A Sysmex Group Company

Interpret NGS Analysis Software – Seamless transition from microarray to NGS in constitutional cytogenetics

Introduction

At OGT we understand that making the change from microarrays to NGS is a daunting prospect, particularly when it comes to data analysis. That's why we've developed the Interpret NGS analysis software for use in conjunction with the CytoSure™ Constitutional NGS panel for constitutional cytogenetics research. To facilitate the easy transition for copy number variant (CNV) and loss of heterozygosity (LOH) analysis as well as streamlining single nucleotide variant (SNV) and insertion/deletion (indel) calling, we've provided familiar visualisation and interpretation tools alongside an intuitive web interface.

This technical note outlines some of the key features we've incorporated to help make the change from arrays to NGS as painless as possible.

CNV and LOH Data Visualisation

Interpret replicates the copy number and LOH data visualisation approach used in OGT's microarray analysis platform, CytoSure Interpret, including the $\log_2$ ratio/B-Allele Frequency scatter plots, segmentation lines and highlighted CNV and LOH calls to ensure users immediately understand the results (Figure 1).

Figure 1: **A** A 10p15.3-p13 deletion and associated LOH viewed in Interpret NGS Analysis Software, **B** the same deletion viewed in CytoSure Interpret

Annotation Tracks

Comprehensive annotation of genomic features (for example genes, exons, segmental duplications) from local and external CNV databases (such as DGV, ClinGen, DECIPHER), is an essential tool of CNV and LOH interpretation. Interpret displays a number of annotation tracks and provides a variety of mechanisms for importing new tracks, including from an existing CytoSure Interpret database (Figure 2).

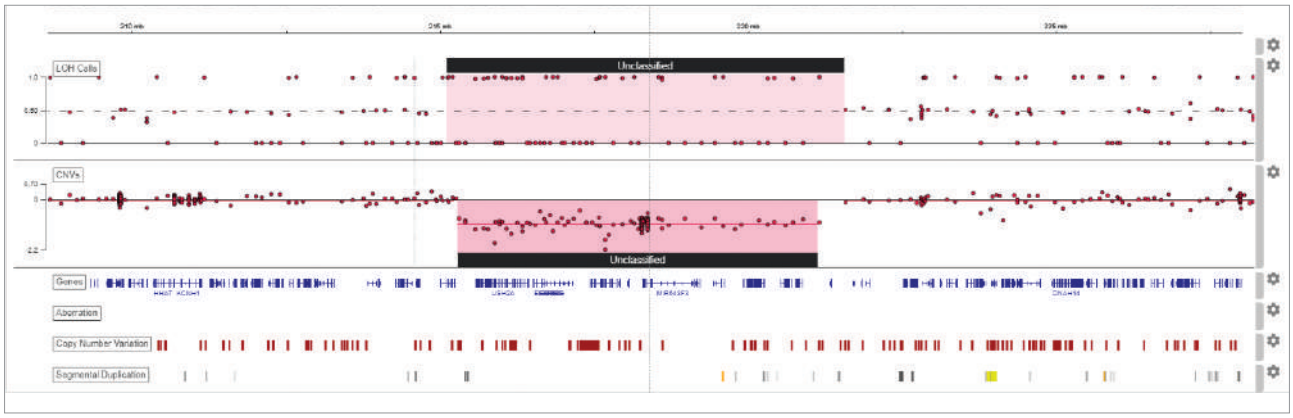


Figure 2: Annotation tracks providing context to a deletion in Interpret

Links to External Resources

Interpret also provides customisable links to external web resources, such as Ensembl, UCSC and DGV, to enable easy access to information relevant to the results and further streamline interpretation (Figure 3).

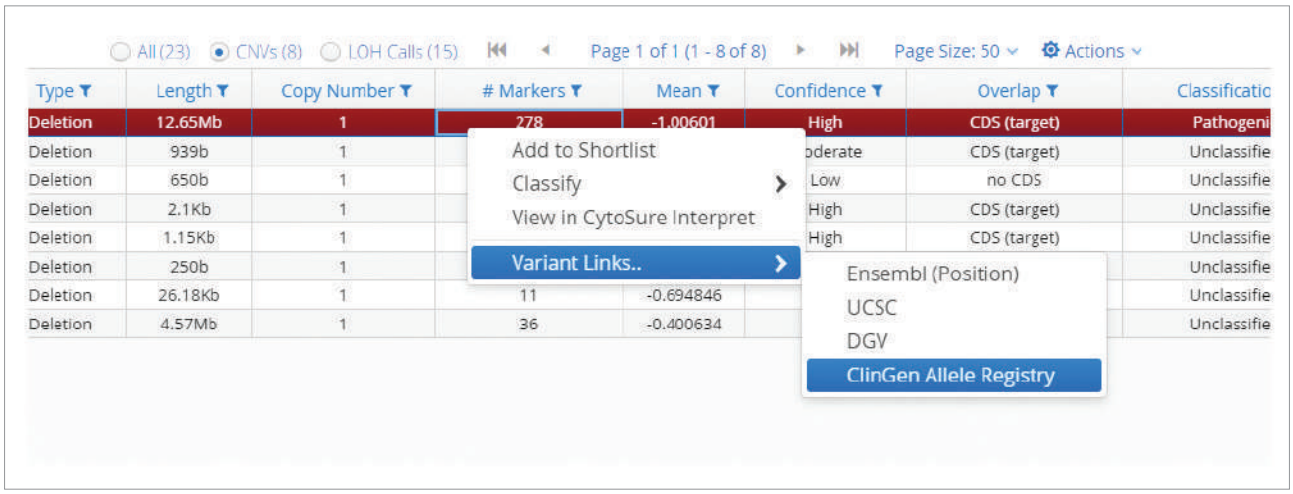


Figure 3: Links to external web resources specific to the selected variant

Automated Workflow

All NGS analysis pipeline and variant detection settings are encapsulated in a customisable analysis protocol, ensuring minimal hands-on time and consistent results generation – simply upload your FASTQ files and select a protocol to produce CNVs, LOH calls and SNVs/indels ready for visualisation and interpretation (Figure 4). In addition, there is no need to process each sample separately, upload a complete run or multiple runs as a single batch for seamless processing (Figure 5).

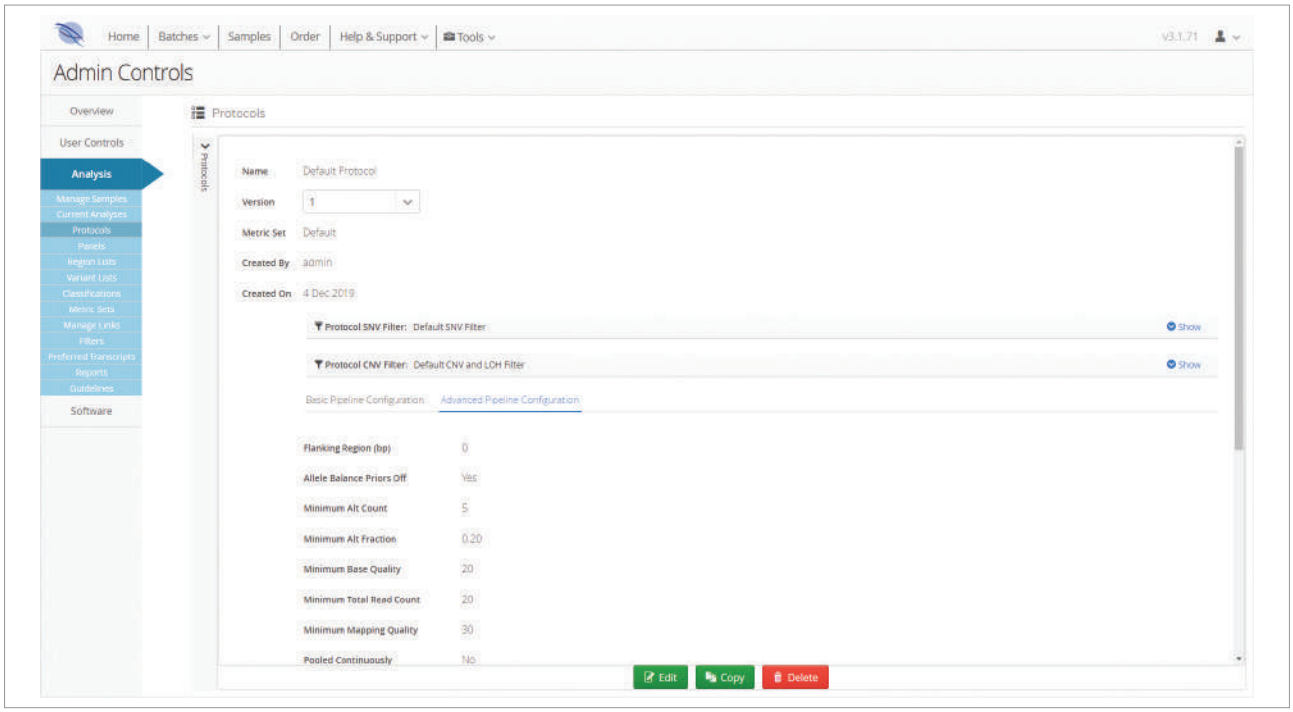


Figure 4: Analysis protocol settings

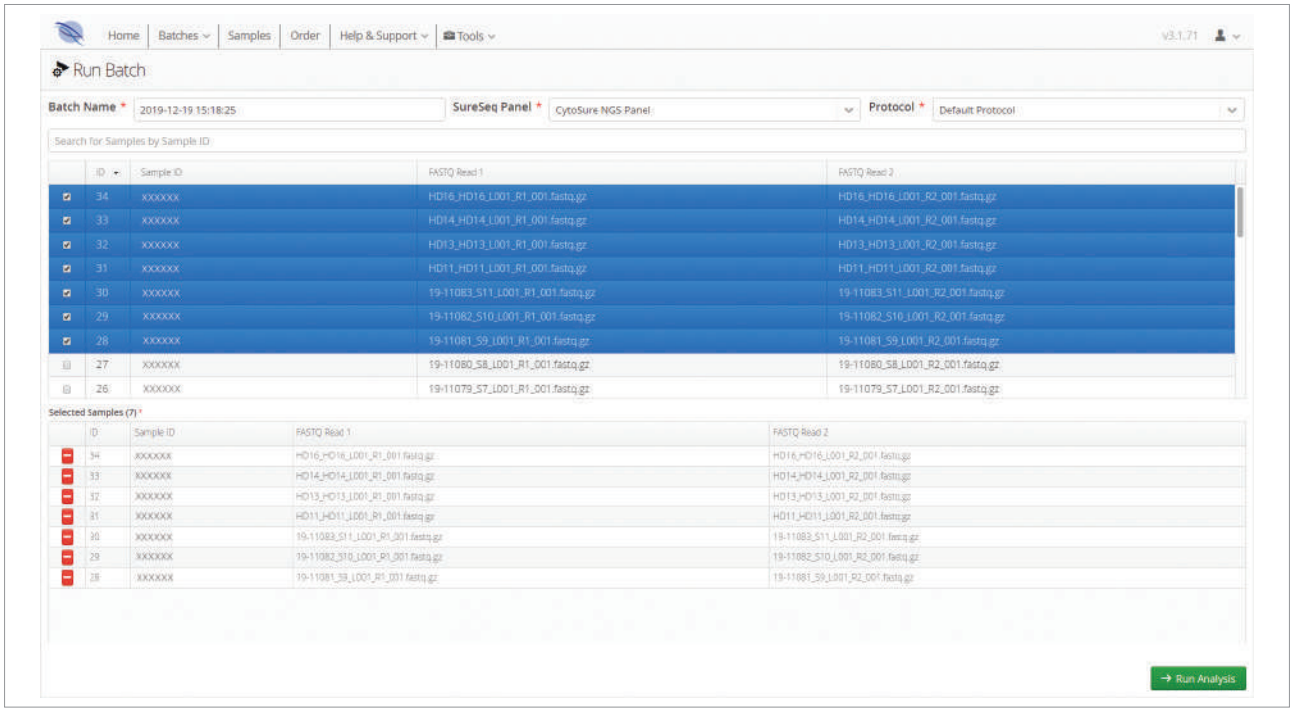


Figure 5: Creation of a batch by selecting samples and an analysis protocol

Easy Results Filtering

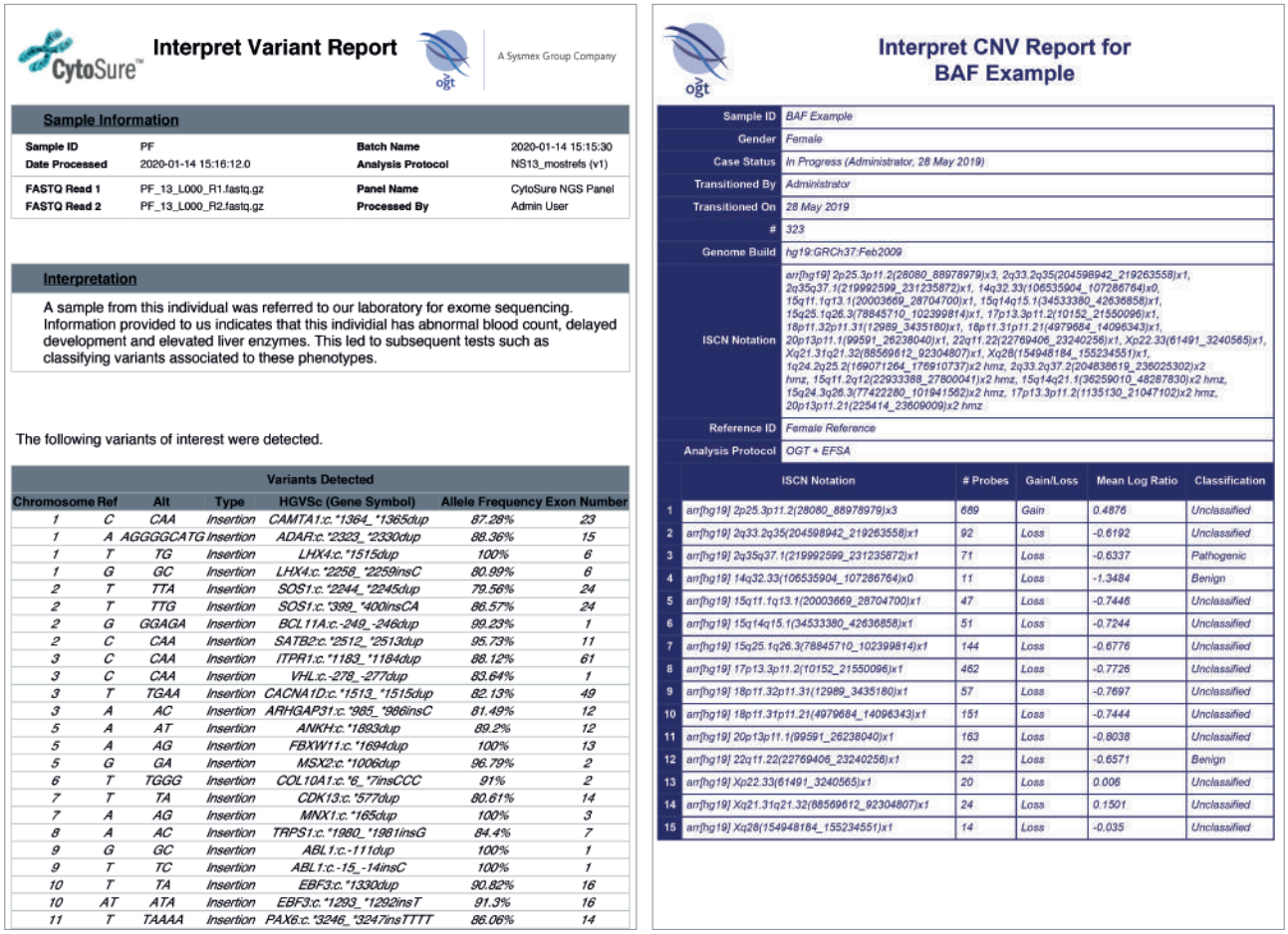
Given the relative size of the CytoSure Constitutional NGS panel, it is inevitable that a large number of calls will be generated by the data analysis pipeline. Simplify your analysis by using predefined protocol filters (Figure 6) or by creating your own. This is especially useful for SNV and indel call filtering and can be achieved through detailed annotation and a highly configurable and user-friendly filtering interface. Easily reduce the number of variants requiring interpretation based on information from external databases such as ExAC, dbSNP and ClinVar, in silico consequence predictions from SIFT and PolyPhen-2 and your own regions of interest (Figure 7).

Figure 6: CNV and LOH results generated automatically by the analysis pipeline, including filtering according to predefined criteria in the analysis protocol (60 CNVs/LOH regions reduced to 15 CNVs and 15 LOH regions).

Figure 7: Filtering of SNVs based on in silico prediction tools to identify causative variants

Flexible Reporting

Interpret enables users to fully customise the layout and format of batch, sample and variant reports through its template-based reporting and plug-in system. Given an example report, OGT can generate a template or plug-in to enable seamless integration with your existing reporting framework (Figure 8).



CytoSure™ **Interpret Variant Report**

A Sysmex Group Company

Sample Information

Sample ID	PF	**Batch Name**	2020-01-14 15:15:30
Date Processed	2020-01-14 15:16:12.0	**Analysis Protocol**	NS13_mostrefs (v1)
FASTQ Read 1	PF_13_L000_R1.fastq.gz	**Panel Name**	CytoSure NGS Panel
FASTQ Read 2	PF_13_L000_R2.fastq.gz	**Processed By**	Admin User

Interpretation

A sample from this individual was referred to our laboratory for exome sequencing. Information provided to us indicates that this individial has abnormal blood count, delayed development and elevated liver enzymes. This led to subsequent tests such as classifying variants associated to these phenotypes.

The following variants of interest were detected.

Variants Detected

Chromosome	Ref	Alt	Type	HGVSc (Gene Symbol)	Allele Frequency	Exon Number
1	C	CAA	Insertion	CAMTA1:c.*1364_*1365dup	87.28%	23
1	A	AGGGGGCATG	Insertion	ADAR:c.*2323_*2330dup	88.36%	15
1	T	TG	Insertion	LHX4:c.*1515dup	100%	6
1	G	GC	Insertion	LHX4:c.*2258_*2259insC	80.99%	6
2	T	TTA	Insertion	SOS1:c.*2244_*2245dup	79.56%	24
2	T	TTG	Insertion	SOS1:c.*398_*400insCA	86.57%	24
2	G	GGAGA	Insertion	BCL11A:c.-249_-246dup	99.23%	1
2	C	CAA	Insertion	SATB2:c.*2512_*2513dup	95.73%	11
3	C	CAA	Insertion	ITPR1:c.*1183_*1184dup	88.12%	61
3	C	CAA	Insertion	VHL:c.-278_-277dup	83.64%	1
3	T	TGAA	Insertion	CACNA1D:c.*1513_*1515dup	82.13%	49
3	A	AC	Insertion	ARHGAP31:c.*985_*986insC	81.49%	12
5	A	AT	Insertion	ANKH:c.*1893dup	89.2%	12
5	A	AG	Insertion	FBXW11:c.*1694dup	100%	13
5	G	GA	Insertion	MSX2:c.*1006dup	96.79%	2
6	T	TGGG	Insertion	COL10A1:c.*6_*7insCCC	91%	2
7	T	TA	Insertion	CDK13:c.*577dup	80.61%	14
7	A	AG	Insertion	MNX1:c.*165dup	100%	3
8	A	AC	Insertion	TRPS1:c.*1980_*1981insG	84.4%	7
9	G	GC	Insertion	ABL1:c.-111dup	100%	1
9	T	TC	Insertion	ABL1:c.-15_-14insC	100%	1
10	T	TA	Insertion	EBF3:c.*1330dup	90.82%	16
10	AT	ATA	Insertion	EBF3:c.*1293_*1292insT	91.3%	16
11	T	TAAAA	Insertion	PAX6:c.*3246_*3247insTTTT	86.06%	14

Interpret CNV Report for BAF Example

Sample ID	BAF Example
Gender	Female
Case Status	In Progress (Administrator, 28 May 2019)
Transitioned By	Administrator
Transitioned On	28 May 2019
#	323
Genome Build	hg19:GRCh37:Feb2009
ISCN Notation	arr[hg19] 2p25.3p11.2(28080_88978979)x3, 2q33.2q35(204598942_219263558)x1, 2q35q37.1(219992599_231235872)x1, 14q32.33(106535904_107286764)x0, 15q11.1q13.1(20003669_28704700)x1, 15q14q15.1(34533380_42636858)x1, 15q25.1q26.3(78845710_102399814)x1, 17p13.3p11.2(10152_21550096)x1, 18p11.32p11.31(12989_3435180)x1, 18p11.31p11.21(4979684_14096343)x1, 20p13p11.1(99591_26238040)x1, 22q11.22(22769406_23240256)x1, Xp22.33(61491_3240565)x1, Xq21.31q21.32(88569612_92304807)x1, Xq28(154948184_155234551)x1, 1q24.2q25.2(169071264_176910737)x2 hmz, 2q33.2q37.2(204838619_236025302)x2 hmz, 15q11.2q12(22933388_27800041)x2 hmz, 15q14q21.1(36259010_48287830)x2 hmz, 15q24.3q26.3(77422280_101941562)x2 hmz, 17p13.3p11.2(1135130_21047102)x2 hmz, 20p13p11.21(225414_23609009)x2 hmz
Reference ID	Female Reference
Analysis Protocol	OGT + EFSA

	ISCN Notation	# Probes	Gain/Loss	Mean Log Ratio	Classification
1	arr[hg19] 2p25.3p11.2(28080_88978979)x3	689	Gain	0.4876	Unclassified
2	arr[hg19] 2q33.2q35(204598942_219263558)x1	92	Loss	-0.6192	Unclassified
3	arr[hg19] 2q35q37.1(219992599_231235872)x1	71	Loss	-0.6337	Pathogenic
4	arr[hg19] 14q32.33(106535904_107286764)x0	11	Loss	-1.3484	Benign
5	arr[hg19] 15q11.1q13.1(20003669_28704700)x1	47	Loss	-0.7446	Unclassified
6	arr[hg19] 15q14q15.1(34533380_42636858)x1	51	Loss	-0.7244	Unclassified
7	arr[hg19] 15q25.1q26.3(78845710_102399814)x1	144	Loss	-0.6776	Unclassified
8	arr[hg19] 17p13.3p11.2(10152_21550096)x1	462	Loss	-0.7726	Unclassified
9	arr[hg19] 18p11.32p11.31(12989_3435180)x1	57	Loss	-0.7697	Unclassified
10	arr[hg19] 18p11.31p11.21(4979684_14096343)x1	151	Loss	-0.7444	Unclassified
11	arr[hg19] 20p13p11.1(99591_26238040)x1	163	Loss	-0.8038	Unclassified
12	arr[hg19] 22q11.22(22769406_23240256)x1	22	Loss	-0.6571	Benign
13	arr[hg19] Xp22.33(61491_3240565)x1	20	Loss	0.006	Unclassified
14	arr[hg19] Xq21.31q21.32(88569612_92304807)x1	24	Loss	0.1501	Unclassified
15	arr[hg19] Xq28(154948184_155234551)x1	14	Loss	-0.035	Unclassified

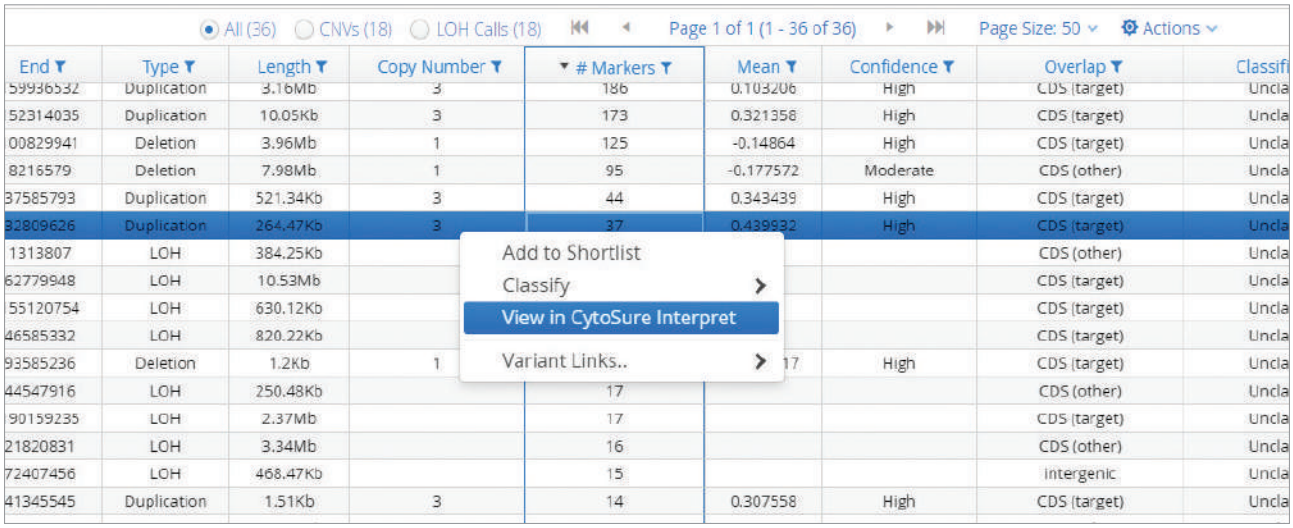
Sample	Genome Build	Chromosome	Start	End	Length	Type	Copy Number	Classification	Depth	Frequency	Overlap	Quality	Confidence	Score	# Markers	Mean	Mean Standard Error
PM	GRCh38	12	21798534	21799184	650b	Duplication	3	Unclassified	29		no CDS		High		5	0.358263	0.0655077
PM	GRCh38	Y	2786847	2787747	900b	Duplication	2	Unclassified	382		CDS (target)		High		18	0.964692	0.0662846
PM	GRCh38	Y	2807519	56884975	54.08Mb	Deletion	0	Unclassified	85	0.34	CDS (other)		High		46	-1.73419	0.22833
PM	GRCh38	X	271049	2781229	2.51Mb	Duplication	2	Unclassified	353	0.39	CDS (target)		High		263	0.91455	0.0125475
PM	GRCh38	16	89267816	89268466	650b	Duplication	3	Unclassified	259		no CDS		High		7	0.595421	0.0681251
PM	GRCh38	X	155746371	156003457	257.09Kb	Duplication	2	Unclassified	238	0.41	CDS (other)		High		33	0.946477	0.038851
PM	GRCh38	X	18636413	18636663	250b	Deletion	0	Unclassified	130		no CDS		High		5	-0.263363	0.0354837
PM	GRCh38	X	30729571	30729771	200b	Deletion	0	Unclassified	117		no CDS		High		4	-0.403956	0.0207025
PM	GRCh38	X	11759659	11759875	216b	Deletion	0	Unclassified	155		CDS (target)		High		4	-0.406383	0.0582947
PM	GRCh38	X	13712798	13713598	800b	Deletion	0	Unclassified	182	0.31	no CDS		High		16	-0.172012	0.024235
PM	GRCh38	X	23395926	23396776	850b	Deletion	0	Unclassified	183		no CDS		High		17	-0.183054	0.0251132
PM	GRCh38	X	38993103	39431026	437.92Kb	Deletion	0	Unclassified	97	0.32	no CDS		High		5	-0.303358	0.0673745
PM	GRCh38	X	41235532	41333602	98.07Kb	Deletion	0	Unclassified	184		no CDS		High		27	-0.108347	0.0215627
PM	GRCh38	X	43744370	43745370	1Kb	Deletion	0	Unclassified	168		CDS (target)		High		20	-0.107585	0.0253335
PF	GRCh38	Y	3491379	11086379	7.59Mb	LOH		Unclassified			CDS (other)			11.9	4		
PF	GRCh38	11	1983850	1994608	10.76Kb	Deletion	0	Unclassified	42	0.53	CDS (other)		High		8	-2.23335	0.129732
PF	GRCh38	19	47840474	47840674	200b	Duplication	3	Unclassified	105		no CDS		High		4	0.502655	0.0774674
PF	GRCh38	1	1516074	1520196	4.12Kb	Duplication	3	Unclassified	314		CDS (target)		High		12	0.321889	0.0491659
PF	GRCh38	1	109597206	109597456	250b	Duplication	3	Unclassified	193		no CDS		High		5	0.39925	0.0345425
PF	GRCh38	12	21798384	21798684	300b	Duplication	3	Unclassified	31		no CDS		High		4	0.634026	0.0560846
PF	GRCh38	15	38354956	38355156	200b	Duplication	3	Unclassified	174		no CDS		High		4	0.307895	0.0438078

Figure 8: Examples of reports generated by Interpret's template and plug-in framework.

CytoSure Interpret Integration

If you are a current user of OGT's CytoSure microarrays and CytoSure Interpret, you can also connect your Interpret NGS Analysis software with your existing CytoSure Interpret database to enable interpretation of your CytoSure NGS panel data in CytoSure Interpret:

- Automatically add cases processed in Interpret to your CytoSure Interpret database.
- Open and interpret these cases in the same way as your CytoSure microarrays.
- View variants detected in CytoSure NGS Panel data in CytoSure Interpret with a single click from within the Interpret user interface (Figure 9).



End	Type	Length	Copy Number	# Markers	Mean	Confidence	Overlap	Classif
59936532	Duplication	3.16Mb	3	186	0.103206	High	CDS (target)	Uncla
52314035	Duplication	10.05Kb	3	173	0.321358	High	CDS (target)	Uncla
00829941	Deletion	3.96Mb	1	125	-0.14864	High	CDS (target)	Uncla
8216579	Deletion	7.98Mb	1	95	-0.177572	Moderate	CDS (other)	Uncla
37585793	Duplication	521.34Kb	3	44	0.343439	High	CDS (target)	Uncla
32809626	Duplication	264.47Kb	3	37	0.439932	High	CDS (target)	Uncla
1313807	LOH	384.25Kb					CDS (other)	Uncla
62779948	LOH	10.53Mb					CDS (target)	Uncla
55120754	LOH	630.12Kb					CDS (target)	Uncla
46585332	LOH	820.22Kb					CDS (target)	Uncla
93585236	Deletion	1.2Kb	1	17		High	CDS (target)	Uncla
44547916	LOH	250.48Kb		17			CDS (other)	Uncla
90159235	LOH	2.37Mb		17			CDS (target)	Uncla
21820831	LOH	3.34Mb		16			CDS (other)	Uncla
72407456	LOH	468.47Kb		15			Intergenic	Uncla
41345545	Duplication	1.51Kb	3	14	0.307558	High	CDS (target)	Uncla

Figure 9: Easy link to open and view CytoSure NGS data in CytoSure Interpret software.

Ordering information

Product	Contents	Cat. No.
CytoSure Constitutional NGS Solution (24)	Bundle of 1x CytoSure Constitutional NGS Panel (24), 1x CytoSure NGS Library Preparation Kit (24) and 1x CytoSure NGS Hybridisation & Wash Kit (24)	502005-B24
CytoSure Constitutional NGS Solution (96)	Bundle of 1x CytoSure Constitutional NGS Panel (96), 1x CytoSure NGS Library Preparation Kit (96) and 1x CytoSure NGS Hybridisation & Wash Kit (96)	502005-B96

ogt
A Sysmex Group Company

Oxford Gene Technology, Begbroke Science Park, Woodstock Road, Begbroke, Oxfordshire, OX5 1PF, UK
T: +44(0)1865 856800 (US: 914-467-5285) E: products@ogt.com W: www.ogt.com

CytoSure: For Research Use Only; Not for Diagnostic Procedures.