

Table S1. Examples of identified BC–ROI combinations.

No.	BCs		BC–ROI combinations			
	Sequence	Read count	BC sequence	ROI sequence	Read count	Proportion of chimeric molecules
1	ATATAGTCATGCCCTAAG ATATAGTCAG <u>CC</u> CCCTAAG	248 3	ATATAGTCATGCCCTAAG	TA ACTGAT	242	1/251 (0.4%)
				TA <u>A</u> CTGCT	2	
			TA <u>A</u> CTGGT	1		
			TA <u>A</u> CGGAT	1		
			T <u>A</u> CTGAT	1		
			ACT <u>G</u> TTGG	1		
			ATATAGTCAG <u>CC</u> CCCTAAG	TA <u>A</u> CTGAT	3	
2	CTAGTATGATACACATCT CTAGTATGATAT <u>A</u> CATCT	99 2	CTAGTATGATACACATCT	TA ACTGAT	96	0/101 (0%)
				TA <u>A</u> CTCAT	2	
			TA <u>A</u> CTGGT	1		
			CTAGTATGATAT <u>A</u> CATCT	TA <u>A</u> CTGAT	2	

Sequences of genuine BCs and ROIs (i.e. the most represented sequences within the groups defined by the algorithm used) are in bold characters. Mutant variants of genuine BCs and ROIs (BCs with up to 2 nucleotide substitutions and ROIs with 1 nucleotide substitution) are in regular font; the substitutions are underlined. A chimeric ROI is in red.