Coverage Report

Coverage report is also available for data nodes containing aligned reads (*Aligned reads, Trimmed reads,* or *Filtered reads*). The purpose of the report is to understand how well the genomic regions of interest are covered by sequencing reads for a particular analysis.

When setting up the task (Figure 1), you first need to specify the *Genome build* and then a *Gene/feature annotation* file, which defines the genomic regions you are interested in (e.g. exome or genes within a panel). The *Gene/feature annotation* can be previously associated with Partek® Flow® via *Library File Management* or added on the fly.

Complete coverage report will contain percentage of bases within the specified genes / features with coverage greater than or equal to the coverage levels defined under *Add minimum coverage levels*. To add a level, click on the **green plus** ✚. Alternatively, to remove it, click on the **red cross** ✖ icon.

As for the *Advanced* options, if *Strand-specificity* is turned on, only reads which match the strand of a given region will be considered for that region's coverage statistics.

Generate target enrichment graphs will generate a graphical overview of coverage across each feature.

When **Use multithreading** is checked, the computation will utilize multiple CPUs. However, if the input or output data is on some file systems like GPFS file system, which doesn't support well on multi-thread tasks, unchecking this option will prevent task failures.

Figure 1. Setting up Coverage report. The example on the figure shows the use of a custom annotation file, CRCTargets, which defines target regions for a targeted resequencing panel

Coverage report result page contains project-level overview and starts with a summary table, with one sample per row (Figure 2). The first few columns show the percentage of bases in the genomic features which are covered at the specified level (or higher) (default: 1×, 20×, 100×). *Average coverage* is defined as the sum of base calls of each base in the genomic features divided by the length of the genomic features. Similarly, *Average quality* is defined as the sum of average quality of those bases that cover the genomic features, divided by the length of covered genomic features. The last two columns show the number of *On-tarted reads* (overlapping the genomic features) and *Off-target reads* (not overlapping the features). The *Optional columns* link enables import of any meta-data present in the data table (Data tab).

Annotation model CRCTargets
Strand-specificity Yes

Sample name	1x	20x	100x	Average coverage	Average quality	On-target reads	Off-target reads
09-129-N_R1	86.97%	68.08%	49.88%	134.57	36.41	2,287,428	11,003,526
09-287-N_R1	88.28%	69.46%	51.37%	136.40	36.40	2,327,443	10,923,264
09-128-T_R1	85.88%	64.21%	35.41%	75.02	36.02	1,300,921	5,662,048
09-29-T_R1	85.36%	64.25%	34.87%	74.37	36.11	1,294,784	5,579,635
09-129-T_R1	88.06%	68.48%	45.34%	102.57	36.38	1,767,675	7,995,955
09-287-T_R1	87.76%	68.94%	50.77%	130.86	36.47	2,290,821	11,034,455
09-128-N_R1	84.79%	62.46%	29.91%	65.87	36.06	1,158,134	4,960,458
09-29-N_R1	85.12%	62.93%	30.66%	67.68	36.02	1,180,681	5,086,942
09-297-T_R1	90.31%	74.04%	59.36%	225.72	36.26	3,963,391	22,279,334
09-297-N_R1	87.90%	68.92%	50.87%	132.60	36.46	2,294,950	10,472,740

Rows per page 25 ▾ |◀ ◀◀ (1 of 1) ▶▶ ▶| **Download**

Figure 2. Coverage report table showing one sample per row and key coverage metrics on columns

Quantification of on- and off-target reads is also displayed in the column chart below the table (Figure 3), showing each sample as a separate column and fraction of on-/off-target reads on the y-axis.
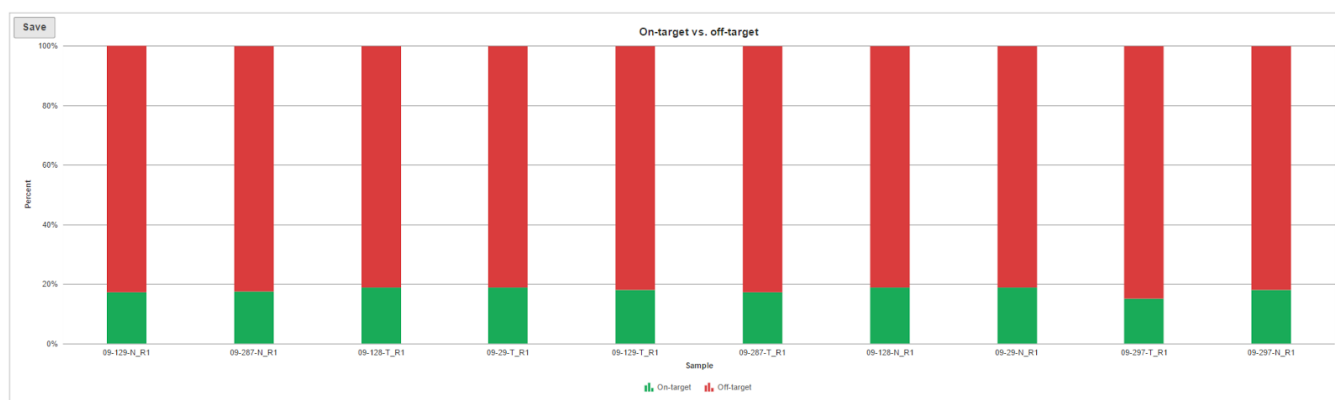

Figure 3. On-target vs. off-target chart. Each column is a sample. Fraction of reads on- or off-target is on the y-axis

Region coverage summary hyperlink opens a new page, with a table showing average coverage for each region (rows), across the samples (columns) (Figure 4).

Region name	09-128-N_R1	09-128-T_R1	09-129-N_R1	09-129-T_R1	09-287-N_R1	09-287-T_R1	09-297-N_R1	09-297-T_R1	09-29-N_R1	09-29-T_R1	
NM_000038	101.81	101.58	197.35	134.44	208.25	237.86	195.12	260.86	105.22	109.97	⬚
NM_000051	83.29	79.94	156.91	132.48	158.60	171.30	160.58	334.94	85.22	91.00	⬚
NM_000057	90.94	97.30	151.06	116.47	160.45	159.24	169.50	243.40	89.82	76.39	⬚
NM_000059	104.75	148.97	204.88	336.17	209.99	191.88	205.16	527.85	109.56	169.56	⬚
NM_000075	39.25	59.74	78.76	69.06	88.63	93.83	92.39	161.78	39.13	47.82	⬚
NM_000077	8.36	9.06	16.47	13.74	20.87	16.47	18.12	35.16	7.22	10.45	⬚
NM_000118	4.01	4.85	5.75	6.41	7.87	6.44	8.32	15.35	4.45	3.56	⬚
NM_000122	36.47	41.73	65.26	59.22	73.06	68.59	74.53	134.75	42.72	44.08	⬚
NM_000123	48.49	66.98	104.08	73.93	105.16	95.31	101.83	261.82	46.28	77.88	⬚
NM_000127	44.96	48.34	86.37	99.41	88.72	102.49	86.22	180.70	43.39	54.45	⬚
NM_000135	23.27	21.99	36.29	33.82	42.09	41.09	46.44	74.21	21.63	27.45	⬚

Figure 4. Region average coverage summary (truncated). Regions are on rows, samples on columns. The icon in the right-most column opens a coverage graph for that region

The **browser icon** in the right-most column (⬚) of the *Region average coverage summary* table opens the *Coverage* graph for the respective region (Figure 5). The horizontal axis is the normalized position within the genomic feature, represented as 1st to 100th percentile of the length of the feature. The vertical axis is coverage. Each line on the plot is a single sample, and the samples are listed below the plot.
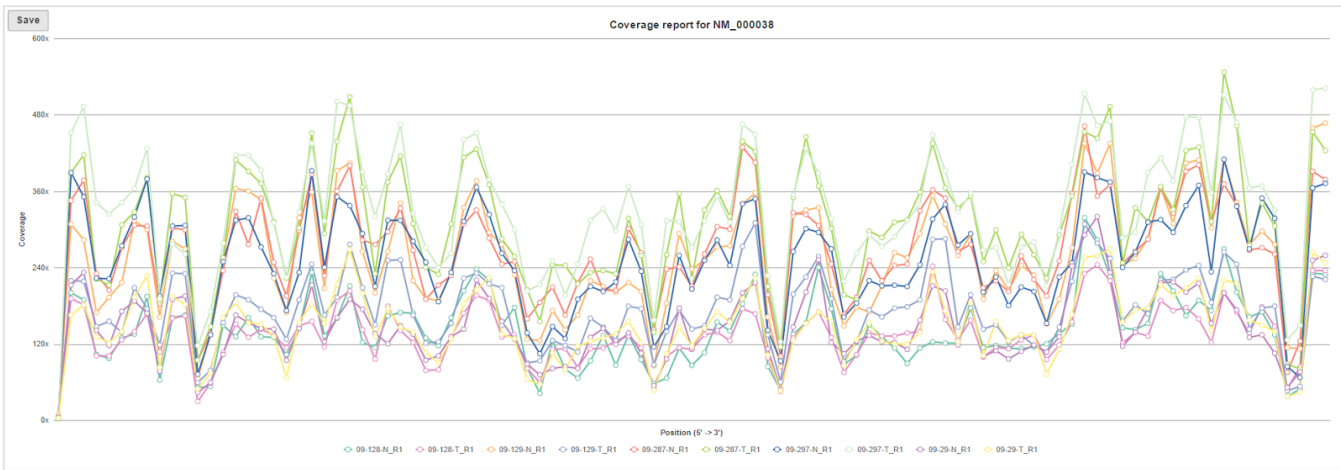
Figure 5. Coverage report of a genomic region (in this example NM_000038). The horizontal axis is the normalized position within the genomic feature, represented as 1st to 100th percentile of the length of the feature. The vertical axis is coverage. Each line on the plot is a single sample, and the samples are listed below the plot

The **Coverage summary** (Figure 6) plot is an overview of coverage across of the targeted genomic features for all the samples in the project. Each line within the plot is a single sample, the horizontal axis is the normalized position within the genomic feature, represented as 1st to 100th percentile of the length of the feature, while the vertical axis show the average coverage (across all the features for a given sample).
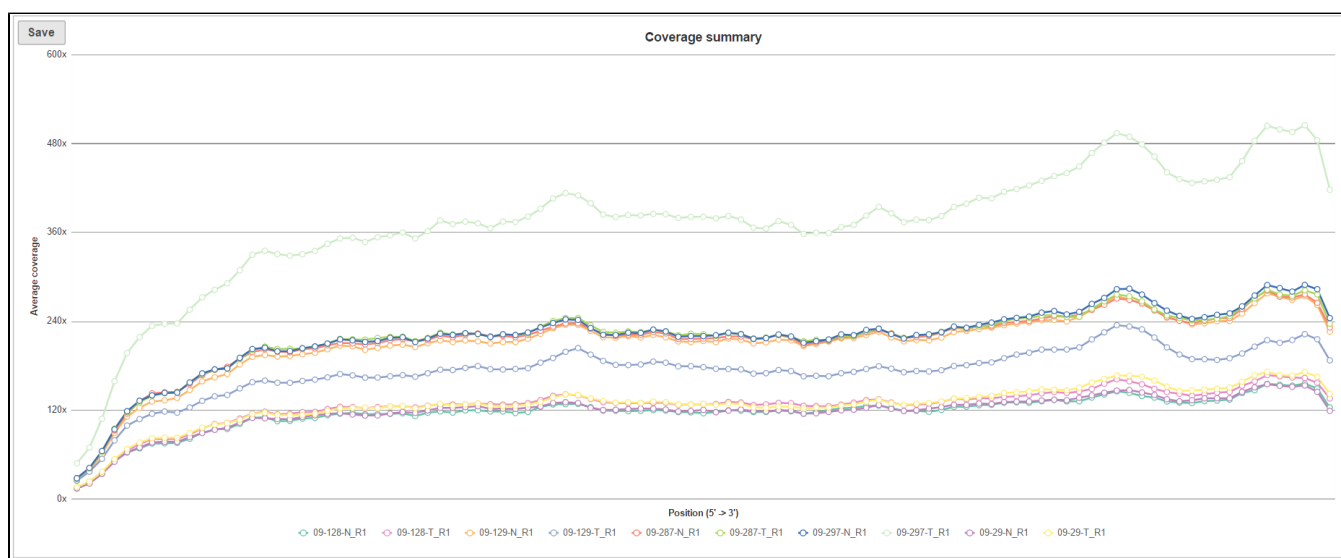
Figure 6. Coverage summary plot with sample shown as lines, normalised genomic position on the horisontal axis and average coverage on the vertical axis

If you need more details about a sample, click on the **sample name** in the Coverage report table (Figure 7). The columns are as follows:

- **Region name**: the genomic feature identifier (as specified in the annotation file)
- **Chromosome**: the chromosome of the genomic feature (or region)
- **Start**: the start position of the genomic feature (1-based)
- **Stop**: the stop position of the genomic feature (2-based, which means the stop position is exclusive)
- **Strand**: the strand of the genomic feature
- **Total exon length**: the length of the genomic feature
- **Reads**: the total number of reads aligning to the genomic feature
- **% GC**: the percentage of GC contents of those reads aligning to the genomic feature
- **% N**: the percentage of ambiguous bases (N) of those reads aligning to the genomic feature
- **(n)x**: the proportion of the genomic feature which is covered by at least n number of alignments. [Note: n is the coverage level that you specified when submitting *Coverage report* task, defaults are 1×, 20×, 100×]
- **Average coverage**: the average sequencing depth across all bases in the genomic feature
- **Average quality**: the average quality score across covered bases in the genomic feature

- : the invokes the *Coverage graph* across the genomic feature, showing the current sample only (Figure 23) (or mouse over to get a preview of the plot)

- : the invokes the *Chromosome view* and browses to the genomic location

Region name	Chromosome	Start	Stop	Strand	Total exon length	Reads	% GC	% N	1x	20x	100x	Average coverage	Average quality	
NM_022114	1	2,985,742	3,355,186	+	8,723	6,752	42.63%	.01%	88.79%	40.03%	18.37%	49.95	35.50	
NM_199454	1	2,985,742	3,355,186	+	8,666	6,752	42.61%	.01%	88.71%	40.30%	18.49%	50.20	35.50	
NM_005356	1	32,716,840	32,751,769	+	2,119	851	53.12%	.01%	78.43%	28.08%	0%	15.52	34.46	
NM_001042771	1	32,739,712	32,751,769	+	2,102	843	53.25%	.01%	81.83%	32.45%	0%	16.96	33.95	
NM_005373	1	43,803,475	43,820,136	+	3,645	2,831	56.19%	.04%	55.94%	33.91%	18.49%	39.15	34.61	
NM_001048171	1	45,794,914	45,806,143	-	1,888	854	56.46%	.01%	91.95%	34.11%	1.32%	19.44	34.63	
NM_001048172	1	45,794,914	45,805,788	-	1,779	842	56.36%	.01%	85.05%	36.20%	1.41%	20.40	34.63	
NM_001048173	1	45,794,914	45,805,788	-	1,776	842	56.34%	.01%	85.02%	36.26%	1.41%	20.40	34.64	
NM_001048174	1	45,794,914	45,805,630	-	1,695	842	56.34%	.01%	89.09%	37.99%	1.47%	21.37	34.64	
NM_001128425	1	45,794,914	45,806,143	-	1,930	854	56.61%	.01%	92.12%	33.37%	1.30%	19.38	34.58	
NM_001293190	1	45,794,914	45,806,143	-	1,891	854	56.48%	.01%	91.96%	34.06%	1.32%	19.44	34.62	
NM_001293191	1	45,794,914	45,805,630	-	1,728	842	56.46%	.01%	89.29%	37.27%	1.45%	21.29	34.60	

Figure 7. Coverage report table, sample-level (truncated). Each genomic feature is a row, coverage metrics are on columns (an example is shown)

Additional Assistance

If you need additional assistance, please visit our support page to submit a help ticket or find phone numbers for regional support.

Your Rating: ☆☆☆☆☆   Results: ★★★★★ 43 rates