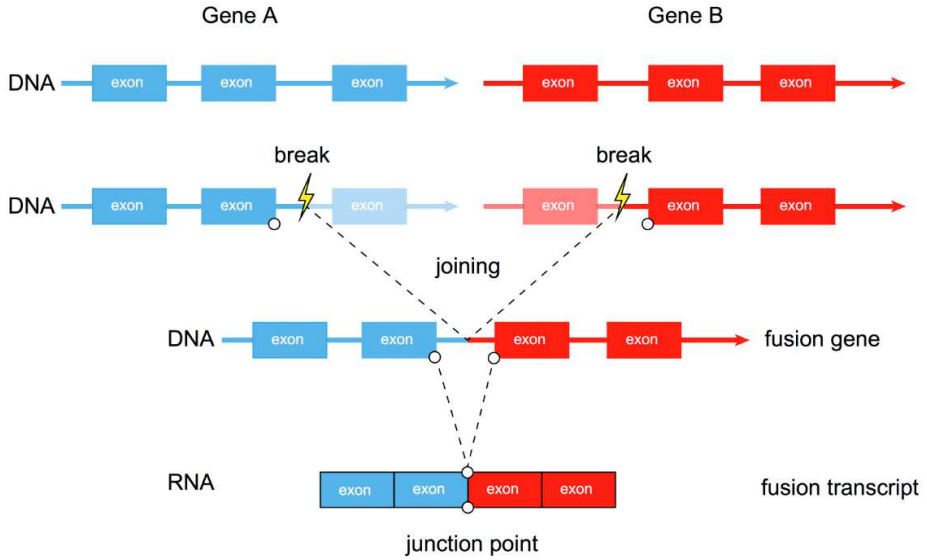
Clinical interpretation for gene fusions

Valentina Nardi, MD

Associate Professor of Pathology, Harvard Medical School
Assistant pathologist, Massachusetts General Hospital

No personal disclosures/conflicts of interest

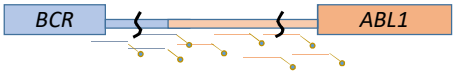
Detection of gene rearrangements by NGS off DNA or RNA



https://www.tumorfusions.org/intro.html

Choice of targeted DNA vs RNA based NGS assay for gene fusion detection

DNA-based



+: finds exact breakpoint

+: can detect "promoter swap" type rearrangements (i.e. IgH rearrangements)

-: needs a lot of baits and sequencing capacity to cover large intronic regions

-: alignment can be tricky and may miss the fusion or not find the partner

-: if small amount of tumor or of DNA the sensitivity for fusion detection is lost

RNA-based

+: detects expressed fusions

+: can work on degraded RNA

(many more copies of expressed fusion than of DNA/cell)

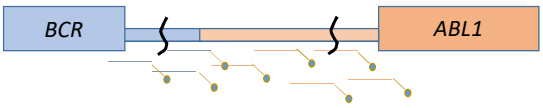
+: detects fusion partner

-: cannot detect promoter swap type rearrangements (i.e. IgH rearrangements)

-: if RNA very degraded the assay will be insensitive

Detection of gene rearrangements by NGS off DNA

When using DNA, a hybrid-capture based approach is generally required in order to cover large intronic regions



Sensitivity of detection depends on:

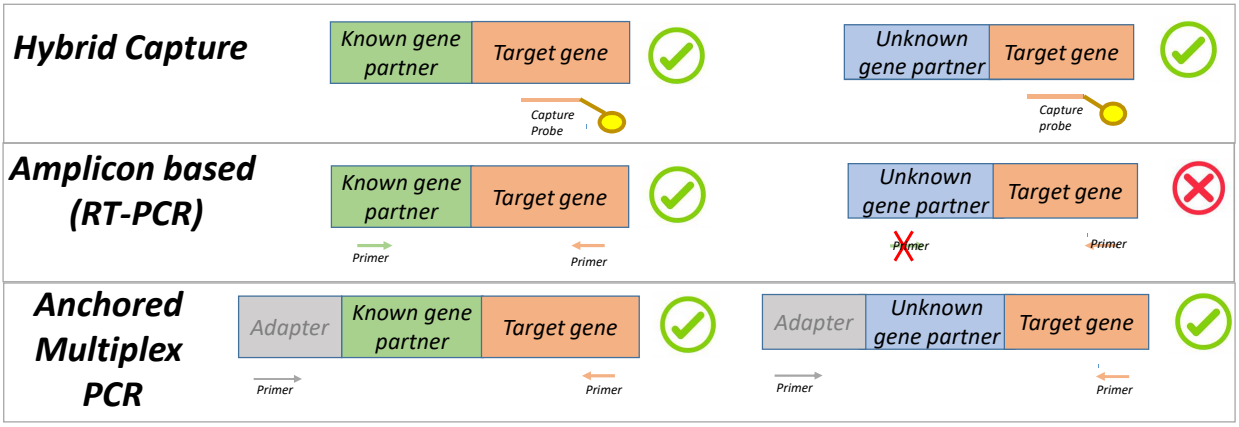
- *Tumor content*
- *Sequencing coverage*
- *Quality of intronic coverage*
  - *Repetitive elements*
  - *Low GC content*
  - *Bait design challenges and stringency*
- *Informatics pipeline*

Detection of rare events (more noise for CNV calls)

(less noise for CNV calls)

Detection of gene rearrangements by NGS off RNA

When using RNA, besides transcriptome sequencing, one can use targeted approaches



Sensitivity of detection depends on:

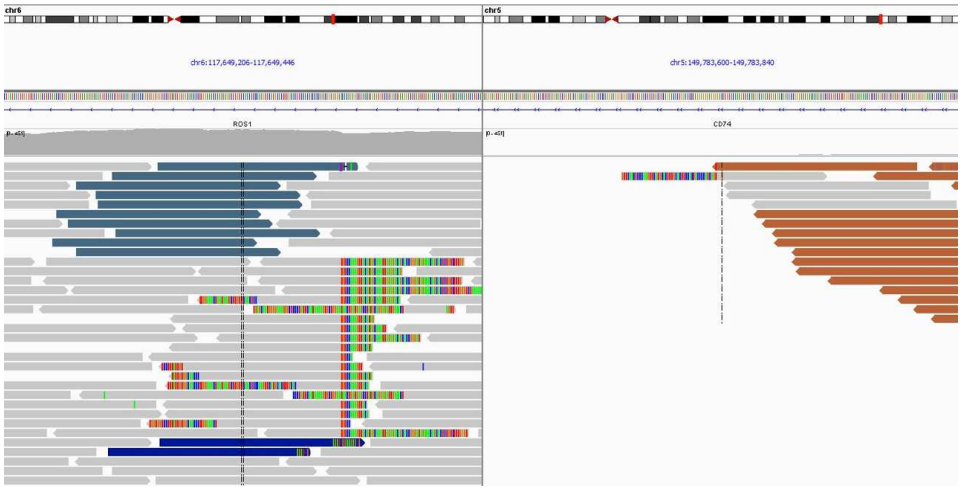
- *Tumor content*
- *Quality of RNA*
- *Sequencing coverage*
- *Expression level of wild type transcripts*
- *Repetitive regions affecting mapping quality*
- *Coverage of pertinent exons*

Analysis of gene rearrangements by NGS

1. *Mapping sequence fragments to a reference (pair-end reads that overlap or span the fusion junction)*
2. *Gene fusion detection tools (filtering steps to reduce false positive calls)*
3. *If fusions are detected, visual inspection and interpretation (novelty, biological significance)*

Example of a fusion detected off DNA

Chromosome 6 *Chromosome 5*

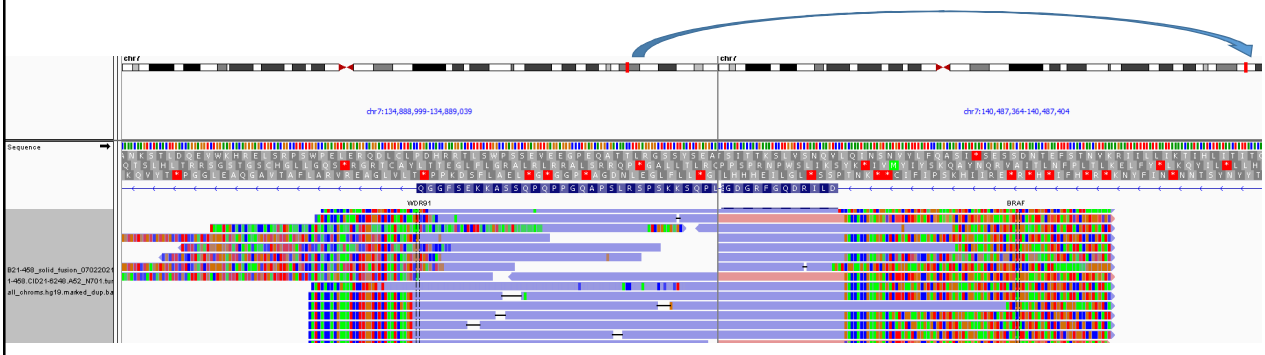


ROS1 intron 32

CD74

Example of an intrachromosomal fusion detected off RNA

Chromosome 7q *Chromosome 7q*



WDR91 exon 6

BRAF exon 9

ATTENTION ATTENTION ATTENTION ATTENTION ATTENTION

False positives! Not all chimeric transcripts are real or pathogenic!

A fusion transcript likely significant/pathogenic

- *It is not a readthrough (the genes are not next to each other i.e. HALC1-COLQ)*

Fusion transcripts more difficult to review, understand frame or Biological significance

- *Inversions/three-way fusions*

Information provided by the variant call file for fusion detection off RNA

 # Reads supporting the fusion

(100s = likely real/driver; single digits, teens= likely artifact or subclonal)

Genes, transcripts IDs, exons involved

(i.e. BCR exon 14 (ENST00000305877), ABL1 2 (ENST00000318560))

Distance from intron-exon boundaries

(i.e. left gene boundary: 0; right gene boundary: 0)

How many bases cross the junction between the 2 genes

(i.e. 5 bases= non specific alignment; vs. 30 bases=confident)

Information provided by the variant call file for fusion detection off RNA

Ideally, also:

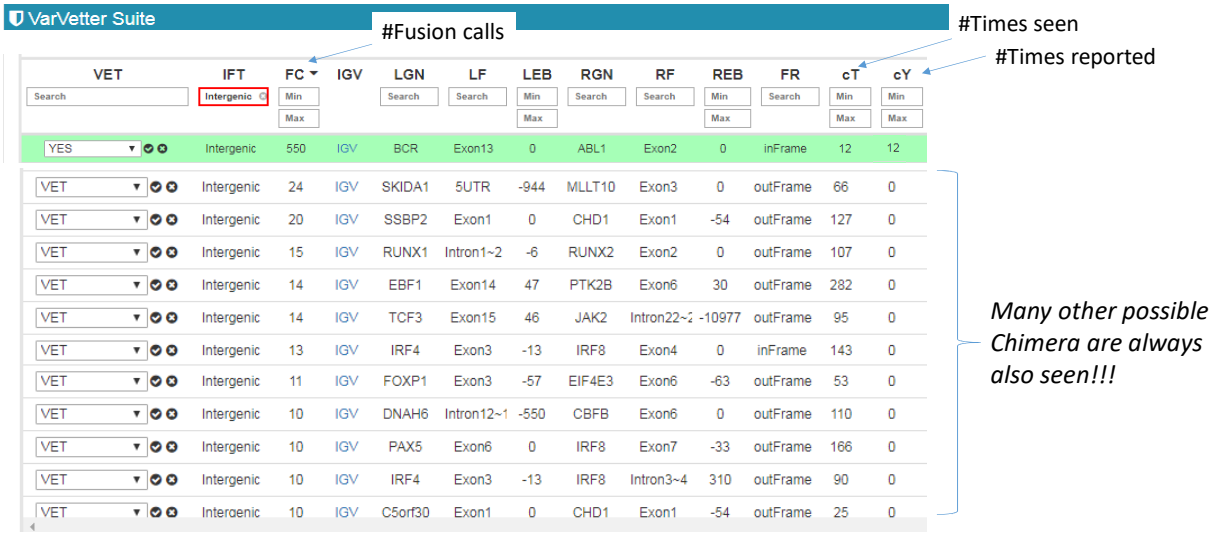
 # Times seen (in the laboratory)
(0=novel, pay attention; 500=artifact)

 # Times reported (in the laboratory)
(1/200= error (i.e. first time seen); 12/12=real)

Already reported in fusion databases or not
(if Yes, you can be more confident, particularly if tumor type fits)

Frame status
(i.e. in-frame vs out-of-frame)

Information provided by the variant call file for fusion detection off RNA



VET	IFT	FC	IGV	LGN	LF	LEB	RGN	RF	REB	FR	cT	cY
Search	Intergenic	Min / Max		Search	Search	Min / Max	Search	Search	Min / Max	Search	Min / Max	Min / Max
YES	Intergenic	550	IGV	BCR	Exon13	0	ABL1	Exon2	0	inFrame	12	12
VET	Intergenic	24	IGV	SKIDA1	5UTR	-944	MLLT10	Exon3	0	outFrame	66	0
VET	Intergenic	20	IGV	SSBP2	Exon1	0	CHD1	Exon1	-54	outFrame	127	0
VET	Intergenic	15	IGV	RUNX1	Intron1~2	-6	RUNX2	Exon2	0	outFrame	107	0
VET	Intergenic	14	IGV	EBF1	Exon14	47	PTK2B	Exon6	30	outFrame	282	0
VET	Intergenic	14	IGV	TCF3	Exon15	46	JAK2	Intron22~2	-10977	outFrame	95	0
VET	Intergenic	13	IGV	IRF4	Exon3	-13	IRF8	Exon4	0	inFrame	143	0
VET	Intergenic	11	IGV	FOXP1	Exon3	-57	EIF4E3	Exon6	-63	outFrame	53	0
VET	Intergenic	10	IGV	DNAH6	Intron12~1	-550	CBFB	Exon6	0	outFrame	110	0
VET	Intergenic	10	IGV	PAX5	Exon6	0	IRF8	Exon7	-33	outFrame	166	0
VET	Intergenic	10	IGV	IRF4	Exon3	-13	IRF8	Intron3~4	310	outFrame	90	0
VET	Intergenic	10	IGV	C5orf30	Exon1	0	CHD1	Exon1	-54	outFrame	25	0

Manual review of fusion reads to verify frame status, direction..

VET	IFT	FC	IGV	LGN	LF	LEB	RGN	RF	REB	FR	cT	cY
Search	Intergenic	Min / Max		Search	Search	Min / Max	Search	Search	Min / Max	Search	Min / Max	Min / Max
YES	Intergenic	550	IGV	BCR	Exon13	0	ABL1	Exon2	0	inFrame	12	7

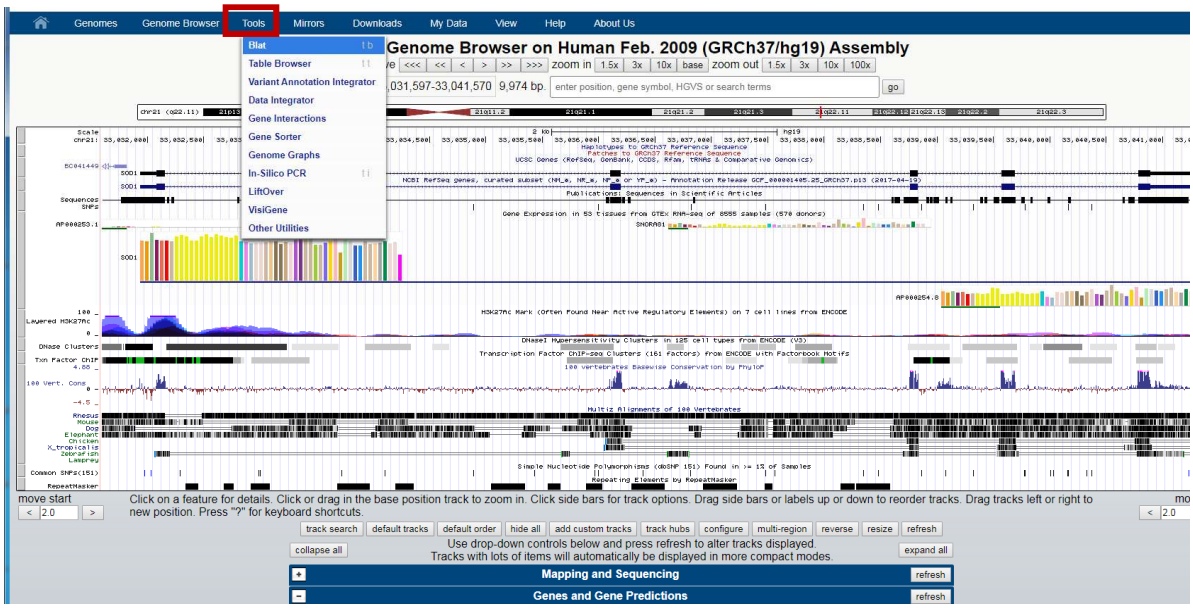
>read_1
TCATTCCGCTGACCATCAATAAGGCAGAAGCCCTTCAGCTGCCAGTAGCATCTGACTTTGAGCCTCAGGGTCTGAGTGAAGCCGCTCGTTGGAA
CTCCAAGGAAAACCTTCTCGCTGGACCCAGTGAAAATGACCCCAACCTTTTCGTTGC

>read_2
GTGAAACTCCAGACTGTCCACAGCATTCCGCTGACCATCAATAAGGAAGAAGCCCTTCAGCGGCCAGTAGCATCTGACTTTGAGCCTCAGGGTC
TGAGTGAAGATCGGAAGAGCACACGTCTGAACTCCAGTCACGCTACGCTATCTCGTA
……..

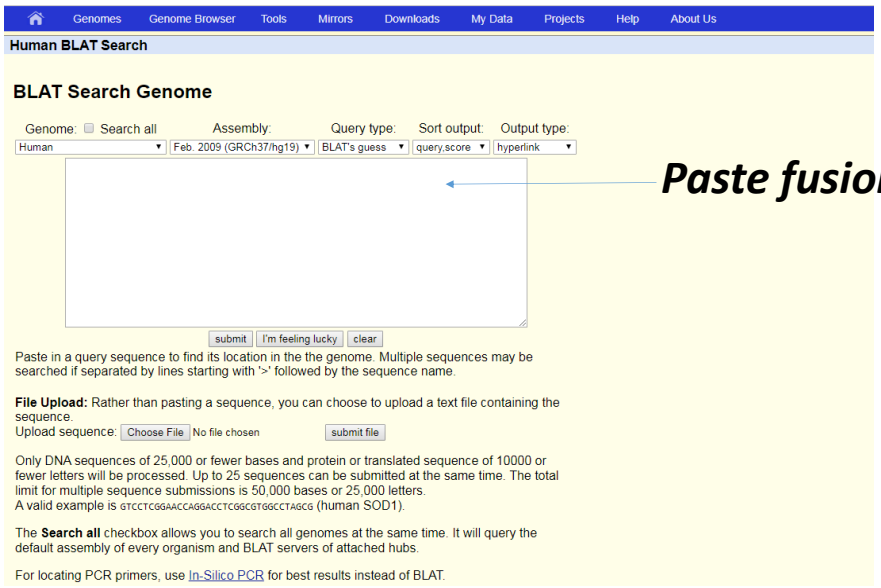
1. ***BLAT*** *the reads with the chimeric transcript and /or*
2. *Use* ***IGV*** *(BAM files) to review the chimeric reads*

Use ***Expasy*** *to ascertain translation*

UCSC genome browser

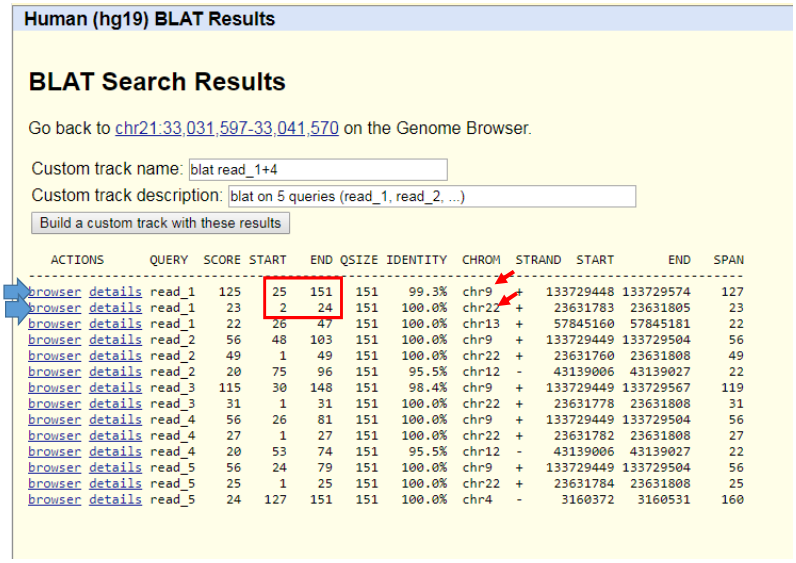


BLAT tool



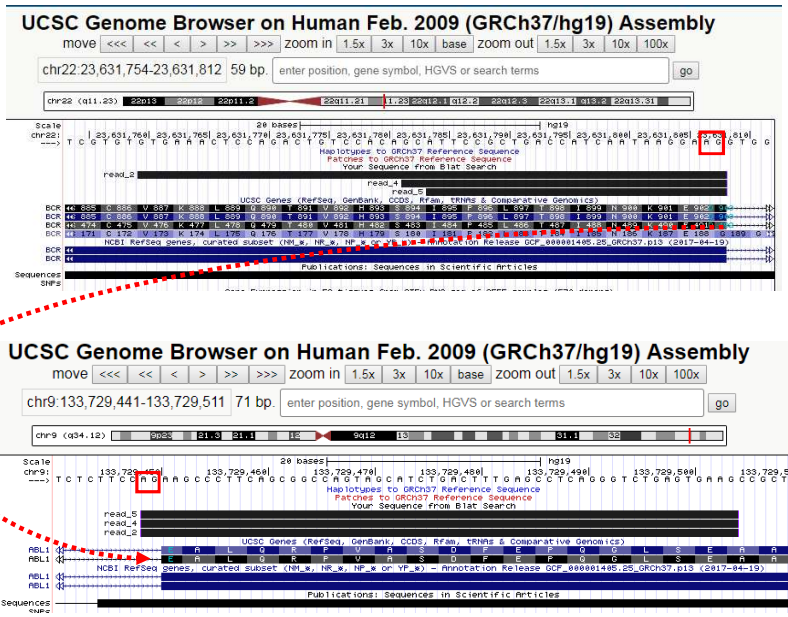
Paste fusion reads here

1. UCSC : BLAT tool to review frame status



ACTIONS	QUERY	SCORE	START	END	QSIZE	IDENTITY	CHROM	STRAND	START	END	SPAN
browser details	read_1	125	25	151	151	99.3%	chr9	+	133729448	133729574	127
browser details	read_1	23	2	24	151	100.0%	chr22	+	23631783	23631805	23
browser details	read_1	22	26	47	151	100.0%	chr13	+	57845160	57845181	22
browser details	read_2	56	48	103	151	100.0%	chr9	+	133729449	133729504	56
browser details	read_2	49	1	49	151	100.0%	chr22	+	23631760	23631808	49
browser details	read_2	20	75	96	151	95.5%	chr12	-	43139006	43139027	22
browser details	read_3	115	30	148	151	98.4%	chr9	+	133729449	133729567	119
browser details	read_3	31	1	31	151	100.0%	chr22	+	23631778	23631808	31
browser details	read_4	56	26	81	151	100.0%	chr9	+	133729449	133729504	56
browser details	read_4	27	1	27	151	100.0%	chr22	+	23631782	23631808	27
browser details	read_4	20	53	74	151	95.5%	chr12	-	43139006	43139027	22
browser details	read_5	56	24	79	151	100.0%	chr9	+	133729449	133729504	56
browser details	read_5	25	1	25	151	100.0%	chr22	+	23631784	23631808	25
browser details	read_5	24	127	151	151	100.0%	chr4	-	3160372	3160531	160

1. UCSC : BLAT tool to review frame status



BCR
Exon 13

Last aa:
…TINKE

ABL1
Exon 2

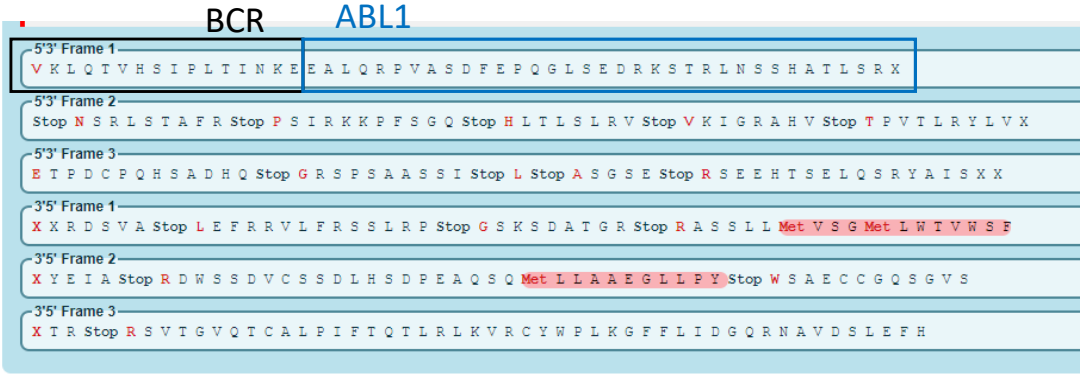
first aa:
…EALQRP

ExPASy translate: tool to review translation of fusion transcript

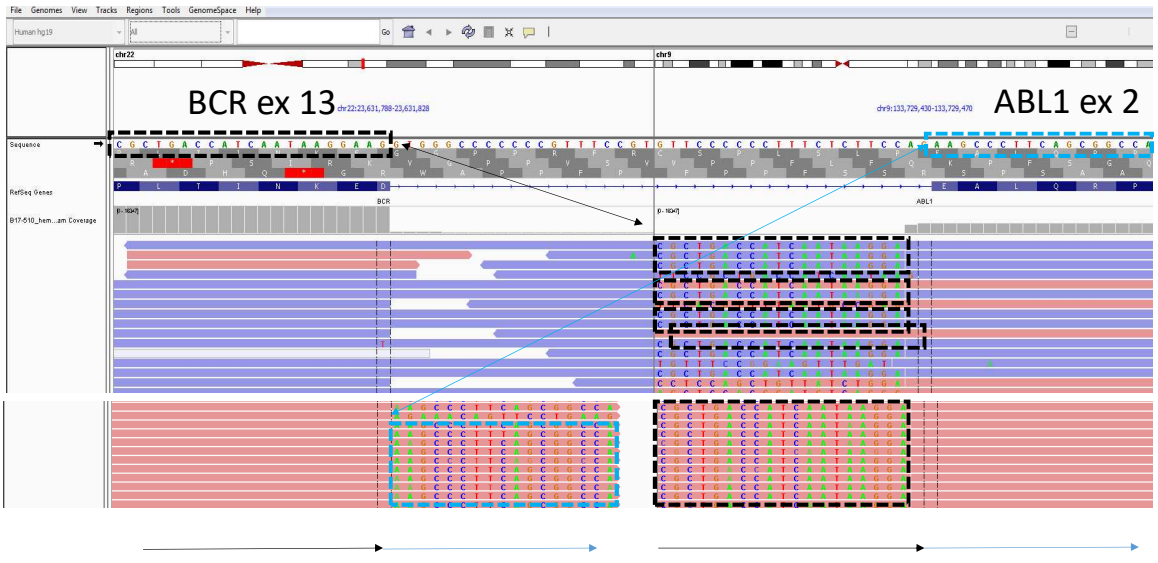
ExPASy translate: tool to review translation of fusion transcript

Last aa: first aa:
…TINKE …EALQRP

BCR ABL1



2. IGV to review fusion reads and direction



Complicated case: 2 fusions? A 3-way fusion? Significance?

AML with complex karyotype that included t(X;8)(p22.1;q22) in addition to the t(8;21)

VET	IFT	FC	IGV	LGN	LF	LEB	RGN	RF	REB	FR	cT	cY	cR	cM	cC	cN	cG	cYG	cA	LCH	LP	LTID	LEID	RCH	RP	RTID
YES	Intergenic	1341	IGV	RUNX1	Exon3	0	RUNX1T1	Exon2	-1	outFrame	9	4	0	0	0	0	0	0	0	21	362317	NM_0010018		8	930295	NM_0011986
YES	Intergenic	309	IGV	RUNX1	Exon3	0	RUNX1T1	Exon2	-32638	outFrame	1	1	0	0	0	0	0	0	0	21	362317	NM_0010018		8	930749	NM_0011986
VET	Intergenic	230	IGV	SKIDA1	5UTR	-1034	MLLT10	Exon2	0	outFrame	116	0	0	0	3	0	0	0	0	10	218135	NM_207371		10	218235	NM_004641
VET	Intergenic	177	IGV	SKIDA1	5UTR	-943	MLLT10	Exon3	0	outFrame	6	0	0	0	2	0	0	0	0	10	218136	NM_207371		10	218277	NM_004641
VET	Intergenic	167	IGV	ETV6	Exon6	-28	SYPL1	Intron1~2	2191	outFrame	0	0	0	0	0	0	0	0	0	12	120374	NM_001987		7	105741	NM_182715
YES	Intergenic	109	IGV	RUNX1	Exon4	-40	RBM10	Exon17	-30	outFrame	1	1	0	0	0	0	0	0	0	21	362068	NM_0010018		X	470416	NM_0012044
VET	Intergenic	92	IGV	RUNX1	Exon3	-47	SYPL1	Intron1~2	2190	outFrame	4	0	0	0	0	0	0	0	0	21	362318	NM_0010018		7	105741	NM_182715
VET	Intergenic	76	IGV	SYPL1	Intron1~2	2163	ETV6	Exon1	0	outFrame	11	0	0	0	0	0	0	0	0	7	105741	NM_182715		12	118030	NM_001987
VET	Intergenic	76	IGV	SYPL1	Intron1~2	2158	MYC	5UTR	101	outFrame	0	0	0	0	0	0	0	0	0	7	105741	NM_182715		8	128748	NM_002467
VET	Intergenic	70	IGV	SYPL1	Intron1~2	2158	MYC	Exon2	0	outFrame	0	0	0	0	0	0	0	0	0	7	105741	NM_182715		8	128750	NM_002467
VET	Intergenic	67	IGV	ETV6	Intron2~3	-9	RNF114	Intron4~5	964	outFrame	0	0	0	0	0	0	0	0	0	12	119920	NM_001987		20	485637	NM_018683
VET	Intergenic	65	IGV	SYPL1	Intron1~2	2163	ETV6	Exon2	6	outFrame	1	0	0	0	0	0	0	0	0	7	105741	NM_182715		12	119053	NM_001987
VET	Intergenic	61	IGV	ETV6	Exon7	-2	CTSK	Intron4~5	-623	outFrame	26	0	0	0	0	0	0	0	0	12	120388	NM_001987		1	150777	NM_000396

RUNX1 ex 3-RUNX1T1 ex 2

RUNX1 ex 4- RBM10 ex17

RUNX1 ex 3-RUNX1T1 ex 2

BLAT Search Results

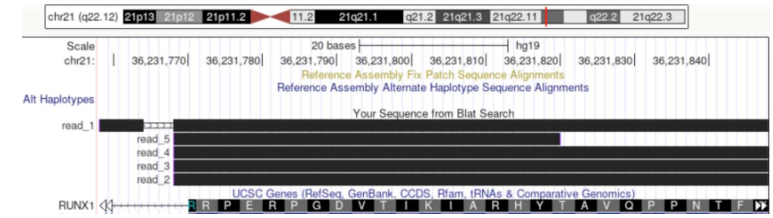
Go back to chr6:151142326-151144807 on the Genome Browser.

Custom track name: blat read_1+4

Custom track description: blat on 5 queries (read_1, read_2, ...)

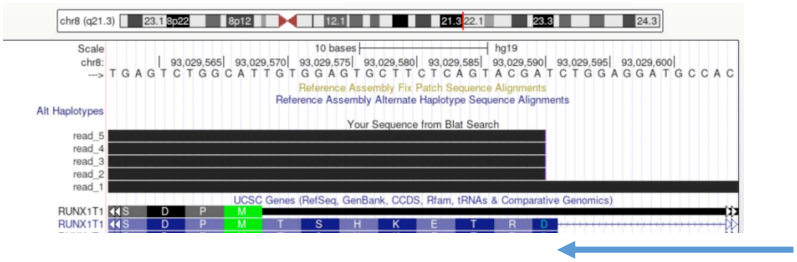
Build a custom track with these results

ACTIONS	QUERY	SCORE	START	END	QSIZE	IDENTITY	CHROM	STRAND	START	END	SPAN
browser details	read_1	84	41	127	151	98.9%	chr21	+	36231759	36231848	90
browser details	read_1	49	1	49	151	100.0%	chr8	+	93029557	93029605	49
browser details	read_2	95	34	128	151	100.0%	chr21	+	36231769	36231863	95
browser details	read_2	34	1	34	151	100.0%	chr8	+	93029557	93029590	34
browser details	read_3	81	34	114	151	100.0%	chr21	+	36231769	36231849	81
browser details	read_3	46	106	151	151	100.0%	chr8	-	93029503	93029548	46
browser details	read_3	34	1	34	151	100.0%	chr8	+	93029557	93029590	34
browser details	read_4	81	34	114	151	100.0%	chr21	+	36231769	36231849	81
browser details	read_4	46	106	151	151	100.0%	chr8	-	93029503	93029548	46
browser details	read_4	34	1	34	151	100.0%	chr8	+	93029557	93029590	34
browser details	read_5	52	34	85	151	100.0%	chr21	+	36231769	36231820	52
browser details	read_5	34	1	34	151	100.0%	chr8	+	93029557	93029590	34
browser details	read_5	20	67	86	151	100.0%	chr6	+	50596397	50596416	20



RUNX1

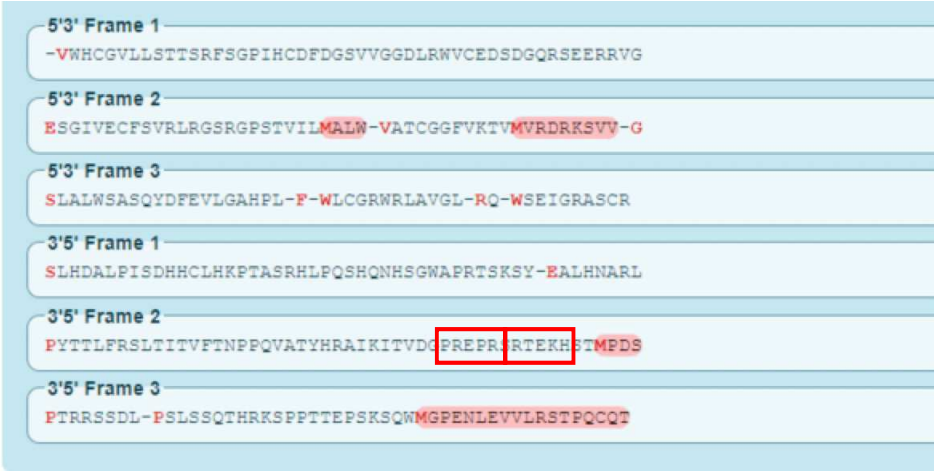
…PREPR



RUNX1T1

…(D)RTEK…

RUNX1-RUNX1T1 is in frame



BLAT Search Results

Go back to chr21:36231759-36231848 on the Genome Browser.

Custom track name: blat read_1+4

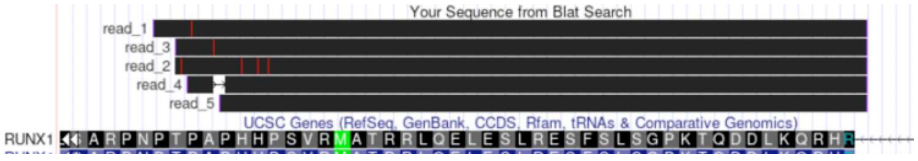
Custom track description: blat on 5 queries (read_1, read_2, ...)

Build a custom track with these results

ACTIONS	QUERY	SCORE	START	END	QSIZE	IDENTITY	CHROM	STRAND	START	END	SPAN
browser details	read_1	128	2	131	151	99.3%	chr21	+	36206771	36206900	130
browser details	read_1	23	129	151	151	100.0%	chrX_kb021648_fix	-	431866	431888	23
browser details	read_1	23	129	151	151	100.0%	chrX	-	47041704	47041726	23
browser details	read_1	21	21	41	151	100.0%	chr1	-	182501970	182501990	21
browser details	read_1	21	13	33	151	100.0%	chr22	+	19958392	19958412	21
browser details	read_2	118	1	126	151	96.9%	chr21	+	36206775	36206900	126
browser details	read_2	28	124	151	151	100.0%	chrX_kb021648_fix	-	431861	431888	28
browser details	read_2	28	124	151	151	100.0%	chrX	-	47041699	47041726	28
browser details	read_2	25	1	28	151	96.5%	chr20	+	47552462	47552492	31
browser details	read_2	25	1	28	151	84.7%	chr1_jh636052_fix	+	2582205	2582230	26
browser details	read_2	23	1	28	151	92.6%	chr5	+	141455982	141456011	30
browser details	read_2	23	1	28	151	92.6%	chr18	+	46606508	46606536	29
browser details	read_2	20	3	22	151	100.0%	chr3	-	138974006	138974025	20
browser details	read_2	20	3	22	151	100.0%	chr22	-	45518000	45518019	20
browser details	read_3	124	6	131	151	99.3%	chr21	+	36206775	36206900	126
browser details	read_3	29	1	33	151	96.8%	chr22	+	19958099	19958412	314
browser details	read_3	23	129	151	151	100.0%	chrX_kb021648_fix	-	431866	431888	23
browser details	read_3	23	129	151	151	100.0%	chrX	-	47041704	47041726	23
browser details	read_3	21	21	41	151	100.0%	chr1	-	182501970	182501990	21
browser details	read_3	20	1	22	151	95.5%	chr5	-	89228110	89228131	22
browser details	read_4	121	2	123	151	100.0%	chr21	+	36206777	36206900	124
browser details	read_4	31	121	151	151	100.0%	chrX_kb021648_fix	-	431858	431888	31
browser details	read_4	31	121	151	151	100.0%	chrX	-	47041696	47041726	31
browser details	read_4	24	1	25	151	100.0%	chr22	+	19958097	19958412	316
browser details	read_4	21	13	33	151	100.0%	chr1	-	182501970	182501990	21
browser details	read_5	118	14	131	151	100.0%	chr21	+	36206783	36206900	118
browser details	read_5	23	129	151	151	100.0%	chrX_kb021648_fix	-	431866	431888	23
browser details	read_5	23	129	151	151	100.0%	chrX	-	47041704	47041726	23
browser details	read_5	23	1	23	151	100.0%	chrX_jh720453_fix	+	1135748	1135770	23
browser details	read_5	23	1	23	151	100.0%	chrX	+	76336747	76336769	23
browser details	read_5	21	21	41	151	100.0%	chr1	-	182501970	182501990	21

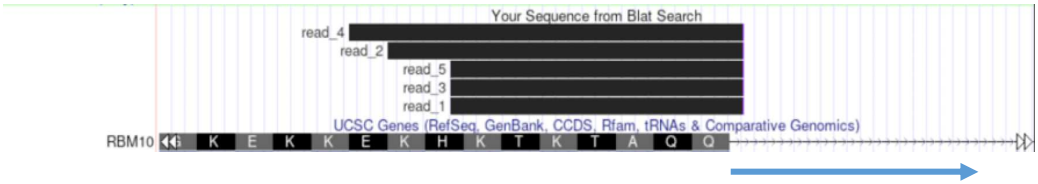
RBM10-RUNX1

RBM10-RUNX1



RUNX1

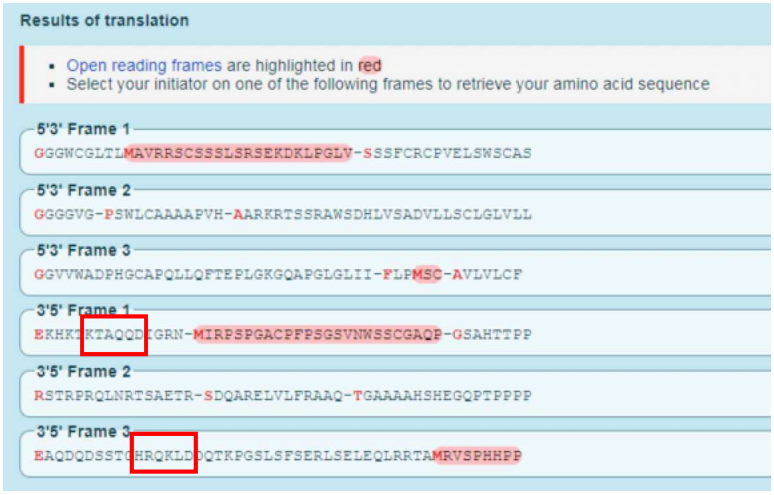
(R)HRQKL..



RBM10

..KTAQQ

RBM10-RUNX1 is not in frame



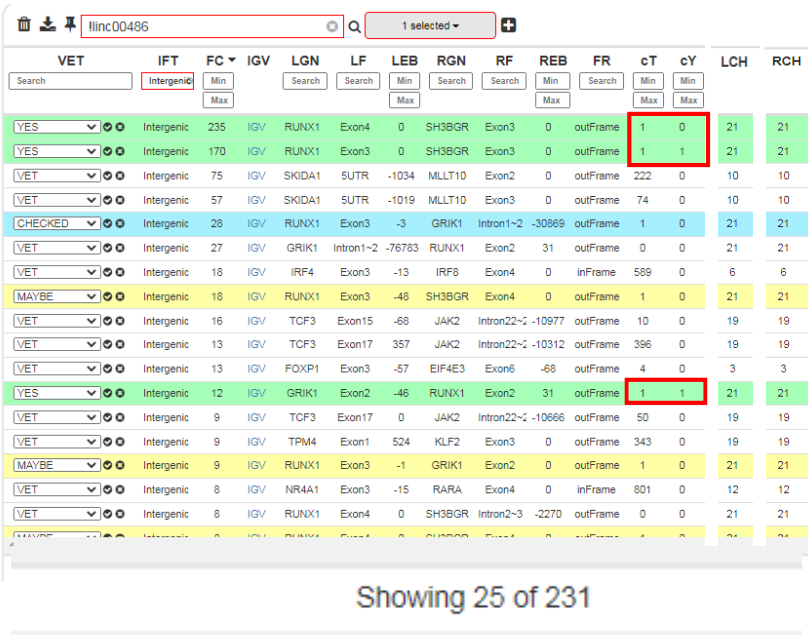
RBM10
..KTAQQ

RUNX1
(R)HRQKLD..

Do you report both?

Complex karyotype that included t(X;8)(p22.1;q22) in addition to the canonical t(8;21)

Another complicated case: 2 fusions? A 3-way fusion?



llinc00486 | 1 selected

VET	IFT	FC	IGV	LGN	LF	LEB	RGN	RF	REB	FR	cT	cY	LCH	RCH
YES	Intergenic	235	IGV	RUNX1	Exon4	0	SH3BGR	Exon3	0	outFrame	1	0	21	21
YES	Intergenic	170	IGV	RUNX1	Exon3	0	SH3BGR	Exon3	0	outFrame	1	1	21	21
VET	Intergenic	75	IGV	SKIDA1	5UTR	-1034	MLLT10	Exon2	0	outFrame	222	0	10	10
VET	Intergenic	57	IGV	SKIDA1	5UTR	-1019	MLLT10	Exon3	0	outFrame	74	0	10	10
CHECKED	Intergenic	28	IGV	RUNX1	Exon3	-3	GRIK1	Intron1~2	-30869	outFrame	1	0	21	21
VET	Intergenic	27	IGV	GRIK1	Intron1~2	-76783	RUNX1	Exon2	31	outFrame	0	0	21	21
VET	Intergenic	18	IGV	IRF4	Exon3	-13	IRF8	Exon4	0	inFrame	589	0	6	6
MAYBE	Intergenic	18	IGV	RUNX1	Exon3	-48	SH3BGR	Exon4	0	outFrame	1	0	21	21
VET	Intergenic	16	IGV	TCF3	Exon15	-68	JAK2	Intron22~2	-10977	outFrame	10	0	19	19
VET	Intergenic	13	IGV	TCF3	Exon17	357	JAK2	Intron22~2	-10312	outFrame	396	0	19	19
VET	Intergenic	13	IGV	FOXP1	Exon3	-57	EIF4E3	Exon6	-68	outFrame	4	0	3	3
YES	Intergenic	12	IGV	GRIK1	Exon2	-46	RUNX1	Exon2	31	outFrame	1	1	21	21
VET	Intergenic	9	IGV	TCF3	Exon17	0	JAK2	Intron22~2	-10666	outFrame	50	0	19	19
VET	Intergenic	9	IGV	TPM4	Exon1	524	KLF2	Exon3	0	outFrame	343	0	19	19
MAYBE	Intergenic	9	IGV	RUNX1	Exon3	-1	GRIK1	Exon2	0	outFrame	1	0	21	21
VET	Intergenic	8	IGV	NR4A1	Exon3	-15	RARA	Exon4	0	inFrame	801	0	12	12
VET	Intergenic	8	IGV	RUNX1	Exon4	0	SH3BGR	Intron2~3	-2270	outFrame	0	0	21	21

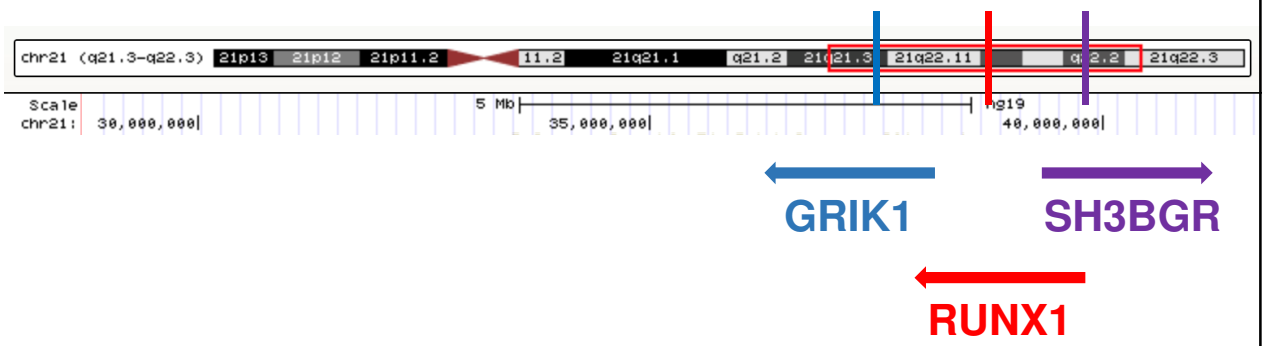
Showing 25 of 231

RUNX1 exon3
SH3BGR exon 3

GRIK1 exon2
RUNX1 exon 2

Complicated case: 2 fusions? A 3-way fusion?

RUNX1, _GRIK1_, and _SH2BGR_ are located all on Chr21q21-q22

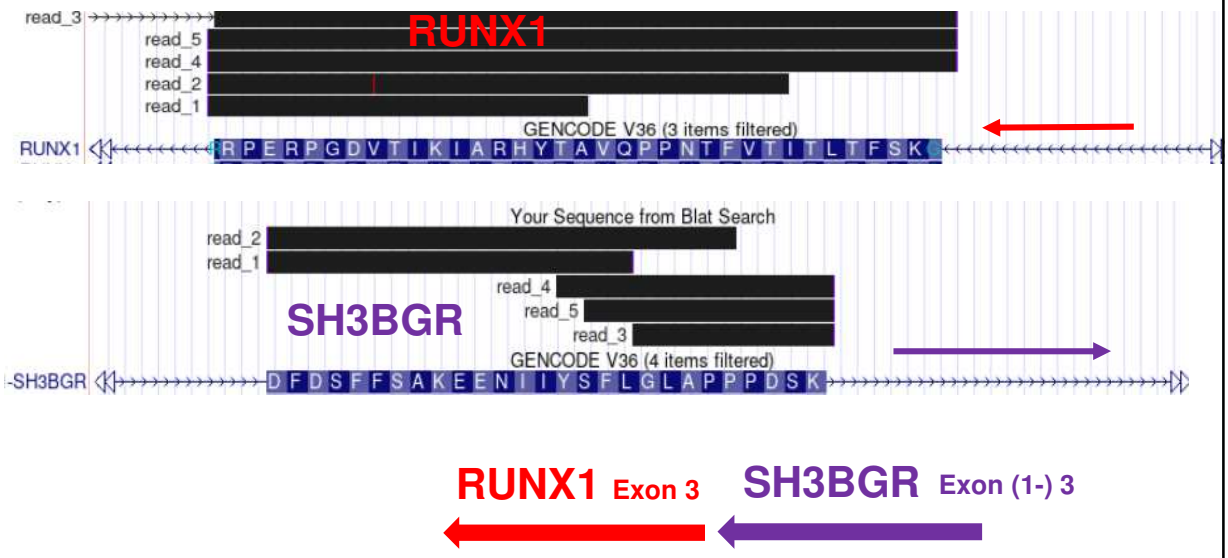


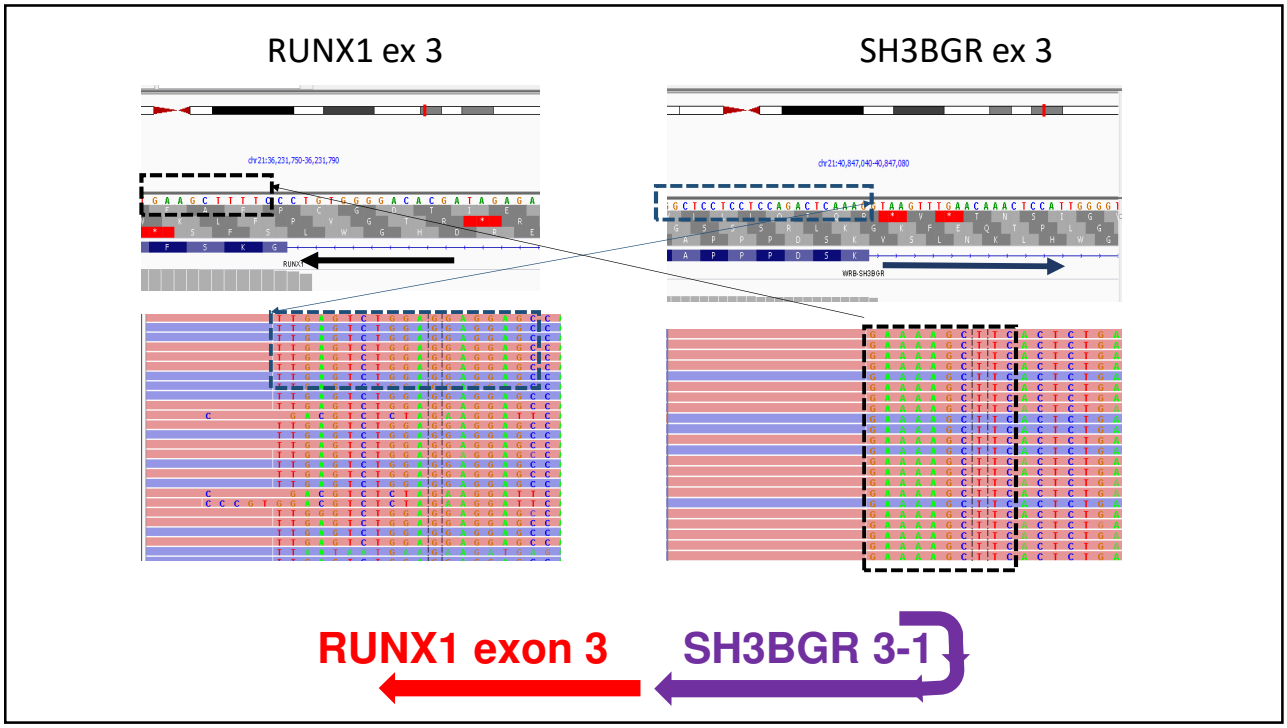
GRIK1 **SH3BGR**

RUNX1

RUNX1- SH3BGR

An inversion leads to SH3BGR exon3-RUNX1 exon 3

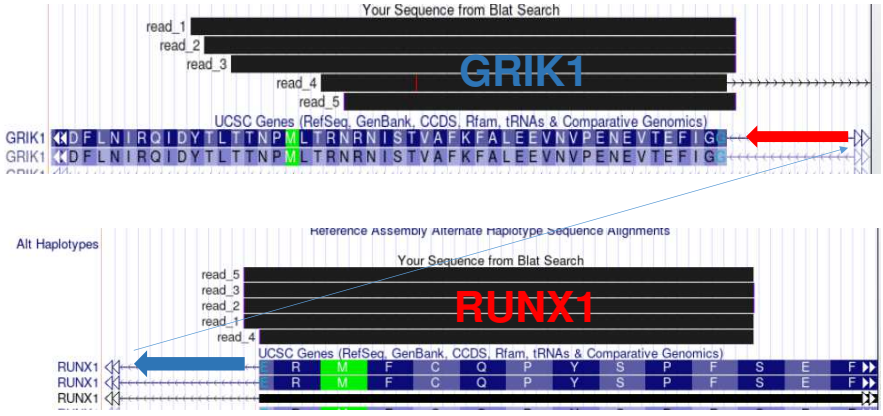




SH3BGR-RUNX1: not in frame or alt frame?

RUNX1-GRIK1

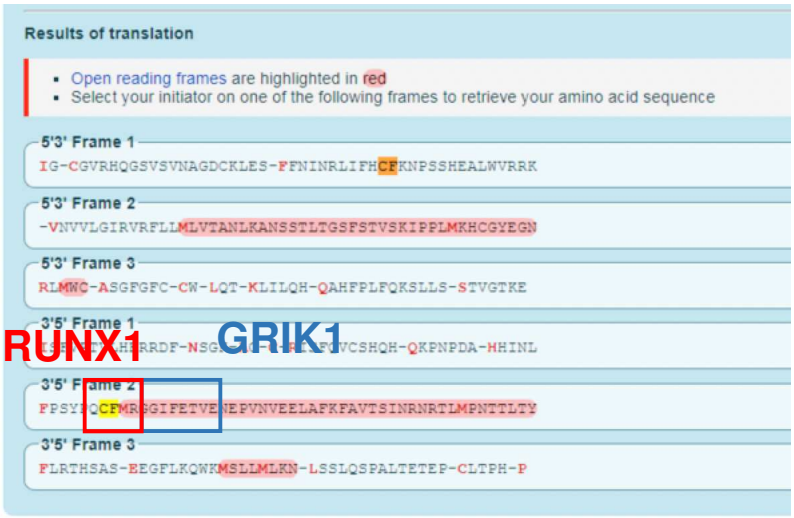
An inversion leads to GRIK1 exon2- RUNX1 exon2

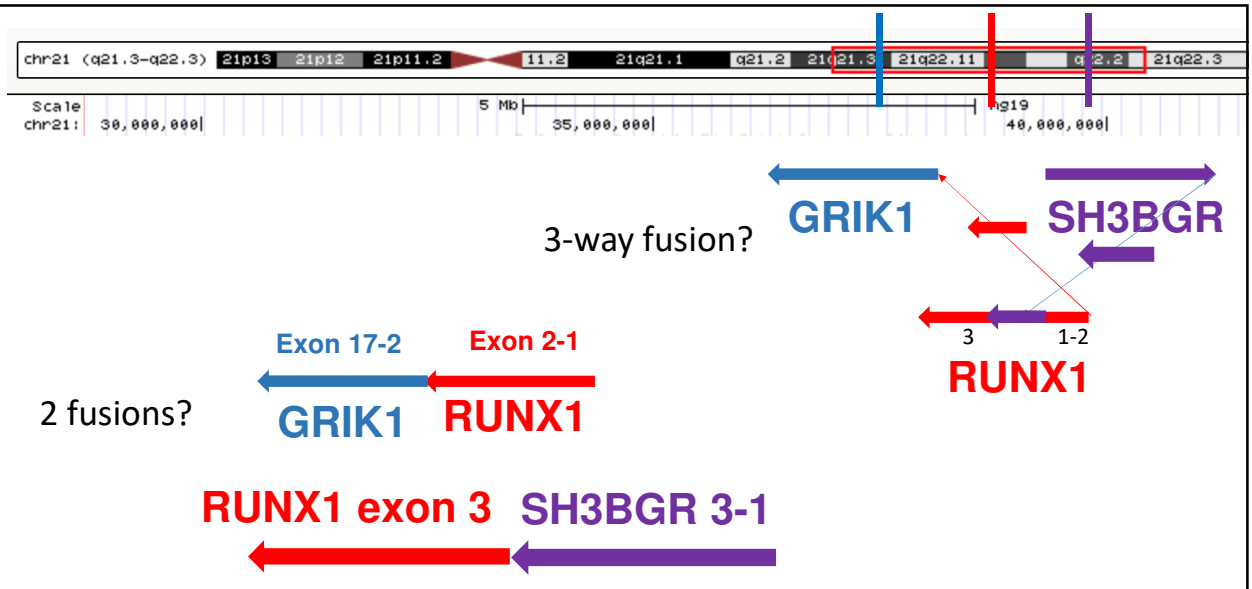


Exon 17-2 Exon 2-1

GRIK1 RUNX1

RUNX1-GRIK1 is in frame





Karyotype: *complex karyotype*
44-46, XY, add(5)(q11.2),-7,+11[5],add(11)(p11.2)[3], der(11)t(8;11)(q?13;p13), der(13;14)(q10;q10),+14,add(14)(p11.2),-16,add(17)(p11.2),+21,ider(21)(q22), +mar1[3],+mar2[2][cp20]

Conclusions

- Gene fusions can be detected off targeted sequencing
- RNA sequencing can detect known and novel fusion partners
- Many chimeric transcripts are usually seen, most of which are artifacts
- Reportable fusions are usually

  - highly expressed
  - not seen in a lot of case
  - in frame
  - biologically significant (i.e. activating kinases)

Thank you!

Clinical interpretation of gene fusions

Valentina Nardi

Question: What are the features that support a reportable gene fusion?

1) Detecting the gene fusion in a high number of reads and specimens
2) Detecting a gene fusion that involves adjacent fusion genes
3) Detecting a gene fusion involving genes with high similarity
4) Detecting an in-frame tyrosine kinase activating gene fusion
5) Detecting a gene fusion involving a gene highly amplified in the specimen sequenced

Answer: 4)

Justification / Reference
Reportable gene fusions will be rather uncommon, seen in a high number of reads, not involving homologous genes, not seen in the setting of a high gene amplification, in frame and often involving and activating kinases.
Ref PMID: 25500544
The landscape and therapeutic relevance of cancer-associated transcript fusions.

Keywords: rearrangement, RNA sequencing, chimera, translocation, chimeric transcripts,

Take Home Summary:
Gene fusion detection off targeted sequencing is effective but requires manual review of the sequenced reads to avoid false positive calls.