GENOVESA · USER GUIDE

RNA fusions in Genovesa

Workflow **Genovesa RNA FusionPlus**: where to find the results, what the Excel and PDF contain and how to read them

Engine	RNA fusion engine v1.0
Workflow	Genovesa RNA FusionPlus v1
Annotation	version 1.7.2, NCBI RefSeq RS_2025_08, GRCh38
Genovesa version	1.10.1
Document version	v1.1 · 2026-10-30
Published by	Bioxsys s.r.o.

Contents

1. Introduction

2. What is new

3. Where to find the results in Genovesa

- 3.1 Results of one examination
- 3.2 Results of the whole project

4. Examination Excel output analysis

- 4.1 Fusions sheet
- 4.2 Technical sheet
- 4.3 Info sheet

5. Project summary Excel output analysis

6. PDF report

- 6.1 Parts of the page
- 6.2 EVIDENCE panel
- 6.3 Exon skipping page

7. How to interpret the results

- 7.1 Confidence
- 7.2 Caller support (1/2 vs 2/2)
- 7.3 Read counts and expression
- 7.4 Reading frame and domains
- 7.5 Nomenclature
- 7.6 Exon skipping and skip fraction

8. Limitations

9. Coming soon

1. Introduction

The **Genovesa RNA FusionPlus** workflow detects **fusion transcripts** in RNA sequencing data (paired-end RNA-seq). Each fusion is described at the level of RefSeq transcripts (partner genes, exons, junction, reading frame, retained protein domains), together with standardised nomenclature and the number of reads supporting it.

The fusion results also include **exon skipping** for selected clinically relevant events:

- **MET** – exon 14 skipping (METex14),
- **EGFR** – variants EGFRvIII, vIIIb, vII, vIIb, vIVa, vIVb,
- **ERBB2** – Δ16 (exon 16 skipping).

Exon skipping is reported in the **same table and the same PDF** as fusions, so all structural findings from RNA are in one place.

This guide is intended for laboratory diagnosticians and clinical geneticists. It describes **where to find the results in Genovesa, what each file, sheet and column contains and how to read the values**. The bioinformatic processing is described only as far as needed.

Which version are you looking at?

The workflow and annotation versions are shown with every result:

- **Excel** – sheet **Info**, rows *Genovesa workflow* (e.g. `Genovesa RNA FusionPlus v1`) and *Annotation version* (e.g. `1.7.2`),
- **PDF** – footer of every page (workflow name and version, annotation version).

When contacting support, please quote the project ID, examination ID and these versions.

Important

The workflow output is **supporting analytical data**, not a clinical diagnosis. The clinical significance of a finding, its confirmation and its inclusion in a report are decided by the laboratory specialist according to the laboratory's validated procedures.

2. What is new

Compared with the previous RNA workflow, Genovesa RNA FusionPlus brings the following changes:

Redesigned caller

A new fusion detection method: more accurate results with fewer artefacts and faster sample processing.

Two independent callers

Every fusion is evaluated by two independent callers. The `caller_support` column (**1/2**, **2/2**) shows how many of them confirmed it.

RefSeq annotation

NCBI RefSeq RS_2025_08 (GRCh38): versioned transcripts (NM_…), exon numbers, reading frame, retained protein domains.

HGVS and VICC nomenclature

Junction description according to HGVS (RNA level) and VICC, including junctions inside an exon (`e.N±k`).

Exon skipping

MET ex14, EGFR (vIII, vII, vIV) and ERBB2 Δ16 as part of the results, with a `skip_fraction` column. Reported from 2 reads; the clinician decides on its significance.

Clear ordering

Rows are ordered **high** → **medium** → **low**. New column `supporting_reads_total` = total number of supporting reads.

New PDF

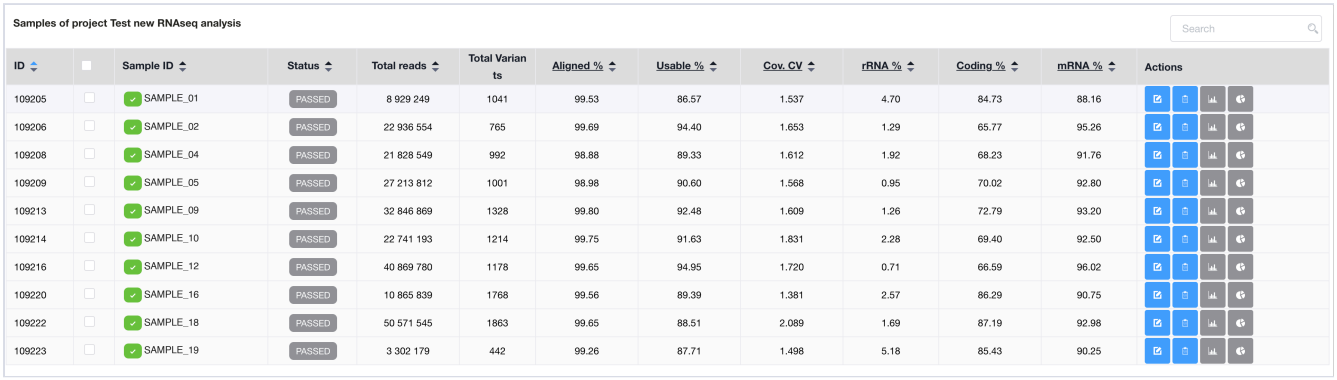
Header with a colour-coded confidence label, NOMENCLATURE and EVIDENCE panels, footer with IDs and versions, bookmarks; one page = one fusion.

Excel for clinicians

Clinically relevant columns in colour-coded groups, explanatory notes on headers, and a summary Excel for the whole project. The TSV remains for bioinformaticians.

3. Where to find the results in Genovesa

Fusion results are available at two levels: **for an individual examination** and **for the whole project**. When the analysis is complete, the examination shows the status **PASSED** in the project overview.



ID	Sample ID	Status	Total reads	Total Variants	Aligned %	Usable %	Cov. CV	rRNA %	Coding %	mRNA %	Actions
109205	SAMPLE_01	PASSED	8 929 249	1041	99.53	86.57	1.537	4.70	84.73	88.16	
109206	SAMPLE_02	PASSED	22 936 554	765	99.69	94.40	1.653	1.29	65.77	95.26	
109208	SAMPLE_04	PASSED	21 828 549	992	98.88	89.33	1.612	1.92	68.23	91.76	
109209	SAMPLE_05	PASSED	27 213 812	1001	98.98	90.60	1.568	0.95	70.02	92.80	
109213	SAMPLE_09	PASSED	32 846 869	1328	99.80	92.48	1.609	1.26	72.79	93.20	
109214	SAMPLE_10	PASSED	22 741 193	1214	99.75	91.63	1.831	2.28	69.40	92.50	
109216	SAMPLE_12	PASSED	40 869 780	1178	99.65	94.95	1.720	0.71	66.59	96.02	
109220	SAMPLE_16	PASSED	10 865 839	1768	99.56	89.39	1.381	2.57	86.29	90.75	
109222	SAMPLE_18	PASSED	50 571 545	1863	99.65	88.51	2.089	1.69	87.19	92.98	
109223	SAMPLE_19	PASSED	3 302 179	442	99.26	87.71	1.498	5.18	85.43	90.25	

Figure 1. Examinations of a project. The **Actions** column is on the right.

3.1 Results of one examination

In the examination row, click the first icon in the **Actions** column, **Edit Examination**. The window contains only the files of the selected sample, organised in the tabs FASTQ Files (download of the input `fastq.gz`), Result Files, MultiQC & Metadata Files, Other Files and BAM/BAI Files.

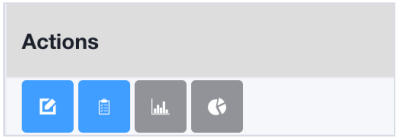


Figure 2. Actions icons of an RNA examination; the first icon opens Edit Examination.

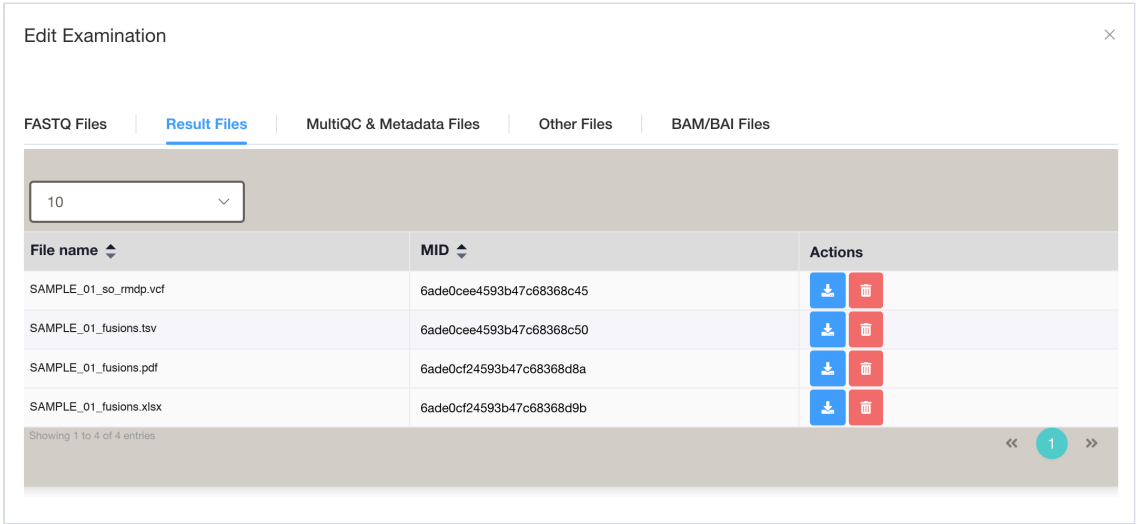


Figure 3. Edit Examination → Result Files: fusion results of the examination (`_fusions.xlsx`, `_fusions.pdf`, `_fusions.tsv`) and the VCF with RNA variants.

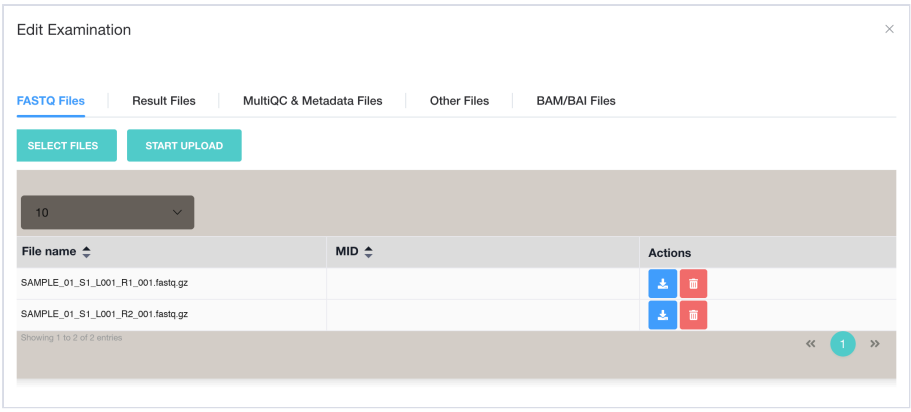


Figure 4. Edit Examination → FASTQ Files: download of the input `fastq.gz` files of the examination.

Tab	File	Purpose
Result Files	`<sample>_fusions.xlsx`	Excel for clinical review (chapter 4)
Result Files	`<sample>_fusions.pdf`	Graphical report, one page per fusion (chapter 6)
Result Files	`<sample>_fusions.tsv`	Complete table for bioinformaticians and further processing
Other Files	`<sample>_fusions_discarded.tsv`	Candidates removed by filters (technical file)
Other Files	`<sample>_refseq_provenance.log`	Record of the analysis run, versions and references (technical file)

You will not normally need the technical files in **Other Files**. They help to find out why a particular candidate was not reported and to audit the processing.

3.2 Results of the whole project

In the project overview, click the **Edit project / upload / download data** icon in **Actions**. The window has the same tabs as for an examination but contains the files of **all examinations in the project**. The fusion results (XLSX, PDF, TSV) are in the **VCF Files** tab.

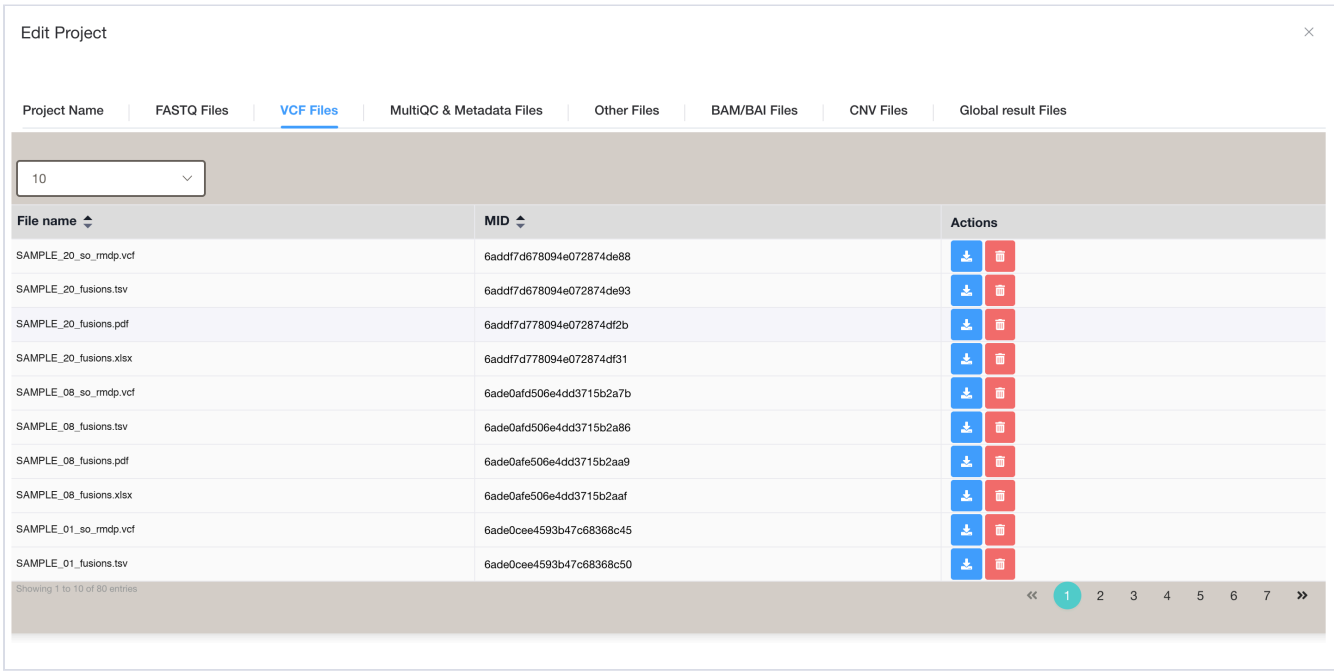


Figure 5. Edit Project → VCF Files: `_fusions.xlsx`, `_fusions.pdf` and `_fusions.tsv` of the individual examinations.

Project summary Excel output (Global Result Files)

The **Global Result Files** tab contains the summary file **`PROJECT_<project ID>_fusions_aggregated.xlsx`** with all fusions of all samples in the project (chapter 5).

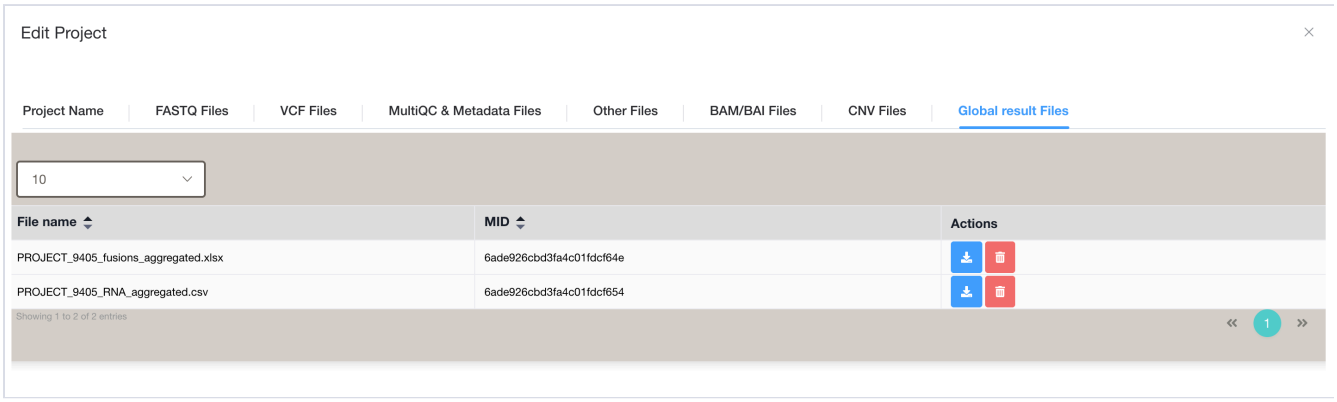


Figure 6. Edit Project → Global Result Files: fusion summary Excel of the project.

When is the summary Excel created?

- It is created automatically **as soon as all examinations in the project have completed their analysis**.
- If a sample is later added to the project or re-analysed, the summary file is **created again and replaces the previous version**. The creation date is on the Info sheet.
- Until the last examination is finished, there is no summary file in Global Result Files (or there is a version from an earlier state of the project).

4. Examination Excel output analysis

The file `<sample>_fusions.xlsx` has three sheets:

Sheet	Content
Fusions	Main table for clinical review: one row = one fusion or exon skipping event
Technical	Technical detection columns for the same rows in the same order
Info	Sample, project and examination ID, date, workflow and annotation version, reference, row counts, column legend

4.1 Fusions sheet

Working with the sheet:

- The header is **frozen** and the first column (*Fusion*) stays visible when scrolling to the right.
- A **filter** is enabled on all columns (e.g. only *Confidence* = *high* or *Caller support* = *2/2*).
- Every header has a **note with an explanation** (hover over the header). The same texts are in the legend on the Info sheet.
- Counts are stored as numbers and *Skip fraction* as a percentage, so you can sort and filter.
- Rows are ordered **high → medium → low**; within the same confidence, fusions come before exon skipping. **The order is the same as the PDF pages.**

Header colour groups: **blue – fusion and nomenclature** **green – evidence** **orange – location**

Confidence cell colour: **high** **medium** **low** (matches the label colour in the PDF).

	A	B	C	D	E	F	G	H	I
1	**Fusion**	**Confidence**	**Type**	**Reading frame**	**VICC nomenclature**	**HGVS (RNA)**	**Exon 5'**	**Exon 3'**	**Tags**
2	TCF7L2:: VTI1A(33367), TCF7L2(61512)	high	duplication	out-of-frame			4		
3	EML4::ALK	high	inversion	in-frame	NM_001410776.1(EML4):e.14:: NM_004304.5(ALK):e.20	NM_001410776.1:r.? _1522::NM_004304.5:r. 3173_?	14	20	
4	SMAD3::SMAD3-DT	high	duplication/ 5'-5'	out-of-frame			3		
5	MYC::SLAIN2	medium	translocation		NM_002467.6(MYC):e.3-2137:: NM_020846.2(SLAIN2):e.2		3	2	
6	MYC::MYC	medium	duplication		NM_002467.6(MYC):e.3-2039:: NM_002467.6(MYC):e.2+382		3	2	

Cut-out of the Fusions sheet, blue column group, first five rows.

	A	B	C	D	E	F	G	H	I	J
1	**Fusion**	**Confidence**	**Supporting reads total**	**Split reads 5'**	**Split reads 3'**	**Discordant mates**	**Coverage 5'**	**Coverage 3'**	**Skip fraction**	**Caller support**
2	TCF7L2::VTI1A(33367), TCF7L2(61512)	high	168	110	44	14	6346	72		2/2
3	EML4::ALK	high	90	47	41	2	2386	457		2/2
4	SMAD3::SMAD3-DT	high	73	31	28	14	3279	57		1/2
5	MYC::SLAIN2	medium	74	24	44	6	65535	63		2/2
6	MYC::MYC	medium	73	32	39	2	61709	65535		1/2

The same rows, green column group (evidence). A coverage of 65,535 is the highest value the report displays; it indicates very high expression (here MYC).

Column reference

Examples are from the **EML4::ALK** fusion; for exon skipping from **MET ex14**. An empty cell means the value is not available or does not apply to that type of row.

Column	Content and how to read it	Example
Fusion	Partner genes as 5′::3′ (HGNC symbols). For exon skipping GENE::GENE. If the breakpoint lies outside a gene (intergenic), the nearest genes are listed with their distance in bp in brackets. If the transcribed strand cannot be determined, the order of partners is not certain.	`EML4::ALK` `MET::MET`
Confidence	Reliability of the finding: **high**, **medium**, **low** (chapter 7.1). For exon skipping it is based on read count and skip fraction.	`high`
Type	Event type: *translocation*, *deletion*, *duplication*, *inversion*; qualifiers *read-through* (transcription continues into the neighbouring gene), *non-canonical_splicing*, *5′-5′ / 3′-3′* (both partners in the same orientation, usually without a functional product), *exon_skipping*.	`inversion` `deletion/exon_skipping`
Reading frame	Reading frame of the fusion transcript: *in-frame* (frame preserved), *out-of-frame* (frameshift), *stop-codon* (stop codon close to the junction); empty = cannot be determined.	`in-frame`
VICC nomenclature	VICC junction description at exon level. Given only when the exon boundaries are verified in the RefSeq model. `e.N` = exon N; `e.N−k` / `e.N+k` = junction k bases before the end / after the start of exon N (inside the exon).	`NM_001410776.1(EML4):e.14::NM_004304.5(ALK):e.20`
HGVS (RNA)	HGVS junction description at RNA level. Given only when the junction sequence is verified against the RefSeq transcript. `?` = unknown outer transcript bounds. For exon skipping, deletion of the skipped exons.	`NM_001410776.1:r.?_1522::NM_004304.5:r.3173_?` `NM_000245.4:r.2888_3028del`
Exon 5′ / Exon 3′	Exon of the 5′ and 3′ partner at the junction; numbering refers to the transcript in columns Transcript 5′ / 3′.	`14 / 20`
Tags	Label of a known event; for exon skipping the event name (METex14, EGFRvIII, ERBB2d16).	`METex14`
Supporting reads total	Number of fragments supporting the event = split reads 5′ + split reads 3′ + discordant mates. Each fragment is counted once, **duplicates excluded**. For exon skipping, the number of reads across the junction that skips the exon.	`90`
Split reads 5′ / 3′	Reads spanning the junction itself, anchored in the 5′ or 3′ partner. The most direct evidence of the junction.	`47 / 41`
Discordant mates	Read pairs with one read in each partner (the junction lies between them). They complement split reads.	`2`
Coverage 5′ / 3′	Number of fragments around the breakpoint (**duplicates included**), an approximate measure of the partner's expression. A low value means a weakly expressed gene; a fusion supported by few reads can then still be relevant. For exon skipping, reads of the normal junctions into and out of the skipped exon.	`2386 / 457`
Skip fraction	Exon skipping only: fraction of transcripts lacking the exon = reads across the skipping junction / (those reads + mean reads of the normal junctions).	`95.9%` (MET)

Column	Content and how to read it	Example
Caller support	`k/n`: how many of n independent callers reported the same breakpoints (chapter 7.2). Empty for exon skipping (the event is evaluated directly from reads spanning the junctions).	`2/2`
Retained protein domains	Protein domains retained in the fusion protein: domains of the 5′ partner | domains of the 3′ partner, with the retained part of each domain in brackets.	`HELP_motif(100%),…|Protein_kinase_domain(100%)`
Breakpoint 5′ / 3′	Genomic position of the breakpoint (GRCh38).	`chr2:42295516` `chr2:29223528`
Site 5′ / 3′	Where the breakpoint lies: *CDS*, *5′UTR*, *3′UTR*, *exon* (non-coding), *intron*, *intergenic*; *splice-site* = exactly at an exon boundary (typical of genuine fusions).	`CDS/splice-site`
Transcript 5′ / 3′	RefSeq transcript used for the nomenclature and exon numbering.	`NM_001410776.1` `NM_004304.5`

Exon numbering depends on the transcript

Exon numbers, VICC and HGVS always refer to the transcript in the *Transcript 5′ / 3′* columns. Literature and databases often use a different transcript, and exon numbers may then differ.

Example: EML4::ALK is reported as `NM_001410776.1(EML4):e.14::NM_004304.5(ALK):e.20`. Transcript NM_001410776.1 contains one additional short exon (exon 7), so its exon 14 corresponds to **exon 13 of transcript NM_019063.5**. This is therefore the EML4::ALK variant commonly referred to as **E13;A20 (variant 1)**. When comparing with the literature or with the result of another method, always compare the same transcript or the genomic breakpoint position.

Exon skipping rows

Exon skipping has its own row on the Fusions sheet: *Type* = `deletion/exon_skipping`, *Fusion* = GENE::GENE, *Tags* = event name, and the *Skip fraction* column is filled in.

	A	B	C	D	E	F	G	H	I	J	K
1	**Fusion**	**Confidence**	**Type**	**VICC nomenclature**	**HGVS (RNA)**	**Tags**	**Supporting reads total**	**Coverage 5'**	**Coverage 3'**	**Skip fraction**	**Caller support**
2	MET::MET	**high**	deletion/exon_skipping	NM_000245.4(MET):e.13::NM_000245.4(MET):e.15	NM_000245.4:r.2888_3028del	METex14	934	66	13	95.9%	

Cut-out of the Fusions sheet: MET ex14 skipping, 934 reads across the exon 13 → exon 15 junction, skip fraction 95.9%.

4.2 Technical sheet

The **Technical** sheet contains **all technical detection columns** for the same rows in the **same order** as the Fusions sheet. For easy orientation it starts with *Fusion*, *Confidence*, *Breakpoint 5′* and *Breakpoint 3′*. The other columns (e.g. sequence around the junction, filters, read identifiers, nomenclature status and the reasons why it could not be determined) are mainly intended for bioinformaticians and technical support.

Very long lists of read identifiers are truncated in Excel (Excel cell limit) with the note *truncated; full value in TSV*. The full value is in the TSV file.

	A	B	C	D	E
1	**Fusion**	**Confidence**	**Breakpoint 5'**	**Breakpoint 3'**	**… technical detection columns …**
2	EML4::ALK	**high**	chr2:42295516	chr2:29223528	…

4.3 Info sheet

The **Info** sheet contains the report metadata followed by the **legend**: a *Sheet – Column – TSV column – Description* table covering all columns of the Fusions and Technical sheets. The descriptions are the same as the header notes.

1	RNA fusion report	
2	**Sample**	SAMPLE_03
3	**Project ID**	9405
4	**Examination ID**	109207
5	**Created**	2026-10-25
6	**Genovesa workflow**	Genovesa RNA FusionPlus v1
7	**Annotation version**	1.7.0
8	**Reference assembly**	GRCh38
9	**Reference bundle sha256**	499a34fe4312e447…
10	**Rows (high / medium / low)**	104 (3 / 8 / 93)
11	**Row order**	Confidence high -> medium -> low, then … order; exon skipping after fusions. Same order as the TSV and the PDF pages.

Cut-out of the Info sheet (abbreviated; reference database checksum shortened).

Row	Meaning
Sample, Project ID, Examination ID	Identification of the sample and examination in Genovesa
Created	Report creation date
Genovesa workflow	Workflow name and version
Annotation version	Version of the fusion annotation (nomenclature, exon skipping, Excel)
Reference assembly, Reference bundle sha256	Reference genome and checksum of the RefSeq transcript database used
Rows (high / medium / low)	Total number of rows and by confidence
Row order	Ordering rule (same as TSV and PDF)

5. Project summary Excel output analysis

The file **`PROJECT_<ID>_fusions_aggregated.xlsx`** (Global Result Files) combines the results of all samples in the project. It is useful for an **overview of the whole run**, comparing samples and quickly finding a particular fusion across the project (e.g. by filtering on a gene or confidence).

Sheet	Content
All fusions	All rows of all samples one below the other. Starts with the **Sample** and **Examination ID** columns, followed by the same columns as the examination Fusions sheet.
<sample> - <examination ID>	One sheet per sample, identical to the Fusions sheet of its Excel. A sample without fusions has an empty sheet.
Technical	Technical columns of all samples (with Sample and Examination ID), with a filter.
Info	Project, date, workflow and version, reference, number of samples, row counts by confidence and a sample table.
Legend	Description of all columns on all sheets.

	Sample	Examination ID	Sheet	Rows	High	Medium	Low	Annotation version
2	SAMPLE_01	109205	SAMPLE_01 - 109205	44	1	1	42	1.7.0
3	SAMPLE_02	109206	SAMPLE_02 - 109206	27	2	6	19	1.7.0
4	SAMPLE_03	109207	SAMPLE_03 - 109207	104	3	8	93	1.7.0
5	SAMPLE_04	109208	SAMPLE_04 - 109208	116	1	12	103	1.7.0
6	SAMPLE_05	109209	SAMPLE_05 - 109209	132	3	17	112	1.7.0

Sample table on the Info sheet of the project summary Excel (first five samples).

Most rows of the summary Excel are usually *low* findings. This is expected: they are mostly weakly supported candidates that are meant for targeted review, not for reporting across the board (chapter 7).

6. PDF report

The file `<sample>_fusions.pdf` has **one page for every row** of the Fusions sheet, in the **same order** (high → medium → low). The PDF opens with the **bookmarks** panel: one bookmark per page in the form `N. FUSION – CONFIDENCE`, with `(exon skipping)` appended for exon skipping, e.g. `2. EML4::ALK – HIGH`, `1. MET::MET – HIGH (exon skipping)`.

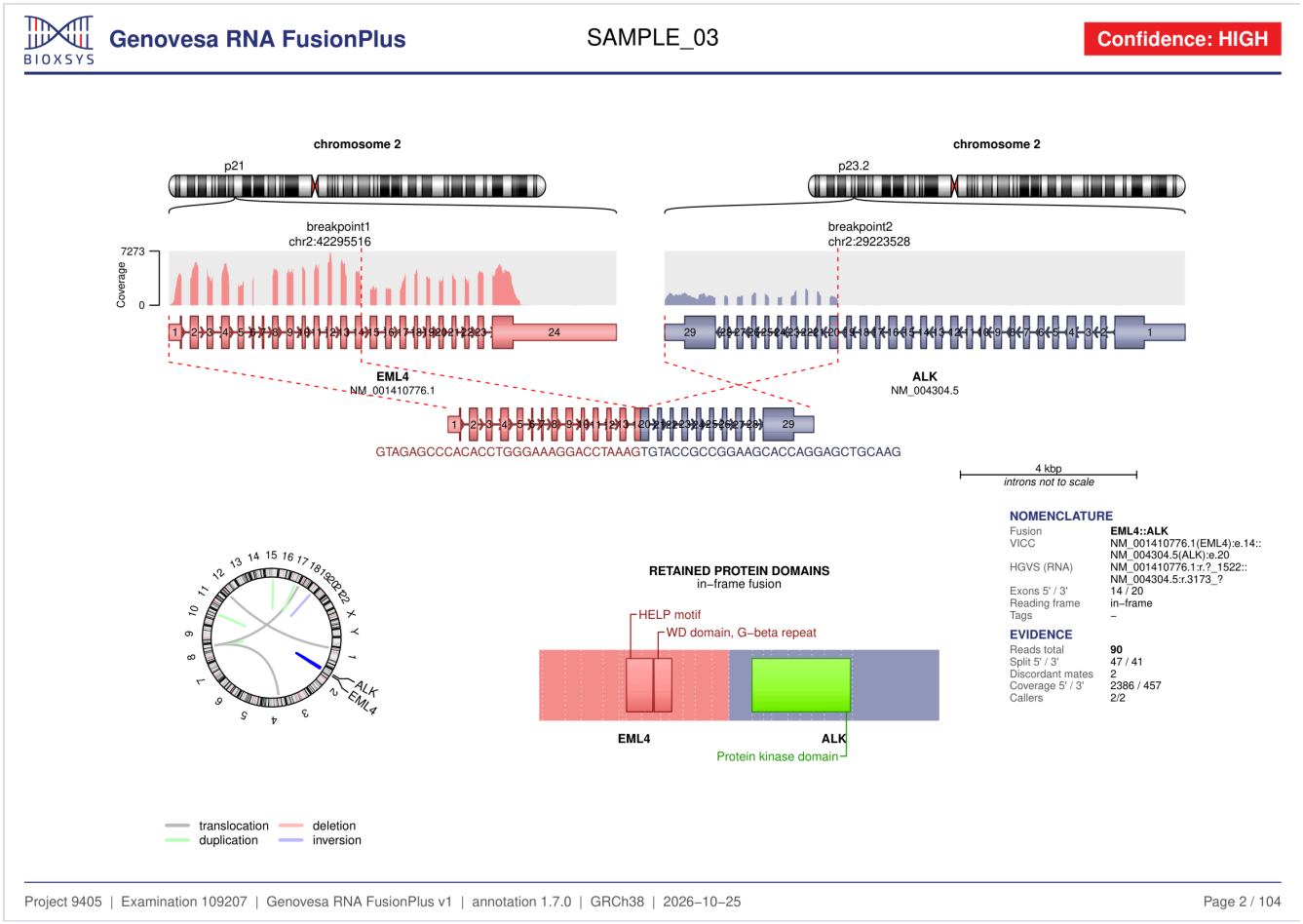


Figure 7. PDF page of the EML4::ALK fusion.

6.1 Parts of the page

Part	What it shows
Header	Logo, workflow name, sample, colour-coded label **Confidence: HIGH** / **MEDIUM** / **LOW**. For exon skipping also the text *exon skipping, skip fraction N %*.
Chromosome ideograms	Chromosome and cytogenetic band of both partners; red mark = breakpoint position.
Coverage	RNA read coverage across the exons of both genes (5′ partner red, 3′ partner blue). Vertical dashed line = breakpoint. For the 3′ partner, coverage is often present only beyond the breakpoint (for ALK exons 20–29), showing that it is specifically the fused part of the gene that is expressed.
Gene models	Exons of both partners with numbers (according to the transcript under the gene name), arrows = direction of transcription. Introns are shortened (*introns not to scale*).
Fusion transcript	Composite transcript: exons of the 5′ partner up to the junction, then exons of the 3′ partner. Below it the sequence around the junction (coloured by partner).

Part	What it shows
Circos plot	Genome overview: the current fusion is highlighted, other high- and medium-confidence fusions of the sample are shown in lighter colours. Colour by type (translocation, duplication, deletion, inversion).
Retained protein domains	Fusion protein diagram: domains retained from the 5′ and 3′ partner; the reading frame is shown under the title (e.g. *in-frame fusion*).
NOMENCLATURE	Fusion, VICC, HGVS (RNA), Exons 5′ / 3′, Reading frame, Tags; same values as in the Excel.
EVIDENCE	Read counts and other support for the finding (table below).
Footer	Project, Examination, workflow and version, annotation version, assembly, PDF creation date, *Page N / M*.

6.2 EVIDENCE panel

Item	Meaning	How to read it (EML4::ALK)
Reads total	Total number of supporting fragments (duplicates excluded) = *Supporting reads total*	`90` – strong support
Split 5′ / 3′	Reads spanning the junction, anchored in the 5′ / 3′ partner	`47 / 41` – junction supported from both sides
Discordant mates	Read pairs split between the two partners	2
Coverage 5′ / 3′	Coverage around the breakpoint (duplicates included), approximate partner expression	`2386 / 457`
Skip fraction	Exon skipping only: fraction of transcripts lacking the exon	– (MET ex14: `95.9 %`)
Callers	How many independent callers reported the same breakpoints (*k*/*n*); “–” for exon skipping	`2/2` – confirmed by both

6.3 Exon skipping page

An exon skipping page has the same layout as a fusion page. Both “partners” are the same gene. The fusion transcript diagram shows the exon preceding the skipped exon joined directly to the exon following it (for MET exon 13 → exon 15). The header shows the skip fraction, the NOMENCLATURE panel the HGVS deletion of the skipped exons and *Tags* the event name.

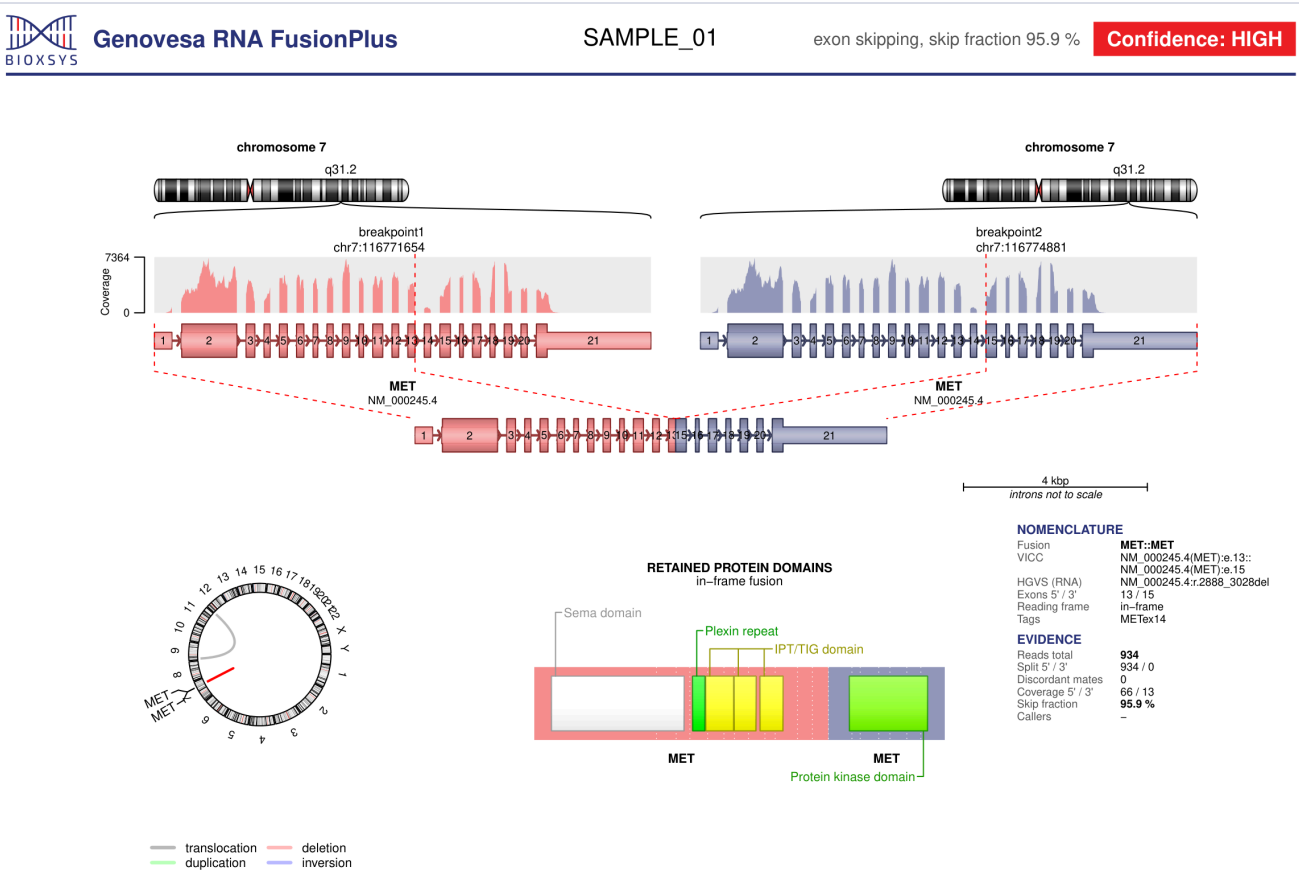


Figure 8. PDF page of MET exon 14 skipping: 934 reads across the exon 13 → 15 junction, skip fraction 95.9%, in-frame, kinase domain retained.

7. How to interpret the results

7.1 Confidence

Level	Meaning	Recommended approach
high	Strong, consistent support of the junction (enough split reads and/or discordant pairs, no signs of a typical artefact).	Assess as a likely genuine finding; consider the clinical context.
medium	Weaker support or circumstances that reduce certainty.	Assess individually (genes, type, frame, domains, caller support).
low	Few reads or features typical of artefacts (e.g. read-through, low expression, ambiguous mapping).	Mostly background. Review in a targeted way (known clinically relevant genes); confirm with another method if suspicious.

For exon skipping, confidence is based on the number of reads across the junction and the skip fraction: **high** ≥ 20 reads and ≥ 10 %, **medium** ≥ 5 reads and ≥ 5 %, **low** ≥ 2 reads. Events with fewer than 2 reads are not reported.

7.2 Caller support (1/2 vs 2/2)

- **2/2** – the fusion was independently reported by both callers with the same breakpoints (10 bp tolerance). This increases confidence that the junction is not an artefact of a single algorithm.
- **1/2** – the fusion was reported by only one caller. This does not mean the fusion is false: the other caller may have different filters or sensitivity (typically with low read counts). **It is a reason for caution, not for dismissing the finding.**
- The second caller always runs; *1/2* therefore never means that the second caller did not run.
- No rows are removed on the basis of caller support.

7.3 Read counts and expression

- Most important are **Supporting reads total** and especially **split reads** (direct evidence of the junction). Support from both sides of the junction (split 5′ and 3′) is stronger than from one side.
- **Coverage** describes the expression of the partners around the breakpoint, not the fusion itself. A weakly expressed gene can carry a relevant fusion with few reads.
- Coverage includes duplicate reads, fusion read counts do not. **Do not calculate “fusion reads / coverage” as an allelic fraction.**
- Absolute read counts depend on sequencing depth and RNA quality; take the examination QC into account when comparing samples.

7.4 Reading frame and domains

- **in-frame** – the fusion protein preserves the reading frame of the 3′ partner. For kinase fusions (ALK, ROS1, RET, NTRK1–3, FGFR…) a functional protein is expected to have an in-frame junction with the kinase domain retained (*Retained protein domains*).
- **out-of-frame** – the frame of the 3′ partner is shifted; a functional fusion protein is less likely. The finding may still matter in other ways (e.g. gene inactivation or promoter swapping); assess in context.
- **stop-codon** / empty – stop codon close to the junction / frame cannot be determined (e.g. breakpoint in a UTR or intron).

7.5 Nomenclature

- **VICC** is given when the exon boundaries are verified in the RefSeq model; **HGVS** only when the junction sequence is verified. Missing HGVS or VICC **does not mean the fusion is false** – only that the exact description could not be determined automatically (the reason is on the Technical sheet).
- A junction **inside an exon** (`e.N–k`, `e.N+k`) is less common and more often an artefact; assess it with extra caution.
- HGVS describes only the junction (outer bounds ?), not the full fusion transcript or the DNA breakpoint.

7.6 Exon skipping and skip fraction

- **Skip fraction** expresses what fraction of the gene's transcripts at that position skip the exon. The higher it is, the more pronounced the aberration.
- **MET ex14**: positive samples usually have a high skip fraction (38–98.5 % in an internal evaluation of 555 samples), background at most a few percent (≤ 2.8 %).
- **ERBB2 Δ16** is also a natural isoform: a low skip fraction occurs in many samples and corresponds to background. ERBB2 Δ16 therefore appears in most samples, typically as *low* with a skip fraction of a few percent.
- Exon skipping is reported **from 2 reads** so that no finding stays hidden. **The clinical significance is decided by the clinician** based on skip fraction, read count, gene expression and context.

Recommendations

- **Confirm** potentially clinically relevant **low** findings and findings with caller support **1/2 with another method** (e.g. RT-PCR, FISH, IHC or another validated method used by your laboratory).
- Check the examination QC (read count, mapped reads, rRNA, coding bases) before drawing a clinical conclusion.
- When comparing with the literature, check the transcript and exon numbering (chapter 4.1).

8. Limitations

- **Sensitivity depends on expression and RNA quality.** A fusion of a weakly expressed gene, or a fusion in a sample with degraded RNA or a low tumour cell content, may not be detected.
- **Only what is in the data can be detected.** A negative result does not mean the fusion is absent from the tumour; consider an independent (orthogonal) method if there is clinical suspicion.
- Fusions with very short read overhangs, with an inserted sequence (linker) or with an intronic breakpoint may not receive automatic nomenclature.
- Exon skipping is evaluated only for the listed events (MET ex14, EGFR vIII/vIIIb/vII/vIIb/vIVa/vIVb, ERBB2 Δ16).
- The annotation is based on the NCBI RefSeq RS_2025_08 reference database (GRCh38); gene names and exon numbers refer to the stated transcript versions.
- The results are supporting analytical data. Clinical interpretation and classification are the responsibility of the laboratory according to its validated procedures.

9. Coming soon

We are working on displaying and reviewing fusions **directly in the Genovesa application**:

- clear **visualisation** of fusions and exon skipping for each examination,
- **comments** and **classification** of fusions,
- **sign-out** of fusions including their inclusion in the **report**,
- display of supporting reads in **IGV**.

We will inform you when the new features become available.

Contact and support

For questions about the results or this guide, please contact Bioxsys technical support. Please quote the project ID, examination ID, workflow name and version and the annotation version (Info sheet or PDF footer).