Supplementary Table S1. *Improvement of merged paired-end reads in alignment and RIS identification*

Merged Paired Sequences	#	
Total reads	2489444	
Alignment filters	**#**	**%**
Repeats	74816	3.005
Unique RISs	1643	1.724
Mean bit score	255.532	
Non-Merged Paired Sequences	#	
Total reads	2489444	
Alignment filters	**#**	**%**
Repeats	85145	3.42
Unique RISs	1585	1.446
Mean bit score	225.131	
Unique RIS Z Score Raetio	5.049	
Unique RIS P Value	0.00001	
Repeat Z Score Ratio	-26.25	
Repeat P Value	0.00001	

The table reports the total number of sequence reads, number of unique retroviral integration site (RISs) identified, mean alignment of blastbit score, and the number of reads excluded because of unreliable mapping in a repetitive region of the genome. The number of sequences filtered out because of unreliable mapping in repeat regions was significantly reduced, and the alignment of blastbit score was improved. Finally, the number of unique integrations identified using merged paired-end reads was significantly higher than unmerged reads. Statistics were calculated using the *z*-ratio for the significance of the difference between two independent proportions on a two-tailed probability with a significance of 0.01.