Esquema de funcionamiento de fastaReader

📑 schizophrenia.fasta

↓

seqsRDD - RDD[(línea, número)]

>FJ804216.1 Homo sapiens disrupted in schizophrenia 1 isoform 46 (DISC1) mRNA, complete cds, alternatively spliced	1
GGAAGGAGCAGGAGGCAGCCCAGGCGGAGCGGGAGGAGCTGGCAGCGGGGCGCATGCCAGGCGGGGGTCC	2
TCAGGGCGCCCCAGCCGCCGCCGGCGGCGGCGGCGTGAGCCACCGCGCAGGCAGCCGGGATTGCTTACCA	3
CCTGCAGCGTGCTTTCGGAGGCGGCGGCTGGCACGGAGGCCGGGCTACATGAGAAGCTCGACAGGGCCTG	4
...	

↓

groupedRDD - RDD[(num_línea_cab,(línea, número))]

1	>FJ804216.1 Homo sapiens disrupted in schizophrenia 1 isoform 46 (DISC1) mRNA, complete cds, alternatively spliced	1
1	GGAAGGAGCAGGAGGCAGCCCAGGCGGAGCGGGAGGAGCTGGCAGCGGGGCGCATGCCAGGCGGGGGTCC	2
1	TCAGGGCGCCCCAGCCGCCGCCGGCGGCGGCGGCGTGAGCCACCGCGCAGGCAGCCGGGATTGCTTACCA	3
1	CCTGCAGCGTGCTTTCGGAGGCGGCGGCTGGCACGGAGGCCGGGCTACATGAGAAGCTCGACAGGGCCTG	4
	...	

↓

sequencesRDD - RDD[(cabecera, secuencia)]

>FJ804216.1 Homo sapiens disrupted in schizophrenia 1 isoform 46 (DISC1) mRNA, complete cds, alternatively spliced	GGAAGGAGCAGGAGGCAGCCCAGGCGGAGCGGGAGGAGCTGG CAGCGGGGCGCATGCCAGGCGGGGGTCCTCAGGGCGCCCCAG CCGCCGCCGGCGGCGGCGGCGTGAGCCACCGCGCAGGCAGCC GGGATTGCTTACCACCTGCAGCGTGCTTTCGGAGGCGGCGGC TGGCACGGAGGCCGGGCTACATGAGAAGCTCGACAGGGCCTG ...
...	