

Copy Number Variation (CNV) detection with NGS

SEQUENCE Pilot ^{CE} module SeqNext

Version 3.5.0
(August 2011)



developed by

JSI medical systems GmbH

Friedhofstr. 5

77971 Kippenheim

GERMANY

phone: +49 78 25 / 86 41 15

fax: +49 78 25 / 86 49 46

email: mail@jsi-medisys.com

JSI medical systems Corp.

1901 Newport Blvd., Suite 350

Costa Mesa, CA 92627

USA

phone: +1-949-999-2092

fax: +1-949-999-2093

email: mail-us@jsi-medisys.com

web: www.jsi-medisys.com

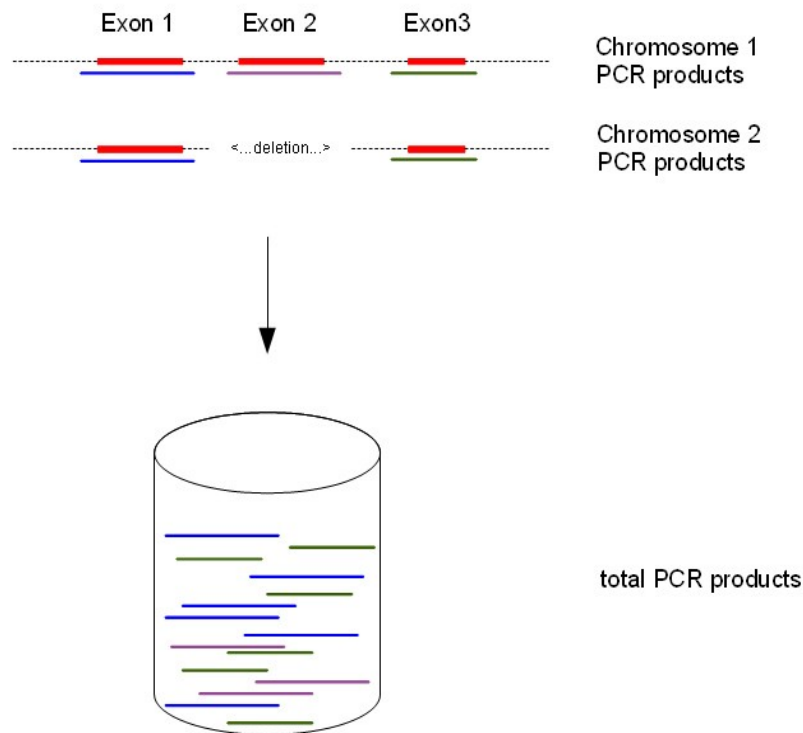
(for research use only)

Table of Contents

1	Background.....	2
2	Required method for CNV detection.....	3
3	Calculation of relative coverage values.....	4
4	How to set up CNV analysis in SeqPilot.....	6
4.1	Genes.....	6
4.2	PCR Products [master file].....	6
4.3	PCR Groups [master file].....	6
4.4	Analysis Mode CNV [master file].....	7
4.5	Operation Run	7
4.6	Operation Joining and Sequence.....	7
4.6.1	Dialogue PCR Group.....	7
4.6.2	CNV Window.....	8
4.6.2.1	Controls.....	8
4.6.2.2	Analysis Modes.....	8
4.6.2.3	Result table.....	8
4.6.2.4	Histograms.....	8
5	Final remarks.....	9

1 Background

Copy Number Variations (CNVs) correspond to relatively large regions of the genome that have been deleted or duplicated. This phenomenon can be limited to a single gene or include a contiguous set of genes. For many years these genomic deletions/duplications have been recognized as pathogenic mutations. Therefore, the determination of gene dosage in combination with sequence variation detection becomes very important in clinical medicine.

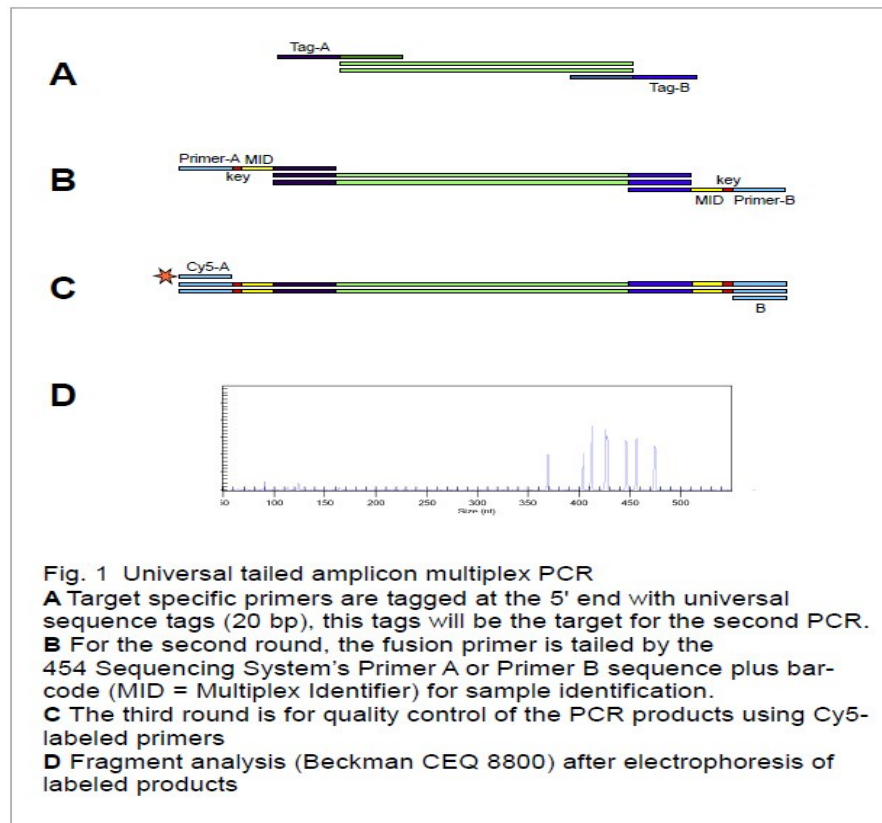


The illustration shows, that the deletion of a genomic region on one chromosome, that is amplified by PCR, does not lead to total dropout of this PCR product. Only the number of PCR products for this region may change. In the illustration for example only 50% of the expected PCR product for exon 2 are present. This reduction of PCR products cannot be detected by usual sequence variation detection.

In order to detect that the PCR product of exon 2 in the illustration is only 50% of the expected value, the PCR product amount (= coverage) of patient samples has to be compared to the PCR product amount of control samples.

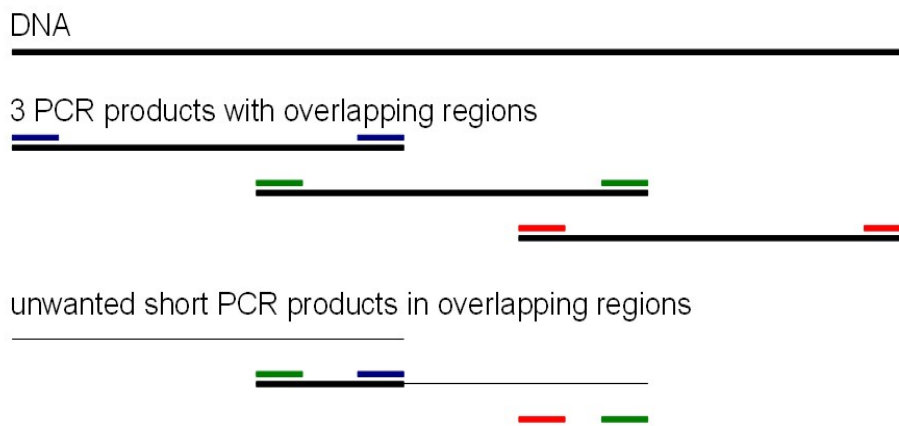
2 Required method for CNV detection

The most common method used in detection of sequence variations is multiplex PCR (Fig 1, Haag et al.).



For reasons such as primer dimers, different melting temperatures..., often not all PCR reactions can be run in ONE multiplex PCR.

If a region of interest is very large and requires multiple amplifications with overlapping regions, minimally TWO multiplex PCR reactions are needed in order to avoid undesirable short PCR products (figure below). For example one multiplex PCR will include fragments No. 1,3,5,7,9,... and the other multiplex PCR will include fragments No. 2,4,6,8,...



In case several multiplex PCRs are used the following rules have to be regarded in SeqNext (see chapters 8.3.1.2 and 9.4 of the Manual - SeqNext for more details):

- First, if there are multiple overlapping PCR products for one large DNA sequence, they must be defined as individual “Amplicon Parts” in “PCR products [master file]”.
- Second, each multiplex PCR must be defined as a separate analysis mode for one PCR group in “Analysis Mode CNV [master file]”.

If there are multiple “Amplicon Parts” defined for a PCR product, they are shown separately in “Analysis Mode CNV [master file]”. Target PCR products of a given multiplex PCR cannot be compared to control PCR products of another multiplex PCR. Each multiplex PCR is analyzed separately and must include target AND control PCR products.

3 Calculation of relative coverage values

Various experimental factors could lead to different absolute amounts of PCR products between different multiplex PCRs. However, the relative values between different PCR products within one multiplex PCR are constant (Fig. 2). Consequently it is very important to have control PCR product from gene(s) without copy number changes such as housekeeping genes for the target PCR product to compare to for establishment of the relative values.

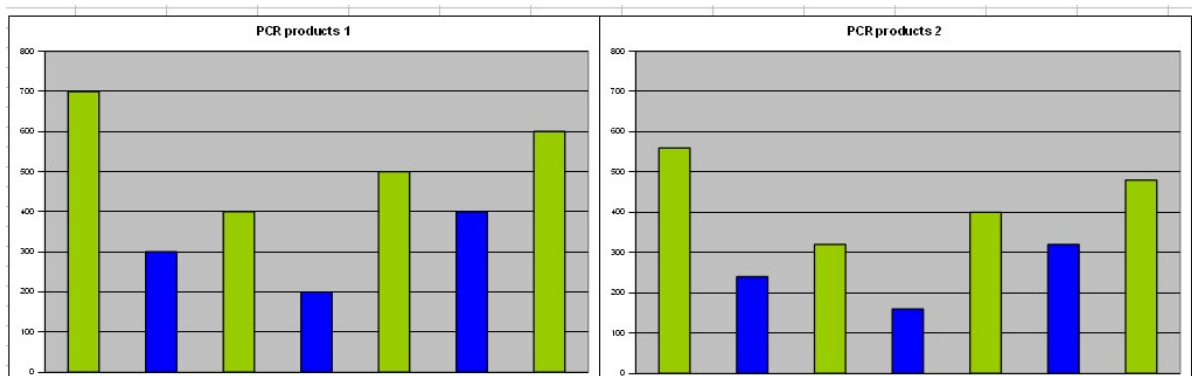


Fig. 2 Absolute values of control and target PCR products. PCR products 2 shows a 20% decrease in absolute values compared to PCR products 1.

The average coverage for all control PCR products is calculated first for each multiplex PCR. Then the relative coverage for each target PCR product is calculated against the average coverage of the control PCR products. These calculations are applied to both the control sample(s) and every patient sample to obtain the relative coverage value for each target PCR product. If multiple control samples were used, the average coverage and standard deviation for each target PCR product will be calculated.

Then the relative coverage of every target PCR product of a patient sample is compared to the average relative coverage of the control samples for the same target PCR product. If the value for a patient target PCR product falls below the pre-defined lower limit (75%) or exceeds the upper limit (125%) of the average control sample coverage, it is regarded to be a deletion or a duplication.

Examples:

multiplex PCR control DNA 1:

Control-Absolute coverage (PCR fragments)	Target-Absolute coverage (PCR fragments)	Relative coverage to Median
300	700	2.33 (700 / 300)
200	400	1.33 (400 / 300)
400	500	1.66 (500 / 300)
	600	2.00 (600 / 300)
Median = 300		

multiplex PCR control DNA 2:

Control-Absolute coverage (PCR fragments)	Target-Absolute coverage (PCR fragments)	Relative coverage to Median
150	360	2.40 (360 / 150)
100	220	1.46 (220 / 150)
200	270	1.80 (270 / 150)
	280	1.86 (280 / 150)
Median = 150		

Then the calculated average relative values for the control samples are:

- 1) 2.365 $((2.33 + 2.40) / 2)$
- 2) 1.395 $((1.33 + 1.46) / 2)$
- 3) 1.730 $((1.66 + 1.80) / 2)$
- 4) 1.930 $((2.00 + 1.86) / 2)$

multiplex PCR patient 1:

Control-Absolute coverage (PCR fragments)	Target-Absolute coverage (PCR fragments)	Relative coverage to Median
330	760	2.30 (760 / 330)
220	410	1.24 (410 / 330)
440	275	0.83 (275 / 330)
	655	1.98 (600 / 330)
Median = 330		

The percentage of the relative coverages compared to the average control values are:

1. 97.251% $((2.30 * 100) / 2.365)$
2. 88.888% $((1.24 * 100) / 1.395)$
3. **47.976%** $((0.85 * 100) / 1.730)$ → **DELETION !!!**
4. 102.590% $((1.98 * 100) / 1.930)$

Therefore, we have a deletion for fragment No. 3. All other fragments are within the range of 75% to 125%.

4 How to set up CNV analysis in SeqPilot

CNV analysis in SeqPilot is an additional function of variations/mutation analysis. These functions are calculated independent of each other in the software and are combined into one analysis result.

4.1 Genes

For genes used, either as control or target, a gene file must be defined in SeqPilot (see Manual SeqNext – Chapter 7.2) or a genome has to be deposit in the software (see Manual SeqNext – Chapter 2).

4.2 PCR Products [master file]

Control and target PCR Products must be defined individually in the operation “PCR Products [master file]”. If there are multiple overlapping PCR products for one large DNA sequence, these have to be defined as “Amplicon Parts”.

4.3 PCR Groups [master file]

Create a PCR Group that includes both control and target PCR Products of the CNV analysis in the operation “PCR Groups [master file]”.

4.4 Analysis Mode CNV [master file]

In the operation “Analysis Mode CNV [master file]” the analysis modes can be defined. The target and control PCR products (housekeeping genes) are established here for result calculation. Limits can be set for copy number changes, default values are lower-75%, upper-125% and multiple copies-175%. If a PCR Group with a defined CNV analysis mode is chosen in the operation *Run*, a CNV analysis and a normal analysis is performed automatically.

Analysis modes must be defined based on your analysis:

You can define several analysis modes for one PCR Group. In case you did several multiplex PCRs for your analysis **one** multiplex PCR should be defined as **one** analysis mode (field *Analysis Mode Name*).

For one analysis mode the control PCR Products should be defined as control, the target PCR Products as target. The relative values between control and target PCR products are then calculated in the analysis.

In case amplicon parts were defined for the PCR products all amplicon parts are listed separately in the PCR Group Products table with consecutive numbers. Therefore amplicon parts can then be separately selected to define the Analyse Mode Products.

4.5 Operation Run

Load the file(s) that include the control and target patients in the operation Run. Define samples and make sure the control sample(s) are listed as *control* in the type column. Select the PCR group that contains the pre-defined control and target genes for CNV analysis. Remember that normal analysis is performed automatically, therefore, other PCR groups or PCR products can also be selected for analysis.

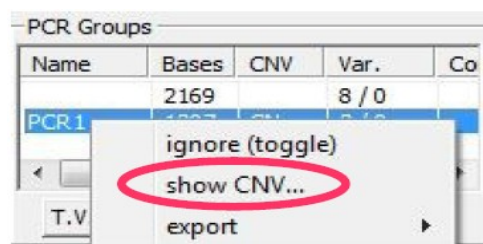
4.6 Operation Joining and Sequence

Control sample(s) should be technically validated first. Therefore select the control sample order(s) in the Joining operation and switch to Sequence operation. Then check all controls and technically validate them as done for normal analysis.

The operation Sequence is the same as the normal analysis. Additionally the *CNV window* can be opened in the dialogue *PCR Group* if a CNV PCR Group is analyzed.

4.6.1 Dialogue PCR Group

This dialogue lists all PCR Groups defined for the Run. When the first line (column Name blank) is selected all PCR groups and PCR products can be viewed. Each individually defined PCR groups are listed below.



There is the entry *CNV* in the *CNV* column in case a *CNV* analysis was performed. Right-click on a *CNV* group and select the entry “show *CNV*...” to open the *CNV* window.

4.6.2 CNV Window

CNV analysis is displayed in this window which allows for manually editing.

4.6.2.1 Controls

The *Control* table lists all joined control result files. Samples defined as controls (*control*) must be technically validated before patients samples (*patient*) are analyzed. Then all technically validated controls loaded on the same date as patients are joined to patient orders automatically. Controls can be joined and removed manually with the *[Control settings...]* button. There are three tabs in this dialogue:

- *Date*: assigns controls loaded at the same day or controls loaded at the selected date range.
- *Default*: shows a list of all default controls with the same PCR Group. Selected controls can be assigned.
- *Selection*: shows a list of all control result files with the same PCR Group (technical validated or not). Selected control can be assigned.

4.6.2.2 Analysis Modes

Lists all defined analysis modes. The detected copy number changes of *active* analysis modes are listed in the *Variation/Mutation* table. The *Result table* and the *histograms* belong to the selected analysis mode.

4.6.2.3 Result table

This table lists all target PCR products defined for the selected analysis mode and the corresponding results. There is an entry in the column *Result* if deletions, duplications or multiple copies were detected. The mutations listed in this table will also be listed in the *Variation/Mutation table* in case the analysis mode is set *active* in the *Analysis Modes* table.

Calculation of the results:

The calculation of the relative coverage for target PCR products of control sample(s) (RPC C.) and patient sample (RPC P.) and the calculation of the ratio relative coverage (ratio RPC) is explained in chapter 3.

If the ratio RPC is around 100% there is probably no copy number change. There will be a copy number change shown if the ratio RPC exceeds the upper and multiple copies limit or falls below the lower limit defined in the operation “Analysis Mode *CNV* [master file]”. These values are also reflected in the *histograms*.

4.6.2.4 Histograms

The upper histogram shows the relative coverage of the patient sample in green and the averaged relative coverage of the control sample(s) in blue. A tool tip box will appear if the mouse pointer is over a bar showing the RPC value of the patient sample, averaged control sample(s) and the standard deviation. The lower histogram shows the ratio RPC in light blue

if it is within the limits defined in the operation “Analysis Mode CNV [master file]”, and in darker blue when it exceeds the limits indicated by the red line.

Right-click a ratio RPC bar, the following context menu appears:

- Set limit value selected entry: changes the limits of the selected PCR product.
- Set limit value all entries: changes the limits of all PCR products.
- Ignore entry (toggle): only active with entries that exceed the limit. If this function is activated, the corresponding entry in the Result column of the Result table will be written in parenthesis with the remark “ignore” in front. The ignored copy number change will not be listed in the *Variation/Mutation* table nor will it be printed in the report.

5 Final remarks

The design of the multiplex PCR has to be done by the customer. Detecting CNV is based on the increase or decrease in coverages due to insertions or deletions. If there are no deviations in coverage, CNV cannot be detected. The success of CNV detection is solely depended on the PCR strategy, not on the software. Therefore JSI cannot guaranty that CNV can be detected for all multiplex PCRs.