Clinker: visualising fusion genes detected in RNA-seq data. Supplementary Figures and Tables.

Breon M Schmidt, Nadia M Davidson, Anthony DK Hawkins,
Ray Bartolo, Ian J Majewski, Paul G Ekert, Alicia Oshlack

Clinker Read Support

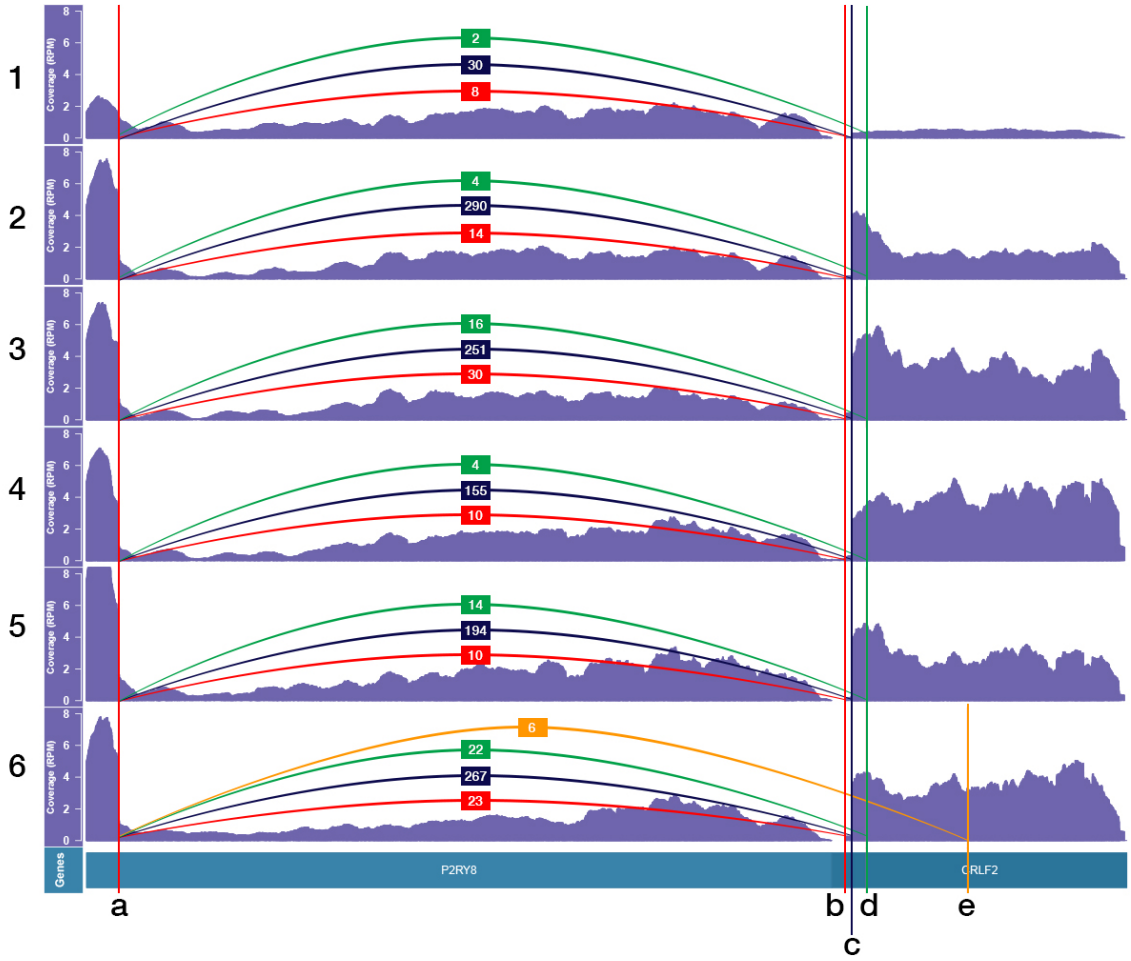

Figure S1. Modified clinker output that demonstrates the coverage and breakpoints of six samples containing the P2RY8-CRLF2 fusion gene.

JAFFA and Clinker spanning read support for the P2RY8-CRLF2 fusions found in the samples above

Sample	Fusion a-b		Fusion a-c		Fusion a-d		Fusion a-e	
	JAFFA	**Clinker**	**JAFFA**	**Clinker**	**JAFFA**	**Clinker**	**JAFFA**	**Clinker**
1	N/A	14	53	290	2	4	N/A	N/A
2	N/A	10	55	155	4	4	N/A	N/A
3	5	10	36	194	10	14	N/A	N/A
4	6	23	57	267	10	22	N/A	6
5	2	8	10	30	3	2	N/A	N/A
6	4	30	51	251	8	16	N/A	N/A

Table S1. Number of spanning reads (reads that split across the breakpoint) discovered by JAFFA (JAFFA filters out breakpoints that are not inframe) and Clinker.

Table S1 indicates that Clinker alignment produces more read support from spanning reads. This is intuitive given that Clinker concatenates the superTranscripts from both genes to form a single reference sequence, which the STAR aligner would more efficiently align to.

Spurious Read Alignments

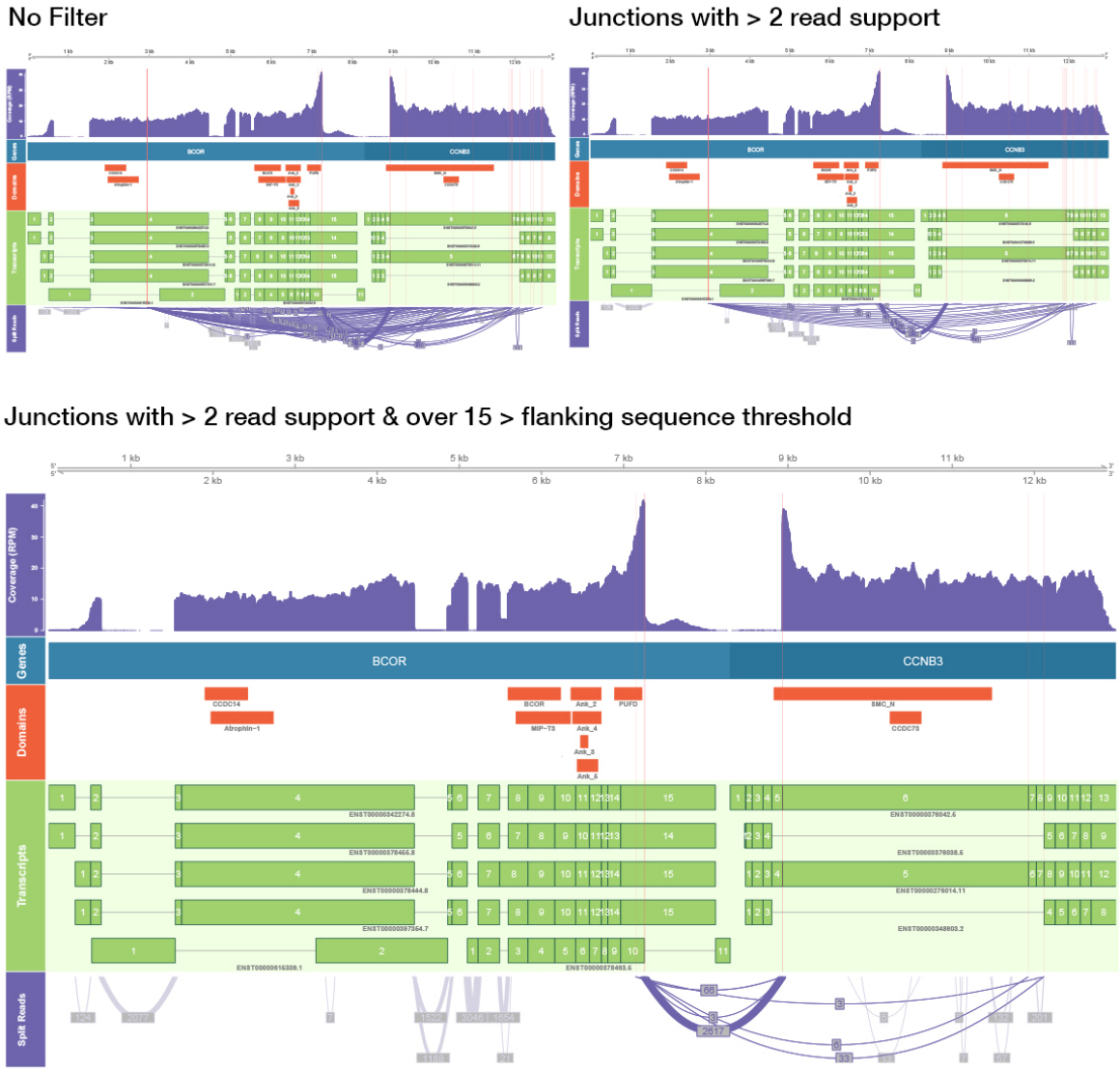

Figure S2. There is substantial reduction of noisy alignments when minimum flanking sequence and minimum junction support thresholds are defined. Minimum flanking sequence is increased to 15 in this case due the large number of split read alignments, which naturally made this a great example of the filtering process.

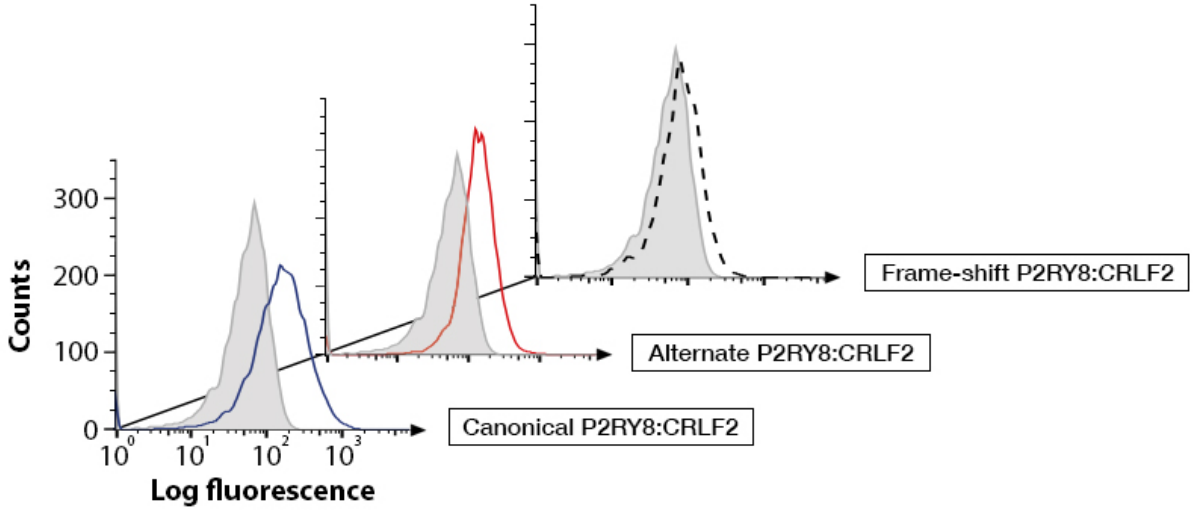

Figure S3. Representative histograms of *CRLF2* expression in BaF3 cells measured by Flow Cytometry. BaF3 cells were transfected with either the canonical *P2RY8-CRLF2* fusion, the alternate in-frame fusion that includes the 1st exon of *P2RY8* and the 5'UTR of *CRLF2* or a third isoform detected in patient samples that results in a frame shift and a premature stop codon. The blue, red and dashed histograms show *CRLF2* expression detected using an anti-*CRLF2* antibody. The gray shaded histogram shows unstained controls. The same unstained control histogram is shown in each case as this data is from a single representative experiment.