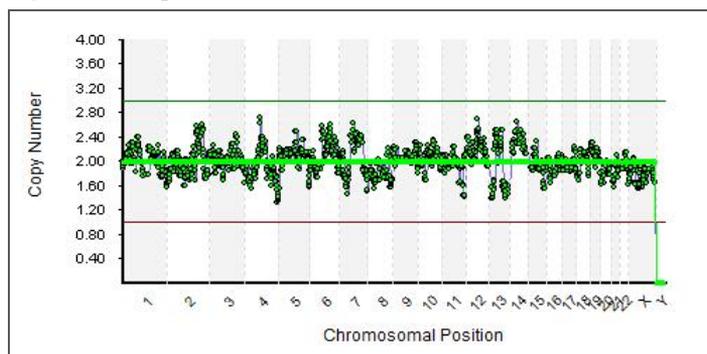
Typically, VeriSeq PGS has an increased dynamic range compared to 24sure, and this applies to sources of noise present in the data. This noise will therefore be more apparent than in 24sure plots. The main causes of noise originate from poor sample quality and amplification artifacts.

Poor sample quality (e.g. degraded DNA) and problems during the DNA amplification step both lead to an increase of the overall noise and spikes in the bin copy numbers within chromosomes, as seen in Figure 6. An overall noise (DLR) score is provided in the sample report, refer to the *VeriSeq PGS-MiSeq QC Assessment Guide (pub # 770-2014-030)* for guidance.

A number of factors can affect the quality of whole genome amplification. These factors include suboptimal embryo biopsy, DNA nicking, incomplete cell lysis, cells undergoing apoptosis, the presence of PCR inhibitors in media, and protocol deviations. These effects are typically more prevalent in single cell biopsies than trophectoderm samples where greater quantities of starting material are available.

A failed amplification can be detected in most cases by observing the lack of a DNA smear in an agarose gel run as part of the SurePlex amplification protocol. However, a smear can occasionally be present even when amplification has failed. Amplification failure can be due to non-human contaminating DNA or occasional self-priming of the SurePlex reaction.

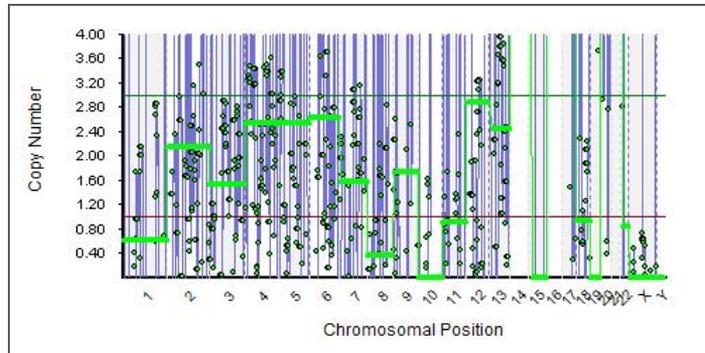
Figure 6 Sample with Increased Overall Noise



Low Read Numbers

Insufficient numbers of sequencing reads will result in noisy profiles and false copy number assignments. An extreme case is shown in Figure 7 where almost no reads are available for the sample. It is critical when analyzing multiple (barcoded) samples in the same run (known as multiplexing) that DNA concentrations are adjusted according to the protocol and the combined library is balanced.

Figure 7 CNV Chart for Sample with Low Numbers of Reads



Switching the chart display to Raw Count will show that the absolute number of reads found for each region is much lower than usual. The Sample QC measures will confirm this observation for the total and filtered number of reads. The *VeriSeq PGS-MiSeq QC Assessment Guide* (pub # 770-2014-030) can be used as guidance on required read numbers.

Protocol Bias

VeriSeq PGS does not require a sample-reference model for data normalization. A specific algorithm step is used to normalize the data and remove any bias from the sample preparation.

If a different sample preparation method than the method specified in the *VeriSeq PGS Library Prep Guide* (part # 15052877) is used, this data correction is likely to fail and produce erroneous copy number calls. The application does not support different amplification methods or protocol deviations. Different amplification methods and protocol deviations will result in noisy data and could lead to false positive and false negative copy number calls.

Sex Chromosome Assessment

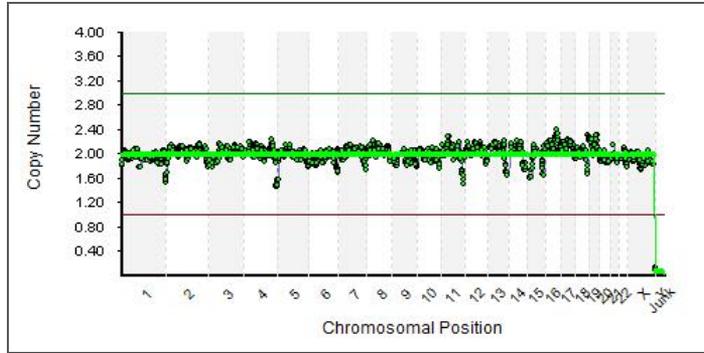
The automatic copy number assessment for sex chromosomes uses the presence of reads for the Y chromosome as an indication for a male sample. Because the unique part of the Y chromosome is comparatively small, the sex chromosome assessment is susceptible to errors if there are not enough reads or the amplification was not optimal.

The exact copy-number level of the sex chromosomes shows slightly more variability than the autosome levels.

Hanging Telomere Artifacts

Telomeric regions have low genomic complexity. They contain a sequence that is repetitive or less unique, and are therefore difficult to analyze. In some suboptimal samples, telomeres show a lower read count and lower copy numbers, as shown in Figure 8.

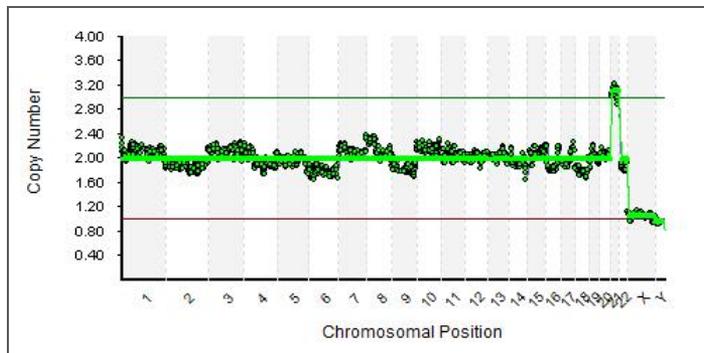
Figure 8 Euploid Sample Plotted with Lower Copy Numbers at the Telomeric Regions



Wave and Step Change Effect Artifacts

In some samples, a slowly varying “wave” is visible. This pattern is an amplification artifact and is specific for the cell or cells isolated. It is most extreme in single cell experiments. An example of an aneuploid sample with wave effect is shown in Figure 9. If multiple cells were analyzed, it is difficult to distinguish between step change artifacts and low-level mosaicism.

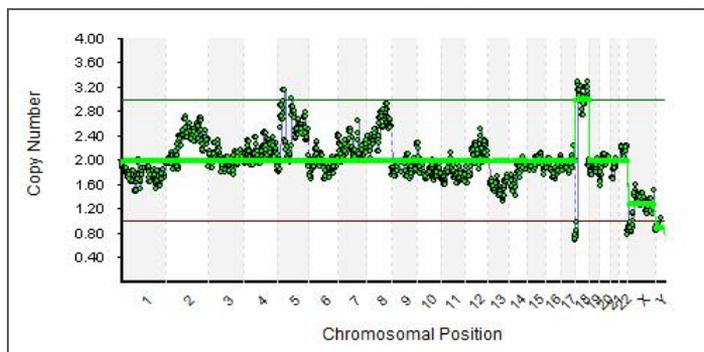
Figure 9 Artifact Showing Alternating Steps for Chromosomes 1–10



Ramping Chromosomes Artifacts

Chromosomes may display continuous copy number increases or decreases (ramps) within a chromosome or across multiple chromosomes, as shown in Figure 10. Amplification failure or copy number variation during the cell cycle can cause these ramps.

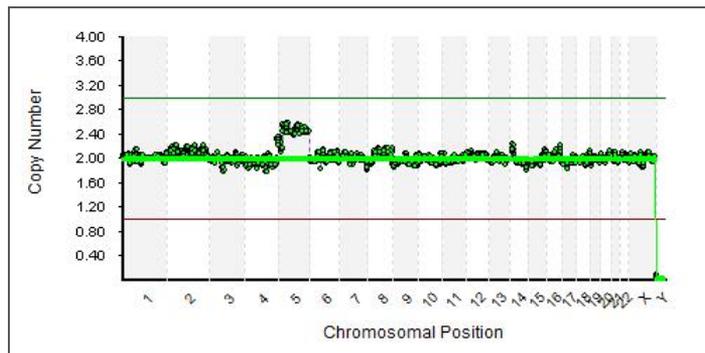
Figure 10 Sample with Increasing Chromosome “Ramp” from Chromosomes 6–8 from a Cell Line Sample with the Expected Karyotype of 46, XY, i(18)(q11.1), i(18)(q11.1)



Mosaic Samples

Trophectoderm samples with multiple cells biopsied may contain cells that have different karyotypes. This mosaicism may be visible in the VeriSeq PGSBlueFuse Multi profile. These changes are currently not automatically called unless they reach more than half the level expected for a full gain or loss. Figure 11 shows a sample with a potential mosaic gain of chromosome 5. The level of mosaicism appears just below 50%.

Figure 11 Female Sample with a Potential Mosaic Gain of Chromosome 5

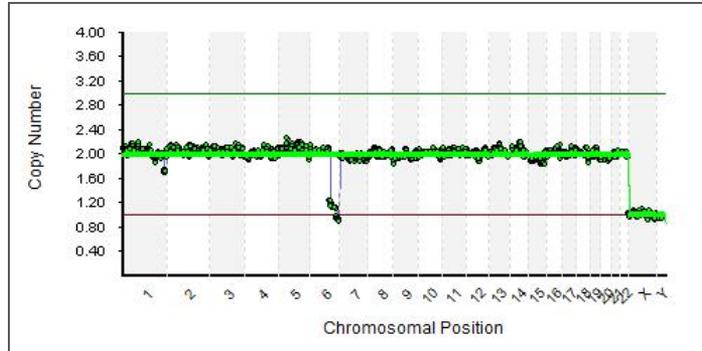


Segmental Copy Number Changes

The VeriSeq PGS algorithms have been developed for whole-chromosome calling. In high quality samples, it is sometimes possible to see subchromosomal imbalances. Figure 12 shows a sample with segmental loss of 6q that can clearly be differentiated from the rest of the genome with its low level of noise.

If the change affects more than 50% of the size of the chromosome, BlueFuse Multi will automatically call a gain/loss for the whole chromosomes. Smaller changes can be called manually.

Figure 12 Sample with a Segmental Loss of 6q



Differences Between VeriSeq PGS and 24sure

The interface of BlueFuse Multi should be immediately recognizable and accessible for existing users of 24sure. However, there are a number of differences in the display of data presented when using VeriSeq PGS that the user should be aware of.

Copy Number Scale

VeriSeq PGS does not require a reference sample to be prepared and used for normalization during the analysis.

VeriSeq PGS operates on a linear copy number scale instead of the logarithmic ratio scale of microarray experiments. Therefore, the concept of XY separation does not apply as it does for 24sure (for male samples in particular). With the normal autosomes expected at copy number value 2, losses and gains are expected at around copy number values 1 and 3 respectively and can be directly read from the CNV chart.

The dynamic range of VeriSeq PGS is higher than 24sure. Therefore, the copy number changes are clearly distinguishable from the normal background in good samples. It also means that any noise will be more apparent and failed samples will be easy to identify. Mosaicism is more clearly visible in VeriSeq PGS.

Karyotype 47,XXY

VeriSeq PGS charts display chromosome counts directly on a copy number scale and do not require a reference sample. This changes the appearance of the sex chromosome counts.

If the sample report is not referred to, an aneuploid karyotype of 47,XXY may be overlooked on the CNV chart and interpreted as 46,XY. Figure 13 shows a karyotype that can be interpreted as 47,XXY with 24sure, even though the gain is not automatically called. Figure 14 shows the same sample processed with VeriSeq PGS. All copy numbers are correctly displayed, but may be overlooked as the copy number line continues from the autosomes.

It is therefore essential for users to familiarize themselves with the copy number scale chart and to remember that chromosomes are labeled and highlighted with alternate shading. The reports generated for every sample or cycle should always be read. All assigned chromosome copy number are listed in the reports.

Figure 13 Potential 47,XXY as Shown with 24sure Against Female Reference

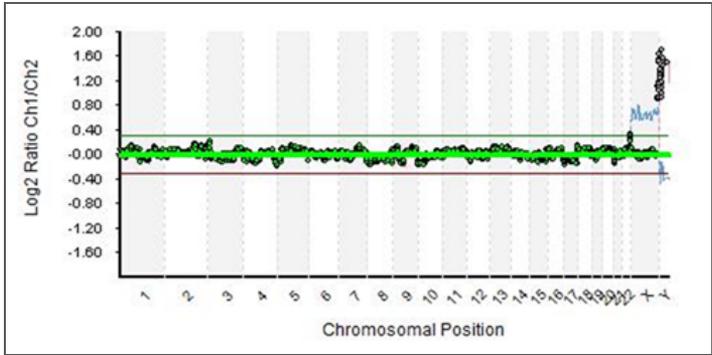
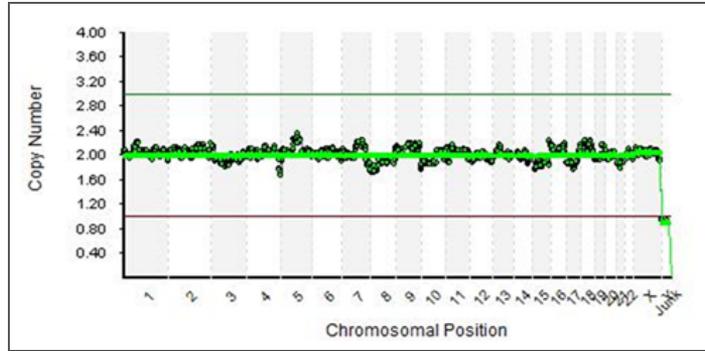


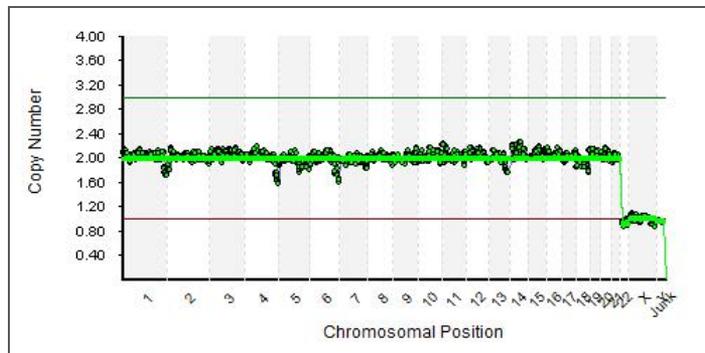
Figure 14 Potential 47,XXY as Shown with VeriSeq PGS



Karyotype 45,XY,-22

Similar caution as described in *Karyotype 47,XXY* should be applied to a loss of chromosome 22 in male samples (Figure 15). The loss can be overlooked during manual calling because it is a small region adjacent to the large chromosome X at the same level. The same precautions should be taken as described in *Karyotype 47,XXY* for correctly identifying 47,XXY.

Figure 15 Potential 45,XY,-22 as Shown with VeriSeq PGS



Limitations

General Limits

VeriSeq PGS was developed for whole-chromosome calling and will not call segmental changes automatically.
Calling of low-level mosaicism in samples is not recommended.

Detecting Karyotype 69,XXY

Like 24sure, VeriSeq PGS is not intended to detect polyploidies. However, the specific case of a potential karyotype 69,XXY could be visually detected with 24sure, as seen in Figure 16. It is more difficult to detect a potential karyotype 69,XXY with VeriSeq PGS, as seen in Figure 17 with a different sample.

In VeriSeq PGS, a difference in the chromosome X and Y levels should be visible. True copy number of 3, 2 and 1 for autosomes, X and Y would in theory appear as 2, 1.33 and 0.66 respectively in the VeriSeq PGS graph.

Figure 16 Potential 69,XXY as Shown with 24sure

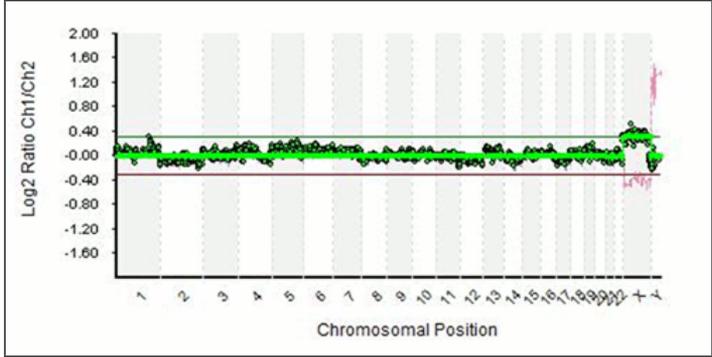
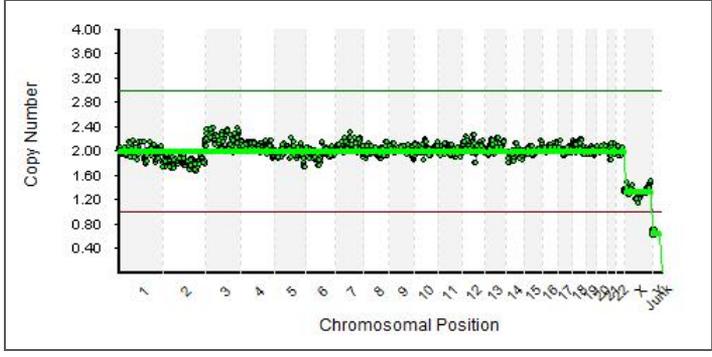


Figure 17 Potential 69,XXY from a Different Sample as Shown with VeriSeq PGS



Technical Assistance

For technical assistance, contact Illumina Technical Support.

Table 1 Illumina General Contact Information

Address	5200 Illumina Way San Diego, CA 92122 USA
Website	www.illumina.com
Email	techsupport@illumina.com

Table 2 Illumina Customer Support Telephone Numbers

Region	Contact Number	Region	Contact Number
North America	1.800.809.4566	Italy	800.874909
Austria	0800.296575	Netherlands	0800.0223859
Belgium	0800.81102	Norway	800.16836
Denmark	80882346	Spain	900.812168
Finland	0800.918363	Sweden	020790181
France	0800.911850	Switzerland	0800.563118
Germany	0800.180.8994	United Kingdom	0800.917.0041
Ireland	1.800.812949	Other countries	+44.1799.534000

Safety Data Sheets

Safety data sheets (SDSs) are available on the Illumina website at support.illumina.com/sds.html.

Product Documentation

Product documentation in PDF is available for download from the Illumina website. Go to support.illumina.com, select a product, then click **Documentation & Literature**.



Illumina

San Diego, California 92122 U.S.A.

+1.800.809.ILMN (4566)

+1.858.202.4566 (outside North America)

techsupport@illumina.com

www.illumina.com