

Fusion Gene Detection

A fusion gene is a hybrid gene that combines parts of two or more original genes. They can form as a result of chromosomal rearrangements (such as translocation, interstitial deletion, or chromosomal inversion) or abnormal transcription and have been shown to act as drivers of malignant transformation or/and progression in various neoplasms (1). The discovery and characterization of fusion genes have been greatly facilitated by the use of NGS (2) and several computational algorithms have been developed to detect them.

This chapter covers will illustrate how to detect fusion genes by:

- [STAR Algorithm](#)
- [TopHat-Fusion Algorithm](#)

STAR Algorithm

General Overview

The [STAR](#) aligner also has the ability to detect fusion genes (referred to as “chimeric alignments”) (5,6). During the first phase of alignment, STAR searches for maximal mappable prefixes (seeds) of sequencing reads. In the second phase, all the seeds that align within user-defined genomic windows are stitched together. If an alignment within one genomic window does not cover the entire read sequence, STAR will try to find two or more windows that cover the entire read. This essentially results in the detection of fusion events, with different parts of reads aligning to distal genomic locations, or different chromosomes, or different strands.

STAR fusion detection is performed in two steps: chimeric alignment of reads with the STAR aligner and fusion detection with STAR-Fusion. Performing fusion detection in two steps is equivalent to running the analysis in “Kickstart” mode, as [described by the authors of STAR-Fusion](#). We recommend using STAR version 2.7.8a (see [Task management](#) to check which version you are running).

To save time, you can import the pre-built STAR-Fusion pipeline from our hosted pipeline page. This pipeline includes the two steps outlined below, where the advanced options for the STAR 2.7.8a alignment have been optimized for fusion detection according to the STAR-Fusion author's recommendations. See [Importing a Pipeline](#) for more information.

Running STAR Chimeric Alignment within Partek Flow

When performing an alignment with STAR, chimeric alignment can be activated by tick-marking the **Chimeric alignment** option in the *Advanced options* of the aligner (the *Advanced options* dialog is reached via the **Configure** link in the setup dialog). When the *Chimeric alignment* checkbox is selected, additional options specific to the fusion search algorithm are shown (Figure 11). For a discussion on the details of the options, see STAR documentation.

Chimeric alignment options

☒ Chimeric alignment

Min chimeric segment length [i](#)

20 [^](#)
[v](#)

Min total score of chimeric segments [i](#)

0 [^](#)
[v](#)

Max difference of total chimeric score from read length [i](#)

20 [^](#)
[v](#)

Min separation between best score and next [i](#)

10 [^](#)
[v](#)

Penalty for non-GT/AG chimeric junction [i](#)

-1 [^](#)
[v](#)

Min chimeric junction overhang [i](#)

20 [^](#)
[v](#)

Max gap between chimeric segments [i](#)

0 [^](#)
[v](#)

☒ Filter alignments with Ns around junction [i](#)

Max multi-alignments for main chimeric segment [i](#)

10 [^](#)
[v](#)

Max chimeric multi-alignments [i](#)

0 [^](#)
[v](#)

Multi-alignment score range *i*

1

^
v

Non-chimeric alignment score drop min *i*

20

^
v

Figure 20. Controls of the STAR fusion gene detection algorithm (aligner defaults are shown)

The output is associated with the *Chimeric junctions* data node (Figure 12), which is a part of the STAR results in addition to *Aligned reads* node and, optionally, *Unaligned reads* node.

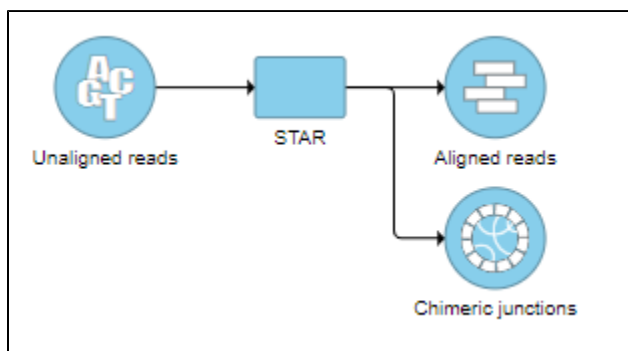


Figure 21. Chimeric results node as a result of STAR's chimeric alignment algorithm

To obtain a .fusion file that summarizes the chimeric reads across samples, select the **Chimeric results** data node and click **Download data** in the toolbox (Figure 13). The file is human-readable and can be opened in a text editor (example in Figure 14). For details refer to STAR's documentation.

X
Chimeric results

Task results

Data summary report

No actions are available for the selected data node.

Download data (12 MB)

Download icon

Figure 22. Chimeric results section of the toolbox, invocable on STAR's chimeric alignment results (data size is an example)

Sample ID	Sample Name	Chromosome	1	Start	1	Stop	1	Chromosome	2	Start	2	Stop	2	Score	Strands	Junction type	Left repeat length	Right repeat length	Read name	First base
1365	SRR2726054	chr8	76664006	76664006	chr23	44653620	44653620					10		SRR2726054.1397981	76664007	68532M	44653552	68M32S		
1364	SRR2726053	chr8	22608348	22608348	chr19	55117766	55117766			rf	0	0	2	SRR2726053.1392551	22608349	49547M	55117767	47S49M		
1366	SRR2726055	chr2	27123549	27123549	chr2	27123493	27123493			fr	0	0	5	SRR2726055.1551	27123503	46M54S	27123439	54M46S		
1364	SRR2726053	chr29	43628049	43628049	chr29	43628141	43628141			rr	0	0	1	SRR2726053.1392911	43628050	67S33M	43628074	67M33S		
1366	SRR2726055	chr23	24637123	24637123	chr23	8441379	8441379			rr	0	0	1	SRR2726055.3541	24637124	58S36M	8440355	34M966N24M36S		
1364	SRR2726053	chr13	28233139	28233139	chr10	26857705	26857705			ff	1	2	1	SRR2726053.1393361	28233095	44M39S	26857706	44S39M		
1366	SRR2726055	chr17	73881475	73881475	chr12	12416463	12416463			rf	1	1	0	SRR2726055.4271	73881476	55S44M	12416464	44S56M		
1359	SRR2726048	chr7	21937437	21937437	chr7	21937334	21937334			rr	0	0	5	SRR2726048.26926152	21937438	68S27M	21937266	68M27S		
1366	SRR2726055	chrM	6558	6558	chrM	6410	6410			rr	0	0	1	SRR2726055.8491	6526	32M68S	6411	32S68M		
1359	SRR2726048	chr19	45129846	45129846	chr19	45129565	45129565			rf	0	0	0	SRR2726048.26926512	45129847	37S28M	45129566	28S32M5S		
1366	SRR2726055	chr19	30277034	30277034	chr19	30277127	30277127			fr	0	0	2	SRR2726055.11321	30277001	33M58S	30277069	58M33S		
1359	SRR2726048	chr16	45416323	45416323	chr16	45416213	45416213			ff	0	0	0	SRR2726048.26927472	45416289	1534M30S	45416214	35S30M		
1366	SRR2726055	chr3	78248055	78248055	chr3	78248591	78248591			ff	0	0	3	SRR2726055.13951	78247984	71M29S	78248592	71S29M		
1359	SRR2726048	chr5	27173077	27173077	chr5	27173793	27173793			rr	0	0	0	SRR2726048.26930402	27173078	77S22M	27173171	42M54S35M22S		
1366	SRR2726055	chr3	88955567	88955567	chr21	57405116	57405116			rf	0	0	1	SRR2726055.17811	88955568	68S27M3S	57405117	30S25M1084N43M		

Figure 23. STAR's .fusion file opened in a text editor (example)

Running STAR-Fusion on Chimeric results

STAR-Fusion v1.10 is wrapped into Partek Flow. STAR-Fusion will process the chimeric output generated by the STAR aligner to map junction reads and spanning reads to a reference annotation set. To run fusion detection, select the **Chimeric results** data node and choose **STAR-Fusion** from the *Variant analysis* menu in the toolbox (Figure 15).



Figure 24. Choose STAR-Fusion from the menu

Choose the STAR-Fusion annotation from the drop-down list. We provide automatic downloads of the [plug-n-play libraries](#) distributed by Trinity Cancer Transcriptome Analysis Toolkit (CTAT) for Human hg38 (Gencode v22 and v37) and hg19 (Gencode v19) assemblies (Figure 16). If you wish to add your own STAR-Fusion library, you can either import a pre-build CTAT library or gather the appropriate files and build it in Partek Flow. [See here](#) for more details on the files you need.

Two side-by-side screenshots of a configuration interface titled 'Select STAR-Fusion annotation'. The left panel is for 'Homo sapiens (human) - hg38' and the right panel is for 'Homo sapiens (human) - hg19'. Both panels have a 'STAR-Fusion annotation' dropdown menu. In the left panel, the dropdown is open, showing options: 'gencode v22', 'gencode v22', 'gencode v37' (highlighted in blue), and 'Add annotation model'. Below the dropdown is an 'Advanced options' section with an 'Option set' dropdown menu set to '-- Default --' and a 'Configure' link. At the bottom of each panel are 'Back' and 'Finish' buttons.

Figure 25. STAR-Fusion task set up

To change any of the advanced options, click the **Configure** link (Figure 17). To run the task, click **Finish**.

Advanced options

Chimeric read filtering

Min percentage multimapping
50

Fusion transcript filtering

Enable filtering
☒

Min junction reads
1

Min fusion support
2

Require long double anchor support
☒

Max promiscuity
10

Min percent dominant promiscuity
20

Aggregate novel junction distance
5

Min novel junction support
3

Min spanning fragments only
5

Min alt percent junction
10.0

Minimum FFPM
0.1

Remove duplicates
☒

Skip EM
☐

Skip FFPM
☐

Annotation filter
☒

RT artifacts filter
☒

Single fusion per breakpoint filter
☒

Downstream analysis of fusion candidates

Examine coding effect
☒

Include FusionInspector
none

Examine coding effect
☐

Apply
Save as new
Cancel

Copyright © 2018 by Partek Incorporated. All Rights Reserved. Reproduction of this material without express written consent from Partek Incorporated is strictly prohibited.

Figure 26. STAR-Fusion advanced options

The resulting *Fusion predictions* task node (Figure 18) can be downloaded to your local machine by selecting the data node and clicking **Download data** from the toolbox. There will be one tab-separated (.tsv) file per sample. To view the full table, double-click the new data node to open the task report (Figure 19). Each row of the table is a fusion event and the columns contain information about each detected fusion.

- *FusionName*: the name of the fusion event, given as *LeftGene--RightGene*. Multiple fusion events can be detected across the same pair of genes, so the *FusionName* of an event is not necessarily unique;
- *JunctionReadCount*: indicates the number of RNA-Seq fragments containing a read that aligns as a split read at the site of the putative fusion junction;
- *SpanningFragCount*: indicates the number of RNA-Seq fragments that encompass the fusion junction such that one read of the pair aligns to a different gene than the other paired-end read of that fragment;
- *est_J*: estimated junction read counts corrected for multiple mappings;
- *est_S*: estimated spanning fragment counts corrected for multiple mappings;
- *SpliceType*: indicates whether the proposed breakpoint occurs at reference exon junctions as provided by the reference transcript structure annotations (Gencode);
- *LeftGene*: name of the first (left) gene;
- *LeftBreakpoint*: genome coordinates for the breakpoint in left gene;
- *RightGene*: name of the second (right) gene;
- *RightBreakpoint*: genome coordinates for the breakpoint in right gene;
- *JunctionReads*: sequence identifiers for all junction reads;
- *SpanningFragments*: sequence identifiers for all spanning fragments;
- *LargeAnchorSupport*: indicates whether there are split reads that provide 'long' (set to 25bp) alignments on both sides of the putative breakpoint;
- *FFPM*: fusion fragments per million reads
- *LeftBreakDinuc*: dinucleotide base pairs at the left breakpoint
- *LeftBreakEntropy*: the Shannon entropy of the 15 exonic bases flanking the left breakpoint
- *RightBreakDinuc*: dinucleotide base pairs at the right breakpoint
- *RightBreakEntropy*: the Shannon entropy of the 15 exonic bases flanking the right breakpoint
- *annots*: provides a simplified annotation for fusion transcript

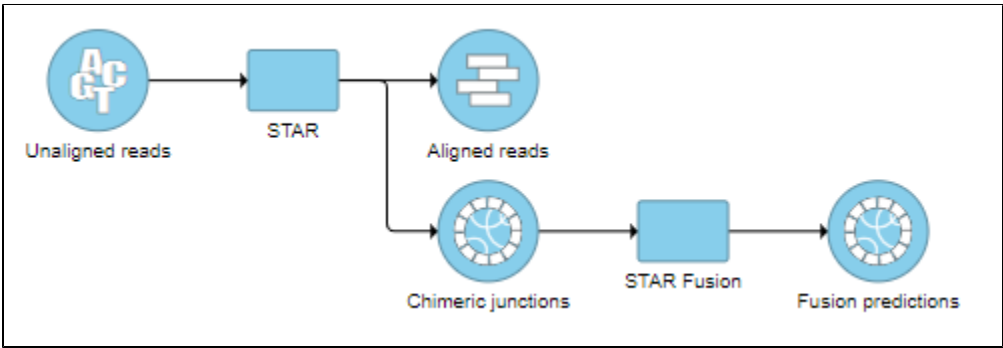


Figure 27. Fusion predictions data node

Optional columns									
#FusionName	JunctionReadCount	SpanningFragCount	est_J	est_S	SpliceType	LeftGene	LeftBreakpoint	RightGene	RightBreakpoint
HLA-S--HLA-A	24	0	24.00	0	ONLY_REF_SPLICE	HLA-S*ENSG00000225851.1	chr6:31382241-	HLA-A*ENSG00000206503.10	chr6:29945451+
ABHD12--PYGB	17	2	17.00	2.00	ONLY_REF_SPLICE	ABHD12*ENSG00000100997.17	chr20:25302219-	PYGB*ENSG00000100994.10	chr20:25295604+
USP6NL--UPF2	9	4	9.00	4.00	ONLY_REF_SPLICE	USP6NL*ENSG00000148429.13	chr10:11597631-	UPF2*ENSG00000151461.18	chr10:11956523-
MTMR3--APOH	7	5	7.00	2.19	ONLY_REF_SPLICE	MTMR3*ENSG00000100330.14	chr22:29883359+	APOH*ENSG00000091583.9	chr17:66229422-
MTMR3--APOH	7	7	7.00	4.81	INCL_NON_REF_SPLICE	MTMR3*ENSG00000100330.14	chr22:29883359+	APOH*ENSG00000091583.9	chr17:66256527-
LAMC3--TM4SF1	4	0	4.00	0	INCL_NON_REF_SPLICE	LAMC3*ENSG00000050555.16	chr9:131026441+	TM4SF1*ENSG00000169908.9	chr3:149371867-
KANSL1--ARL17B	3	0	3.00	0	ONLY_REF_SPLICE	KANSL1*ENSG00000120071.11	chr17:46094560-	ARL17B*ENSG00000228696.7	chr17:46352930-

Rows per page 25 (1 of 1) Download

Figure 28. STAR-Fusion fusion prediction table

TopHat-Fusion Algorithm

General Overview

TopHat-Fusion is a version of [TopHat](#) with the ability to align reads across fusion points and detect fusion genes resulting from breakage and re-joining of two different chromosomes or from rearrangements within a chromosome (3). It is independent of gene annotation and can discover fusion products from known genes, unannotated splice variants of known genes or completely unknown genes.

The reads are first aligned to the genome. The unaligned reads resulting from this initial alignment are split into multiple 25 bp sequences which are, in turn, aligned to the genome by Bowtie. The TopHat-Fusion algorithm identifies the cases where the first and the last 25 bp segments are aligned to either two different chromosomes or two locations on the same chromosome (spacing is defined by the user). The whole read is used to identify a fusion point. After the initial fusion candidates are defined, all the segments from the initially unaligned reads are realigned against the fusion points (as well as intron boundaries and indels). The resulting alignments are combined with the full read alignments.

The most up-to-date TopHat-Fusion version implemented in Partek® Flow® when the manual was written (2.1.0) focuses on fusions due to chromosomal rearrangements, while fusions resulting from read-through transcription or trans-splicing were not supported. For details as well as discussion of TopHat-Fusion options, see TopHat-Fusion home page (4).

Running TopHat-Fusion within Partek Flow

TopHat-Fusion is integrated in the TopHat 2 task and is invoked by using the **Fusion search** check box in the *Alignment options* dialog (Figure 1).

Select Bowtie 2 index

Assembly

Bos taurus (cow) - bosTau8

Aligner index

Whole genome

Alignment options

Generate unaligned reads

Fusion search

Advanced options

Option set

-- Default --

Configure

Back

Finish

Figure 29. Activating TopHat-Fusion algorithm for detection of fusion genes (bovine genome shown as an example)

The output is generated as a new data node *Fusion results* (Figure 2) stemming as part of the if the TopHat 2 align reads task (in addition to *Aligned reads* node and, optionally, *Unaligned reads* node).

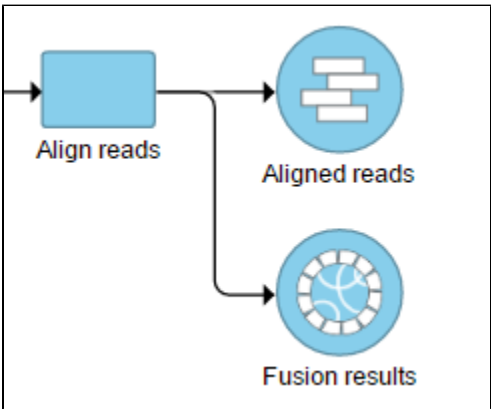


Figure 30. Fusion results node as a result of TopHat-Fusion algorithm

Selecting the **Fusion results** data node opens the task menu, with four options (Figure 3): *Data summary report*, Fusion report, Fusion attribute report, and *Download data*.

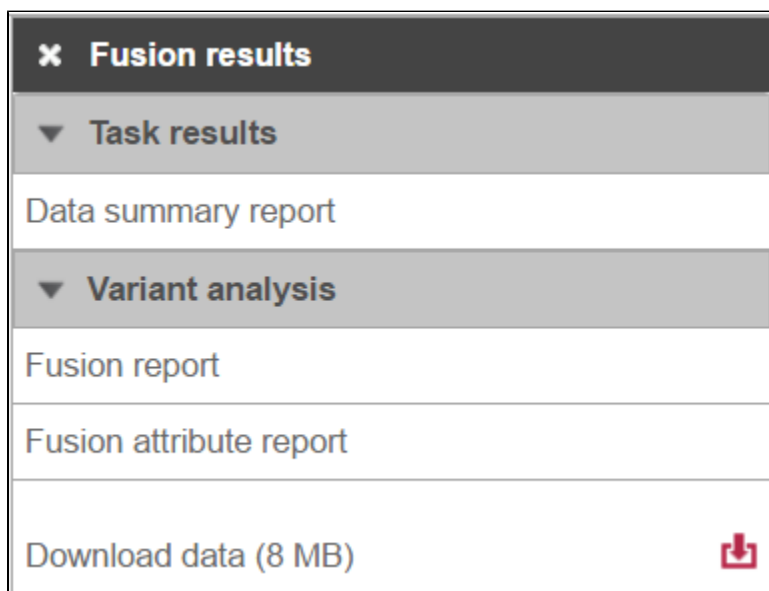


Figure 31. TopHat-Fusion results section of the toolbox, invocable on TopHat-Fusion's results (data size is an example)

Clicking the **Download data** downloads a *.fusion file to the local computer. The file is human-readable and can be opened in a text editor (example in Figure 4). For details refer to TopHat-Fusion documentation.

Sample ID	Chromosome 1	Start 1	Stop 1	Chromosome 2	Start 2	Stop 2	Score	Strands	Spanning reads	Mate pairs	Spanning mate pairs	Contradicting reads	Left bases	Right bases
1359	chr1	2457029	2457029	chr1	61073455	61073455	17.000000	fr	1	0	0	39	22	
1359	chr1	44490488	44490488	chr1	45380539	45380539	24.000000	rr	1	0	0	0	26	74
1359	chr1	55265903	55265903	chr1	81059272	81059272	18.000000	rf	1	0	0	0	25	43
1359	chr1	64878115	64878115	chr1	64881045	64881045	6.000000	fr	1	0	0	0	44	56
1359	chr1	65265854	65265854	chr1	65310830	65310830	8.000000	rf	1	0	0	0	42	58
1359	chr1	81202375	81202375	chr1	81202497	81202497	17.000000	rr	1	0	0	20	26	43
1359	chr1	83532957	83532957	chr1	83532991	83532991	8.000000	rf	1	0	0	2	29	21
1359	chr1	83585258	83585258	chr1	83585402	83585402	16.000000	fr	1	0	0	1	64	34
1359	chr1	83587841	83587841	chr1	83587923	83587923	15.000000	fr	1	0	0	25	21	36
1359	chr1	93139878	93139878	chr1	93139942	93139942	27.000000	rf	1	0	0	27	67	23
1359	chr1	93140018	93140018	chr1	93140065	93140065	19.000000	rf	1	0	0	0	42	23
1359	chr1	93140064	93140064	chr1	93140093	93140093	24.000000	fr	1	0	0	0	74	26
1359	chr1	93140131	93140131	chr1	93140187	93140187	24.000000	rr	1	0	0	0	26	73
1359	chr1	93140303	93140303	chr1	93140343	93140343	11.000000	rf	1	0	0	0	53	39
1359	chr1	104146745	104146745	chr1	104146834	104146834	10.000000	fr	1	0	0	799	22	32

Figure 32. TopHat-Fusion's .fusion file opened in a text editor (example)

A list of annotated fusion genes, in a form of *Fusion report* can be obtained by first selecting the **Fusion report** task node (Figure 2) and then the **Task report** link from the task menu (Figure 3). Since the task provides an annotated report, an annotation file needs to be specified first (Figure 5).

Select Annotation file

Assembly

Bos taurus (cow) - bosTau8

Gene/feature annotation

RefSeqBosTau8

Back

Finish

Figure 33. Selecting an annotation file to annotate TopHat-Fusion results (an example)

The resulting *Fusion report* task node (Figure 6) can be double-clicked to reveal the full table (Figure 7).

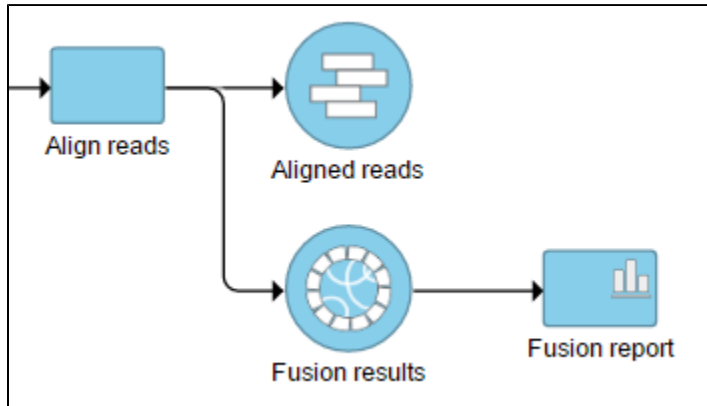


Figure 34. Fusion report task node as a result of annotating Fusion results generated by TopHat-Fusion algorithm

Each row of the table in Figure 7 is a potential fusion event, with the columns providing the following information.

- **Sample ID:** sample in which the fusion event was identified
- **Chromosome 1:** chromosome hosting the first (left) segment of the fusion transcript
- **Stop 1:** end of the first (left) segment of the fusion transcript
- **Chromosome 2:** chromosome hosting the second (right) part of the fusion transcript
- **Start 2:** beginning of the second (right) segment of the fusion transcript
- **Gene1:** gene on the left side of the fusion
- **Gene2:** gene on the right side of the fusion
- **Spanning reads:** number of reads which were unaligned during the initial phase of TopHat and where only one mate is used as evidence of the fusion event
- **Mate Pairs:** number of reads which were unaligned during the initial phase of TopHat and where both mates are used as evidence of the fusion event
- **Spanning mate pairs:** number of reads where both mates were aligned during the initial phase of TopHat, but their pairing is discordant (e.g. different chromosomes, different orientation etc.)
- **Contradicting reads:** number of reads which do not support the fusion
- **Left bases:** number of bases on the left side of the fusion
- **Right bases:** number of bases on the right side of the fusion

All the columns can be sorted by using the **arrow buttons** in column headers, while the type-in boxes can be used for searching. TopHat-Fusion does not report exact start and stop position for each side of the fusion event. It has a single location for the end of the upstream segment (**Stop 1**) and the beginning of the downstream segment (**Start 2**). Therefore, columns **Start 1** and **Stop 2** are added for (internal) consistency with other Partek Flow tools.

Disrupted Genes ☐ Gene/Gene fusions ☐										
Sample ↕	Chromosome 1 ↕	Start 1 ↕	Stop 1 ↕	Chromosome 2 ↕	Start 2 ↕	Stop 2 ↕	Gene 1 ↕	Gene 2 ↕	Spanning reads ▾	Mate pairs ↕
SRR2726052	chr25	1521772	1521772	chr3	97883754	97883754	RPS2	Intergenic	784	0
SRR2726048	chr25	1521780	1521780	chr3	97883746	97883746	RPS2	Intergenic	593	0
SRR2726048	chr5	27173890	27173890	chr7	1616286	1616286	KRT18	Intergenic	572	0
SRR2726050	chr25	1521772	1521772	chr3	97883754	97883754	RPS2	Intergenic	503	0
SRR2726055	chr1	142999251	142999251	chr25	39346135	39346135	Intergenic	ACTB	468	0
SRR2726049	chr1	142999251	142999251	chr25	39346135	39346135	Intergenic	ACTB	433	0
SRR2726048	chr10	73628749	73628749	chr5	27218568	27218568	KRT8	KRT8	433	0
SRR2726052	chr1	142999251	142999251	chr25	39346135	39346135	Intergenic	ACTB	399	0

Figure 35. Fusion report of TopHat-Fusion fusion gene detection algorithm. Each row represents a fusion gene candidate (an example is shown) (table trunc)

The checkboxes **Disrupted Genes** and **Gene/Gene fusions** are filter tools. When selected, **Disrupted Genes** removes all the rows (fusion events) which have no genes assigned to it, i.e. those that merge two intergenic regions. However, if there is a fusion between a gene and an intergenic region, it will be kept in the table. The **Gene/Gene fusions** filters in only those fusion events which have an annotated gene on both sides of the breakpoint. In other words, only gene to gene fusions are kept in the table.

Another table which can be generated based on a **Fusion results** node is the **Fusion attribute report** (Figure 3). When the option is selected, it brings up the dialog shown in Figure 8. First, you need to specify one or more categorical attributes (**Select attribute(s) to test**), which have at least two categories (see [Data tab](#)). Second, you need to specify an annotation file, using the **Assembly** and **Gene/feature annotation** drop-down lists.

Select attribute(s) to test ⓘ
☐ Conception

Select Annotation file

Assembly
Bos taurus (cow) - bosTau8 ▼

Gene/feature annotation
RefSeqBosTau8 ▼

Back
Finish

Figure 36. Selecting attributes to be tested for association with fusion events (the attribute *Conception* and the annotation files are an example)

A new data node, **Fusion attribute report**, is generated in the *Analysis* tab (Figure 9) and it provides access to the **Task report** link in the task menu.

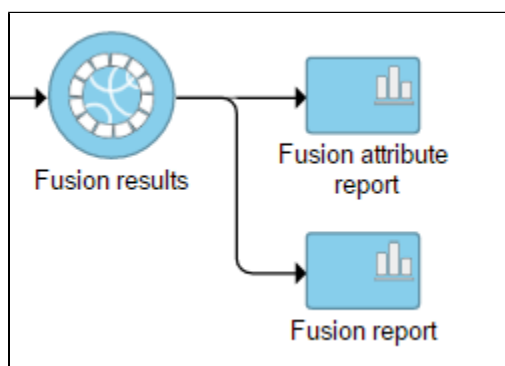


Figure 37. Fusion attribute report node as a result of annotating Fusion results generated by TopHat-Fusion algorithm

The output, *Fusion report table* (Figure 10) resembles the basic TopHat-Fusion output (Figure 7); each row of the table is a single fusion event while the information on the merged segments is on the columns.

- *Chromosome 1*: chromosome hosting the first (left) segment of the fusion transcript;
- *Start 1*: beginning of the first (left) segment of the fusion transcript;
- *Stop 1*: end of the first (right) segment of the fusion transcript;
- *Chromosome 2*: chromosome hosting the second (right) segment of the fusion transcript;
- *Start 2*: beginning of the second (right) segment of the fusion transcript;
- *Stop 2*: end of the second (left) segment of the fusion transcript;
- *Gene1*: gene on the left side of the fusion;
- *Gene2*: gene on the right side of the fusion;
- *% in (category name)*: fraction of samples within the category with the fusion event.

The checkboxes *Disrupted Genes* and *Gene/Gene fusions* are filter tools. When selected, **Disrupted Genes** removes all the rows (fusion events) which have no genes assigned to it, i.e. those that merge two intergenic regions. However, if there is a fusion between a gene and an intergenic region, it will be kept in the table. The **Gene/Gene fusions** filters in only those fusion events which have an annotated gene on both sides of the breakpoint. In the other words, only gene to gene fusions are kept in the table.

Disrupted Genes ☐

Gene/Gene fusions ☐

Chromosome 1	Start 1	Stop 1	Chromosome 2	Start 2	Stop 2	Gene 1	Gene 2	% in AI	% in SCNT
chr24	241	157987176	chr24	55	155176998	Intergenic	Intergenic	100.00%	100.00%
chr23	27143136	27160019	chr5	101331607	101331607	Intergenic	A2M	20.00%	0%
chr8	1781248	1782686	chr8	1809276	1814469	Intergenic	AADAT	20.00%	60.00%
chr11	74526474	104678215	chr2	107015487	107017081	Intergenic	AAMP	20.00%	20.00%
chr13	49582241	62225197	chr19	43635406	43635555	Intergenic	AARSD1	20.00%	0%
chr11	32468565	108435269	chr8	96274805	96310440	Intergenic	ABCA1	20.00%	60.00%
chr1	39336517	104147162	chr25	1803574	1825998	Intergenic	ABCA3	60.00%	20.00%
chr19	43218151	53329762	chrX	81158057	81158066	Intergenic	ABCB7	40.00%	20.00%
chr23	23845612	46872464	chr3	49168642	49225516	Intergenic	ABCD3	40.00%	20.00%
chr1	51715122	105890390	chr17	13334252	13358513	Intergenic	ABCE1	80.00%	100.00%

Figure 38. Fusion attribute report of TopHat-Fusion fusion gene detection algorithm. Each row represents a fusion gene candidate (the example shows com

References

1. Annala MJ, Parker BC, Zhang W, Nykter M. Fusion genes and their discovery using high throughput sequencing. *Cancer Lett.* 2013;340:192-200.
2. Costa V, Aprile M, Esposito R, Ciccodicola A. RNA-Seq and human complex diseases: recent accomplishments and future perspectives. *Eur J Hum Genet.* 2013;21:134-142.
3. Kim D, Salzberg SL. TopHat-Fusion: an algorithm for discovery of novel fusion transcripts. *Genome Biology.* 2011;12:R72
4. TopHat-Fusion. An algorithm for discovery of novel fusion transcripts. http://tophat.cbcb.umd.edu/fusion_index.html Accessed on April 25, 2014
5. Dobin A, Davies CA, Schlesinger F et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* 2013;29:15-21.
6. Haas B.J, Dobin A, Li B. *et al.* Accuracy assessment of fusion transcript detection via read-mapping and de novo fusion transcript assembly-based methods. *Genome Biol.* 2019;20:213 (2019)

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:  39 rates