SARS-CoV-2 orthologs of pathogenesis-involved small viral RNAs of SARS-CoV

Ali Ebrahimpour Boroojeny and Hamidreza Chitsaz

Department of Computer Science, Colorado State University

http://chitsazlab.org
chitsaz@chitsazlab.org

Abstract

Background The COVID-19 pandemic clock is ticking and the survival of many of mankind’s modern institutions and or survival of many individuals is at stake. There is a need for treatments to significantly reduce the morbidity and mortality of COVID-19.

Hence, we delved deep into the SARS-CoV-2 genome, which is the virus that has caused COVID-19. SARS-CoV-2 is from the same family as SARS-CoV in which three small viral RNAs (svRNA) were recently identified [1]; those svRNAs play a significant role in the virus pathogenesis in mice.

Contribution In this paper, we report potential orthologs of those three svRNAs in the SARS-CoV-2 genome. Instead of off-the-shelf search and alignment algorithms, which failed to discover the orthologs, we used a special alignment scoring that does not penalize C/T and A/G mismatches. RNA bases C and U both can bind to G; similarly, A and G both can bind to U, hence, our scoring.

To validate our results, we confirmed the discovered orthologs are fully conserved in all the publicly available genomes of various strains of SARS-CoV-2; the loci at which the SARS-CoV-2 orthologs occur are close to the loci at which SARS-CoV svRNAs occur. We also report potential targets for these svRNAs. We hypothesize that the discovered orthologs play a role in pathogenesis of SARS-CoV-2, and therefore, antagomir-mediated inhibition of these SARS-CoV-2 svRNAs inhibits COVID-19.

1 Introduction

The world is now struggling with a pandemic known as COVID-19 which is caused by a novel coronavirus that was first identified in December 2019 in a local sea food market in Wuhan, China [2]. Due to the similarity of its genomic sequence to that of Severe Acute Respiratory Syndrome (SARS-CoV), which is a member of the subgenus of Sarbecovirus, the aforementioned novel coronavirus was named SARS-CoV-2. Phylogenetic studies have found a bat origin for this virus [3, 4]. As of April 29th, 2020, this virus has infected more than 3,150,000 people in 185 countries, caused more than 227,000 cases of death, and has become a global health concern leading to massive lock downs and quarantine all around the world.
Since the emergence of SARS-CoV in China in 2002, which infected around 8,000 people world-wide, multiple research efforts have tried to understand that virus and to suggest potential treatments. Despite the fact that no vaccines or antivirals have been approved to date for any of coronaviruses, improvements on reducing the severity of the disease and mortality rate have been reported. Because of the similarity of the recent fast-spreading coronavirus SARS-CoV-2, which shares more than 79% of its genomic sequence with SARS-CoV, one plausible way to understand how it works and suggest possible treatments would be to port what has been found for SARS-CoV previously to SARS-CoV-2.

Non-coding RNAs (ncRNAs) are, as the name suggests, RNAs that do not translate to proteins. Although it is likely the case that some of them do not play a major role in the cell [5, 6], some have crucial functions, such as transfer RNAs (tRNAs), ribosomal RNAs (rRNAs), micro RNAs (miRNAs), etc. Some of the ncRNAs, such as miRNAs, play a role in post-transcriptional regulation of gene expression. They, through a procedure called gene silencing, bind with the complementary parts of the target RNAs, and prevent the translation of those RNAs through cleavage of their strand, shortening their poly-A tail, or downgrading the efficiency of their translation by making some nucleotides unavailable to the ribosomes [7, 8].

First viral ncRNA was identified by Reich et al. [9]. Since then a plethora of viral-associated ncRNAs have been identified and this has been accelerated by the advances in technology [10]. Especially, deep sequencing has facilitated the detection of small virus-associated RNAs [11, 12]. Some of these ncRNAs are known to be responsible in counteracting the antiviral defense mechanism that are present in the host cells, mostly through inhibition of protein kinase R (PKR) [13]. Therefore, they aid in the life cycle of the virus [14, 15], such as svRNAs in influenza A virus that are involved in the mechanisms this virus uses for switching between transcription and replication [12].

It had been well-known that nuclear and DNA viruses encode miRNAs [16] that play a role in persistence [17] of the virus as well as changing the transcriptome in the host cell [18]. Using the deep sequencing technologies, it had been revealed that cytoplasmic RNA viruses also express ncRNAs [11, 12, 1] and most of them induce various cytoplasmic pathways to express their ncRNAs [12]. Flaviviruses can be mentioned as examples of cytoplasmic RNA viruses, which are very sensitive to interferons and have evolved a variety of mechanisms to avoid their action [19]. It has been shown that ncRNAs in flaviviral RNA binds to genes responsible for regulation of antiviral state of the host cell and affects the interferon response agains the virus [20].

A recent research has reported three small viral RNAs (svRNAs) that are derived from the genomic regions of SARS-CoV [1]. Morales et al. have shown the presence of these positive sense svRNAs, which are “mapped to nsp3 at the 5′ end of the Replicase gene and the N gene (svRNA-N) at the 3′ end of the genomic RNA (gRNA)” [1], by using specific small RNA RT-qPCR assays. Their experiments on a mouse model of the infection [21, 22] show that these svRNAs contribute to SARS-CoV pathogenesis, and also suggest a potential antiviral treatment using antagonim- mediated inhibition of these svRNAs.

Small non-coding RNAs that play an improtant roll have been shown to be highly conserved among genuses and families. Given that SARS-CoV-2 is in the the same subgenus as SARS-CoV and their genomic sequence has more than 79% similarity, if orthologs of the svRNAs found in SARS-CoV be present and found in SARS-CoV-2, then (antagomir-mediated) inhibition of those svRNA orthologs is expected to lead to reduction of SARS-CoV-2 titers and help decrease severity of COVID-19 in the majority of patients, as was shown to be the case for SARS-CoV [1]. What is needed is the sequence of the three aforementioned svRNAs in SARS-CoV-2 genome, from which
the corresponding antagonimirs are simply designed through base pair complementarity.

To target SARS-CoV-2 svRNAs, we first characterize the sequence of three svRNAs in SARS-CoV-2. We achieve that goal by aligning the sequences of SARS-CoV svRNAs to the SARS-CoV-2 genome sequences. After investigating many of the available sequenced genomes of SARS-CoV-2 that have been reported in various locations around the globe, we discovered the presence of three svRNAs that were highly conserved orthologs of the svRNAs that played a role in pathogenesis of SARS-CoV. This *in silico* discovery still needs to be confirmed using *in vitro* and *in vivo* experiments, but our findings reported in the following sections suggest strong likelihood of our hypothesis.

2 Methods

In order to find the svRNAs in SARS-CoV-2 that are orthologs of those in SARS-CoV, we selected all of the complete genomic sequences of SARS-CoV-2, that were made available on the NCBI portal as of March 27th (173 complete sequences) as well as 27 more randomly chosen ones among the more recent uploads on the NCBI portal. These genomic sequences are from different states and countries, including but not limited to: New York, Washington, California, Illinois, Utah, China, Japan, South Korea, Italy, India, Brazil, Germany, Australia, Turkey, and Greece.

We used a variant of the algorithm introduced by Smith and Waterman, which is known in the community as the *fit alignment* [23], as the core of investigation for the svRNAs of interest. The algorithm devised by Smith and Waterman is itself a variant of the *global alignment* algorithm known as Needleman-Wunsch algorithm [24], and can be used to find the regions of two genomic sequences that are similar. *Fit alignment* is a variant of this algorithm that searches a reference genome for a subsection that is highly similar to another shorter sequence.

We used our own in-house alignment tool that implements a faster variant of local and fit alignment algorithms, based on the idea of limiting the search to considering only a subset of the alignment search space that represents high similarity and pruning the regions that fall below the desired threshold. Also, when aligning two sequences, there might be multiple alignments with the same score as the highest score. Therefore, our tool keeps track of all the alignments with a score equal to the highest score. For this specific problem, we also added the feature of choosing the best alignment that exists in all the other reference sequences. However, we should note that, for all the sequences considered (the ones in Table 1) the loci reported for the exact match to our suggested svRNAs have also the highest score of alignment with the svRNAs reported by Morales *et al.* in their respective sequence.

3 Results

Figures 1, 2, and 3 show the sequences that we have found for three svRNAs in SARS-CoV-2 that are orthologs of the three aforementioned svRNAs in SARS-CoV. To further test our hypothesis, we searched for all these svRNAs in 200 different complete reference sequences of the virus. Our three svRNAs are wholly present, without any mutation, in all the reference sequences. Table 1 shows the NCBI ID of each of these sequences, as well as the string loci of each of the svRNAs. As you can see 199 out of 200 tested sequences are present in the table and contain the exact match of the proposed RNAs. The missing entry of the table is LR757997 which showed the presence of the third svRNA at loci 28604 but does not contain the other two because there is a gap in the
sequence from loci 3001 to 3235 (filled with Ns), and this is the region where the first two svRNAs reside in according to the loci values in the table for these two columns.

We identify a non-detrimental mismatch by : in Figures 1, 2, and 3. RNA-RNA binding energies is mainly governed by Watson-Crick base pairing, namely A-U, G-U, and C-G. Particularly, U can pair with both A and G. Hence, an A vs. G mismatch (substitution) in an ortholog RNA is non-detrimental for the binding to a target RNA. Similarly, G can pair with both U and C. Hence, an U/T vs. C mismatch (substitution) in an ortholog RNA is non-detrimental for the binding to a target RNA.

**Figure 1:** Alignment of the first svRNA in SARS-CoV and its identified ortholog in SARS-CoV-2. Top: nsp3.1 svRNA sequence GAGGAAGAAGAGGACGAT in SARS-CoV according to [1]. Bottom: Ortholog of nsp3.1 svRNA sequence GATGAAGAAGAAGGTGAT in SARS-CoV-2. | represents a match, — represents a gap (indel), empty represents a mismatch, and : represents a non-detrimental mismatch.

**Figure 2:** Alignment of the second svRNA in SARS-CoV and its identified ortholog in SARS-CoV-2. Top: nsp3.2 svRNA sequence GAGGAAGAAGAGGACGAT in SARS-CoV according to [1]. Bottom: Ortholog of nsp3.2 svRNA sequence GAAGAAGAGCAAGAAGAAGATT in SARS-CoV-2. | represents a match, — represents a gap (indel), empty represents a mismatch, and : represents a non-detrimental mismatch.

**Figure 3:** Alignment of the third svRNA in SARS-CoV and its identified ortholog in SARS-CoV-2. Top: N svRNA sequence AGGAACTGGCCAGAAGCTTC in SARS-CoV according to [1]. Bottom: Ortholog of N svRNA sequence AGGAACTGGCCAGAAGCTGGAC in SARS-CoV-2. | represents a match, — represents a gap (indel), empty represents a mismatch, and dotted line represents a potential omission. It is possible that some or all of the last four (4) nucleotides GGAC of the bottom sequence are dropped (omitted).
| NCBI ID   | svRNA1 | svRNA2 | svRNA3 |
|-----------|--------|--------|--------|
| LC528232  | 3056   | 3182   | 28615  |
| LC528233  | 3061   | 3187   | 28620  |
| LC529905  | 3053   | 3179   | 28612  |
| LC529996  | 3033   | 3159   | 28592  |
| LR757995  | 3038   | 3164   | 28597  |
| LR757998  | 3028   | 3154   | 28587  |
| MN908947  | 3053   | 3179   | 28612  |
| MN938384  | 3021   | 3147   | 28580  |
| MN975262  | 3053   | 3179   | 28612  |
| MN985325  | 3053   | 3179   | 28612  |
| MN988668  | 3052   | 3178   | 28611  |
| MN988669  | 3052   | 3178   | 28611  |
| MN988713  | 3053   | 3179   | 28612  |
| MN994467  | 3053   | 3179   | 28612  |
| MN994468  | 3053   | 3179   | 28612  |
| MN996527  | 3020   | 3146   | 28579  |
| MN996528  | 3053   | 3179   | 28612  |
| MN996529  | 3041   | 3167   | 28600  |
| MN996530  | 3039   | 3165   | 28598  |
| MN996531  | 3040   | 3166   | 28599  |
| MN997409  | 3053   | 3179   | 28612  |
| MT007544  | 3053   | 3179   | 28612  |
| MT012098  | 3040   | 3166   | 28596  |
| MT019529  | 3053   | 3179   | 28612  |
| MT019530  | 3053   | 3179   | 28612  |
| MT019531  | 3053   | 3179   | 28612  |
| MT019532  | 3053   | 3179   | 28612  |
| MT019533  | 3053   | 3179   | 28612  |
| MT020781  | 3053   | 3179   | 28612  |
| MT020880  | 3053   | 3179   | 28612  |
| MT020881  | 3053   | 3179   | 28612  |
| MT027062  | 3053   | 3179   | 28612  |
| MT027063  | 3053   | 3179   | 28612  |
| MT027064  | 3053   | 3179   | 28612  |
| MT039873  | 3050   | 3176   | 28609  |
| MT039887  | 3053   | 3179   | 28609  |
| MT039888  | 3053   | 3179   | 28612  |
| MT039890  | 3053   | 3179   | 28612  |
| MT044257  | 3053   | 3179   | 28612  |
| MT044258  | 3029   | 3155   | 28588  |
| MT049951  | 3053   | 3179   | 28612  |
| MT050493  | 3033   | 3159   | 28592  |
| MT066156  | 3053   | 3179   | 28612  |
| MT066175  | 3053   | 3179   | 28612  |
| MT066176  | 3053   | 3179   | 28612  |
|   | MT072688  | 3038 | 3164 | 28597 |
|---|-----------|------|------|-------|
| 46| MT093571  | 3053 | 3179 | 28612 |
| 47| MT093631  | 3040 | 3166 | 28599 |
| 48| MT106052  | 3053 | 3179 | 28612 |
| 49| MT106053  | 3053 | 3179 | 28612 |
| 50| MT106054  | 3053 | 3179 | 28612 |
| 51| MT118835  | 3053 | 3179 | 28612 |
| 52| MT121215  | 3053 | 3179 | 28612 |
| 53| MT123290  | 3056 | 3182 | 28615 |
| 54| MT123291  | 3050 | 3176 | 28609 |
| 55| MT123292  | 3053 | 3179 | 28612 |
| 56| MT123293  | 3047 | 3173 | 28606 |
| 57| MT126808  | 3053 | 3179 | 28612 |
| 58| MT135041  | 3053 | 3179 | 28612 |
| 59| MT135042  | 3053 | 3179 | 28612 |
| 60| MT135043  | 3053 | 3179 | 28612 |
| 61| MT135044  | 3053 | 3179 | 28612 |
| 62| MT152824  | 3051 | 3177 | 28610 |
| 63| MT159705  | 3053 | 3179 | 28612 |
| 64| MT159706  | 3053 | 3179 | 28612 |
| 65| MT159707  | 3053 | 3179 | 28612 |
| 66| MT159708  | 3053 | 3179 | 28612 |
| 67| MT159709  | 3053 | 3179 | 28612 |
| 68| MT159710  | 3053 | 3179 | 28612 |
| 69| MT159711  | 3053 | 3179 | 28612 |
| 70| MT159712  | 3053 | 3179 | 28612 |
| 71| MT159713  | 3053 | 3179 | 28612 |
| 72| MT159714  | 3053 | 3179 | 28612 |
| 73| MT159715  | 3053 | 3179 | 28612 |
| 74| MT159716  | 3038 | 3164 | 28597 |
| 75| MT159717  | 3053 | 3179 | 28612 |
| 76| MT159718  | 3053 | 3179 | 28612 |
| 77| MT159719  | 3053 | 3179 | 28612 |
| 78| MT159720  | 3053 | 3179 | 28612 |
| 79| MT159721  | 3053 | 3179 | 28612 |
| 80| MT159722  | 3053 | 3179 | 28612 |
| 81| MT163716  | 3053 | 3179 | 28612 |
| 82| MT163717  | 3047 | 3173 | 28606 |
| 83| MT163718  | 3053 | 3179 | 28612 |
| 84| MT163719  | 3053 | 3179 | 28612 |
| 85| MT184907  | 3053 | 3179 | 28612 |
| 86| MT184908  | 3053 | 3179 | 28612 |
| 87| MT184909  | 3053 | 3179 | 28612 |
| 88| MT184910  | 3053 | 3179 | 28612 |
| 89| MT184911  | 3053 | 3179 | 28612 |
| 90| MT184912  | 3053 | 3179 | 28612 |
| 91| MT184913  | 3053 | 3179 | 28612 |
|   | MT184913  | 3053 | 3179 | 28612 |
|---|-----------|------|------|-------|
|92 | MT188339  | 2999 | 3125 | 28558 |
|93 | MT188340  | 2999 | 3125 | 28558 |
|94 | MT188341  | 2999 | 3125 | 28561 |
|95 | MT192759  | 3026 | 3152 | 28585 |
|96 | MT192765  | 3046 | 3172 | 28605 |
|97 | MT192772  | 3053 | 3179 | 28612 |
|98 | MT192773  | 3052 | 3178 | 28611 |
|99 | MT226610  | 3053 | 3179 | 28612 |
|100| MT233519  | 2999 | 3125 | 28558 |
|101| MT233520  | 2999 | 3125 | 28558 |
|102| MT233521  | 2999 | 3125 | 28558 |
|103| MT233522  | 2999 | 3125 | 28558 |
|104| MT233523  | 2999 | 3125 | 28558 |
|105| MT240479  | 3017 | 3143 | 28576 |
|106| MT240489  | 2979 | 3105 | 28538 |
|107| MT246450  | 3022 | 3148 | 28581 |
|108| MT246451  | 3008 | 3134 | 28567 |
|109| MT246452  | 3046 | 3172 | 28605 |
|110| MT246453  | 2979 | 3105 | 28538 |
|111| MT246454  | 3044 | 3170 | 28603 |
|112| MT246455  | 3008 | 3134 | 28567 |
|113| MT246456  | 2996 | 3122 | 28555 |
|114| MT246457  | 2997 | 3123 | 28553 |
|115| MT246458  | 2927 | 3053 | 28486 |
|116| MT246459  | 3047 | 3173 | 28606 |
|117| MT246460  | 3053 | 3179 | 28612 |
|118| MT246461  | 3021 | 3147 | 28580 |
|119| MT246462  | 3053 | 3179 | 28612 |
|120| MT246463  | 2927 | 3053 | 28486 |
|121| MT246464  | 3002 | 3128 | 28561 |
|122| MT246465  | 2931 | 3057 | 28490 |
|123| MT246466  | 3047 | 3173 | 28606 |
|124| MT246467  | 3050 | 3176 | 28609 |
|125| MT246468  | 2993 | 3119 | 28552 |
|126| MT246469  | 3019 | 3145 | 28578 |
|127| MT246470  | 3035 | 3161 | 28594 |
|128| MT246471  | 3021 | 3147 | 28580 |
|129| MT246472  | 2968 | 3094 | 28527 |
|130| MT246473  | 2979 | 3105 | 28535 |
|131| MT246474  | 3043 | 3169 | 28602 |
|132| MT246475  | 3033 | 3159 | 28592 |
|133| MT246476  | 3014 | 3140 | 28573 |
|134| MT246477  | 3022 | 3148 | 28581 |
|135| MT246478  | 3040 | 3166 | 28599 |
|136| MT246479  | 2979 | 3105 | 28538 |
|   |   |   |   |
|---|---|---|---|
| 138 | MT246480 | 3052 | 3178 | 28611 |
| 139 | MT246481 | 3021 | 3147 | 28580 |
| 140 | MT246482 | 2996 | 3122 | 28555 |
| 141 | MT246483 | 2927 | 3053 | 28486 |
| 142 | MT246484 | 3021 | 3147 | 28580 |
| 143 | MT246485 | 2925 | 3051 | 28484 |
| 144 | MT246486 | 3021 | 3147 | 28580 |
| 145 | MT246487 | 3037 | 3163 | 28596 |
| 146 | MT246488 | 3026 | 3152 | 28585 |
| 147 | MT246489 | 3003 | 3129 | 28562 |
| 148 | MT246490 | 2997 | 3123 | 28556 |
| 149 | MT251972 | 3016 | 3142 | 28575 |
| 150 | MT251973 | 3049 | 3175 | 28608 |
| 151 | MT251974 | 3002 | 3128 | 28561 |
| 152 | MT251975 | 3025 | 3151 | 28584 |
| 153 | MT251976 | 3051 | 3177 | 28610 |
| 154 | MT251977 | 2979 | 3105 | 28538 |
| 155 | MT251978 | 3051 | 3177 | 28610 |
| 156 | MT251979 | 3002 | 3128 | 28561 |
| 157 | MT251980 | 3002 | 3128 | 28558 |
| 158 | MT253696 | 2999 | 3125 | 28558 |
| 159 | MT253697 | 2999 | 3125 | 28558 |
| 160 | MT253698 | 2999 | 3125 | 28558 |
| 161 | MT253699 | 2999 | 3125 | 28558 |
| 162 | MT253700 | 2999 | 3125 | 28558 |
| 163 | MT253701 | 2999 | 3125 | 28558 |
| 164 | MT253702 | 2999 | 3125 | 28558 |
| 165 | MT253703 | 2999 | 3125 | 28558 |
| 166 | MT253704 | 2999 | 3125 | 28558 |
| 167 | MT253705 | 2999 | 3125 | 28558 |
| 168 | MT253706 | 2999 | 3125 | 28558 |
| 169 | MT253707 | 2999 | 3125 | 28558 |
| 170 | MT253708 | 2999 | 3125 | 28558 |
| 171 | MT253709 | 2999 | 3125 | 28558 |
| 172 | MT253710 | 2999 | 3125 | 28558 |
| 173 | MT327745 | 3049 | 3175 | 28608 |
| 174 | MT328032 | 3053 | 3179 | 28612 |
| 175 | MT334563 | 3052 | 3178 | 28611 |
| 176 | MT345880 | 3006 | 3132 | 28565 |
| 177 | MT350251 | 3037 | 3163 | 28596 |
| 178 | MT350266 | 3053 | 3179 | 28612 |
| 179 | MT350282 | 3053 | 3179 | 28612 |
| 180 | MT359866 | 3043 | 3169 | 28602 |
| 181 | MT370944 | 3013 | 3139 | 28572 |
| 182 | MT370968 | 2989 | 3115 | 28548 |
| 183 | MT370975 | 2998 | 3124 | 28557 |
|    |    MT371019 | 2998  | 3124  | 28557 |
|----|------------|-------|-------|-------|
| 185| MT371024   | 2981  | 3107  | 28540 |
| 186| MT371034   | 3006  | 3132  | 28565 |
| 187| MT371035   | 2998  | 3124  | 28557 |
| 188| MT371036   | 2998  | 3124  | 28557 |
| 189| MT371037   | 2989  | 3115  | 28548 |
| 190| MT371048   | 3053  | 3179  | 28612 |
| 191| MT371568   | 2929  | 3055  | 28488 |
| 192| MT371572   | 2945  | 3071  | 28504 |
| 193| MT372481   | 3048  | 3174  | 28607 |
| 194| MT374112   | 3051  | 3177  | 28610 |
| 195| MT375470   | 3021  | 3147  | 28580 |
| 196| MT385448   | 3050  | 3176  | 28609 |
| 197| MT394529   | 3042  | 3168  | 28601 |
| 198| MT394864   | 2999  | 3125  | 28558 |
| 199| MT396242   | 3028  | 3154  | 28587 |

Table 1: Start loci of the three svRNAs in various strains of SARS-CoV-2 genome.
3.1 Potential targets

We also analyzed the potential target genes of these svRNAs in human body. To this end, we considered more than 32,000 known RNAs that are transcribed in human body. We processed the sequence of each of these RNAs and removed the intron regions of each. Finally, we aligned the reverse complement of our svRNAs to all the exon sequences. A highly similar region to the reverse complement of an svRNA is highly complementary to the svRNA and hence suggests a high chance of interacting with that svRNA. Tables showing the results of that alignment for the regions that had an alignment score higher than 0.7 with these three svRNAs are available in the Appendix Tables 2, 3, and 4. To compute the score of the alignment, we used a reward of one (1) for a pair of matching nucleotides and a penalty of negative one (-1) for substitutions and insertion/deletions (indels). In the end, the total score was divided by the length of each svRNA to normalize the scores.

4 Analysis

Our proposed svRNAs for SARS-CoV-2 are highly conserved versions of the svRNAs of SARS-CoV. Presence of these svRNAs in all the 200 reference sequences that we used to test our hypothesis increases the possibility of correctness of our claim. Also, our proposed svRNAs occur at very similar loci in different reference sequences, and these loci are almost the same as the ones for the original svRNAs reported for SARS-CoV. The fact that the two viruses are in the same subgenus makes our hypothesis more plausible. However, still this hypothesis has to be verified experimentally.

As mentioned earlier, the complete tables showing the possible target RNAs of our proposed svRNAs are available in the supplementary document. However, it is worth mentioning some of them. The second highest match to the reverse complement of the first svRNA is HIF3A transcript which is a transcriptional regulator in adaptive response to low oxygen levels. Silencing this gene affects the reaction of the body in response to hypoxia. The second best match with the reverse complement of the second svRNA is MEX3B transcript which is a member of MEX3 translational regulators. MEX3 are RNA-binding proteins that are evolutionarily conserved and their in vivo functions is yet to be fully characterized.

5 Conclusion

In this paper, we reported three potential svRNAs, which are orthologs of SARS-CoV svRNAs, in the SARS-CoV-2 genome. To validate our results, we confirmed the discovered orthologs are fully conserved in all the publicly available genomes of various strains of SARS-CoV-2; the loci at which the SARS-CoV-2 orthologs occur are close to the loci at which SARS-CoV svRNAs occur. Furthermore, our proposed svRNAs occur at very similar loci in different reference sequences, and these loci are almost the same as the ones for the original svRNAs reported in SARS-CoV.

We also reported potential targets for these svRNAs. We hypothesize that the discovered orthologs play a role in pathogenesis of SARS-CoV-2, and therefore, antagomir-mediated inhibition of these SARS-CoV-2 svRNAs inhibits COVID-19. This in silico discovery still needs to be confirmed using in vitro and in vivo experiments.
References

[1] Lucía Morales, Juan Carlos Oliveros, Raúl Fernandez-Delgado, Benjamin Robert tenOever, Luis Enjuanes, and Isabel Sola. Sars-cov-encoded small rnas contribute to infection-associated lung pathology. *Cell host & microbe*, 21(3):344–355, 2017.

[2] Hongzhou Lu, Charles W Stratton, and Yi-Wei Tang. Outbreak of pneumonia of unknown etiology in wuhan china: the mystery and the miracle. *Journal of Medical Virology*.

[3] Roujian Lu, Xiang Zhao, Juan Li, Peihua Niu, Bo Yang, Honglong Wu, Wenling Wang, Hao Song, Baoying Huang, Na Zhu, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *The Lancet*, 395(10224):565–574, 2020.

[4] Yushun Wan, Jian Shang, Rachel Graham, Ralph S Baric, and Fang Li. Receptor recognition by the novel coronavirus from wuhan: an analysis based on decade-long structural studies of sars coronavirus. *Journal of virology*, 94(7), 2020.

[5] Jürgen Brosius. Waste not, want not–transcript excess in multicellular eukaryotes. *TRENDS in Genetics*, 21(5):287–288, 2005.

[6] Alexander F Palazzo and Eliza S Lee. Non-coding rna: what is functional and what is junk? *Frontiers in genetics*, 6:2, 2015.

[7] Marc Robert Fabian, Nahum Sonenberg, and Witold Filipowicz. Regulation of mrna translation and stability by micrornas. *Annual review of biochemistry*, 79:351–379, 2010.

[8] David P Bartel. Micrnas: target recognition and regulatory functions. *cell*, 136(2):215–233, 2009.

[9] Paul R Reich, Bernard G Forget, Sherman M Weissman, and James A Rose. Rna of low molecular weight in kb cells infected with adenovirus type 2. *Journal of molecular biology*, 17(2):428–439, 1966.

[10] Kazimierz T Tykowski, Yang Eric Guo, Nara Lee, Walter N Moss, Tenaya K Vallery, Mingyi Xie, and Joan A Steitz. Viral noncoding rnas: more surprises. *Genes & development*, 29(6):567–584, 2015.

[11] Poornima Parameswaran, Ella Sklan, Courtney Wilkins, Trever Burgon, Melanie A Samuel, Rui Lu, K Mark Ansel, Vigo Heissmeyer, Shirit Einav, William Jackson, et al. Six rna viruses and forty-one hosts: viral small rnas and modulation of small rna repertoires in vertebrate and invertebrate systems. *PLoS pathogens*, 6(2), 2010.

[12] Jasmine T Perez, Andrew Varble, Ravi Sachidanandam, Ivan Zlatev, Muthiah Manoharan, Adolfo Garcia-Sastre, et al. Influenza a virus-generated small rnas regulate the switch from transcription to replication. *Proceedings of the National Academy of Sciences*, 107(25):11525–11530, 2010.

[13] Joan Steitz, Sumit Borah, Demian Cazalla, Victor Fok, Robin Lytle, Rachel Mitton-Fry, Kasandra Riley, and Tasleem Samji. Noncoding rns of viral origin. *Cold Spring Harbor perspectives in biology*, 3(3):a005165, 2011.

[14] Amiya K Banerjee. Transcription and replication of rhabdoviruses. *Microbiological reviews*, 51(1):66, 1987.

[15] Bryan R Cullen. Viral and cellular messenger rna targets of viral micrornas. *Nature*, 457(7228):421–425, 2009.

[16] Bryan R Cullen. Viruses and micrornas: Riscy interactions with serious consequences. *Genes & development*, 25(18):1881–1894, 2011.

[17] Benjamin R Tenoever. Rna viruses and the host microrna machinery. *Nature Reviews Microbiology*, 11(3), 2013.

[18] David P Bartel. Micrnas: genomics, biogenesis, mechanism, and function. *cell*, 116(2):281–297, 2004.

[19] Michael S Diamond. Mechanisms of evasion of the type i interferon antiviral response by flaviviruses. *Journal of interferon & cytokine research*, 29(9):521–530, 2009.

[20] Katell Bidet, Dhivya Dadlani, and Mariano A Garcia-Blanco. G3bp1, g3bp2 and caprin1 are required for translation of interferon stimulated mrnas and are targeted by a dengue virus non-coding rna. *PLoS pathogens*, 10(7), 2014.
[21] Anjeanette Roberts, Damon Deming, Christopher D Paddock, Aaron Cheng, Boyd Yount, Leatrice Vogel, Brian D Herman, Tim Sheahan, Mark Heise, Gillian L Genrich, et al. A mouse-adapted sars-coronavirus causes disease and mortality in balb/c mice. *PLoS pathogens*, 3(1), 2007.

[22] Marta L DeDiego, Enrique Álvarez, Fernando Almazán, María Teresa Rejas, Elaine Lamirande, Anjeanette Roberts, Wun-Ju Shieh, Sherif R Zaki, Kanta Subbarao, and Luis Enjuanes. A severe acute respiratory syndrome coronavirus that lacks the e gene is attenuated in vitro and in vivo. *Journal of virology*, 81(4):1701–1713, 2007.

[23] Temple F Smith and Michael S Waterman. Comparison of biosequences. *Advances in Applied Mathematics*, 2(4):482–489, 1981.

[24] Saul B Needleman and Christian D Wunsch. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of molecular biology*, 48(3):443–453, 1970.
Appendices

Tables 2, 3, and 4 show the matches with a score of at least 70% between more than 32,000 RNAs transcribed in the human body and reverse-complement of our proposed svRNAs for Sars-CoV-2. The ones with a high score are likely to be target genes of the corresponding svRNA. The entries of each table are sorted based on the normalized score of the alignment.

Tables 5, 6, and 7 show the matches with a score of at least 80% between our proposed svRNAs and human reference genome. Patch release 13 of build 38 was used for as the reference genome.
| Gene ID            | Score | Matched Sequence                      | Loci  |
|-------------------|-------|---------------------------------------|-------|
| ENSG000000168421  | 0.94  | ATCACCTTCTTCTTCTCAGTC                 | 20    |
| ENSG000000196218  | 0.89  | ATCACCTTCTTCTTCTTCAGTC               | 15459 |
| ENSG000000124440  | 0.89  | ATCACCTTCTTCTTCTTCAGTC               | 5383  |
| ENSG000000234545  | 0.83  | ATTCATCTTCTTCTTCATC                  | 1521  |
| ENSG000000186867  | 0.83  | ATCACCTTCTTCTTCTTCATC                | 692   |
| ENSG000000108100  | 0.83  | ATCACCTTCTTCTTCTTCATC                | 4064  |
| ENSG00000119632   | 0.83  | ATCACCTTCTTCTTCTTCATC                | 1524  |
| ENSG00000168038   | 0.83  | ATCACCTGCTTCTTCTTCATC                | 5803  |
| ENSG00000229913   | 0.83  | ATCACCTTCTTCTTCTTCATC                | 274   |
| ENSG00000137767   | 0.78  | ATCA-CTTCTTCTTCTTCATC                | 315   |
| ENSG00000135902   | 0.78  | ATCACCTTCTTCTTCTTCATC                | 929   |
| ENSG00000136758   | 0.78  | ATCAGCCTTCTTCTTCTTCATC               | 1553  |
| ENSG00000250448   | 0.78  | ATCACCTTCTTCTTCTTCTTCATC             | 1192  |
| ENSG00000105675   | 0.78  | ATCACCATCTTCTTCTTCATC                | 2635  |
| ENSG0000013375    | 0.78  | ATCACCTTACTTCTTCTTCATC               | 6738  |
| ENSG00000151553   | 0.78  | ATCACCTTCTTCTTCTTCATC                | 3956  |
| ENSG00000157388   | 0.78  | ATCACCTTCTTCTTCTTCATC                | 5845  |
| ENSG00000173227   | 0.78  | ATCACCTTCTTCTTCTTCATC                | 10806 |
| ID     | Accession  | Score | Sequence 1   | Sequence 2         | Length |
|--------|------------|-------|--------------|--------------------|--------|
| 19     | ENSG00000163485 | 0.78  | ATCACCTTC-T-TCTTCATC | AT-CACCTTCTTCTTCATC | 1300   |
| 20     | ENSG00000103599  | 0.78  | ATACCTTTCTTCTTCATC | ATCACCTTCTTCTTCATC | 5911   |
| 21     | ENSG00000148935  | 0.78  | TTCACCTTCTTCTTCATC | ATCACCTTCTTCTTCATC | 2061   |
| 22     | ENSG00000106278  | 0.78  | ATCATCTTTCTCTATCATC | ATC-ACCTTCTTTCTTCATC | 7929   |
| 23     | ENSG00000115232  | 0.78  | ATACCTTTCTTCTTCATC | ATCACCTTCTTCTTCATC | 1025   |
| 24     | ENSG00000184083  | 0.78  | ATCTACCTTTCTTCATC | ATCACCTTCTTCTTCATC | 4572   |
| 25     | ENSG00000188687  | 0.78  | ATCACCTTTCTTCTTCATC | ATCACCTTCTTCTTCATC | 3941   |
| 26     | ENSG000000279010 | 0.78  | ACCACCTTCTTCTTCATC | ATCACCT-CTTCTTCTTCATC | 4323   |
| 27     | ENSG000000188352 | 0.78  | ATCACCTTCTTCTTCATC | ATCACCTTCTTCTTCATC | 188    |
| 28     | ENSG00000276460  | 0.78  | ATGCACCTTTCTTCTTCATC | ATCACCTTCTTCTTCATC | 685    |
| 29     | ENSG00000100181  | 0.78  | ATGACCTTTCTTCTTCATC | ATCACCTTCTTCTTCATC | 8513   |
| 30     | ENSG00000081148  | 0.78  | ATCATCTACTTCTTCTTCATC | ATCACCTTCTTCTTCATC | 3543   |
| 31     | ENSG00000173208  | 0.78  | ATCACCTACTTATATTCATC | ATCACCTTCTTCTTCATC | 4148   |
| 32     | ENSG00000116783  | 0.78  | CTACCTCTCTCTCCTTCATC | ATCACCTTCTTCTTCATC | 3911   |
| 33     | ENSG00000237654  | 0.78  | ATCACCTTCTTCAAACCTCCATC | ATCACCTTCTTCTTCATC | 1902   |
| 34     | ENSG00000117834  | 0.78  | AT-ACCTTTCTTCTTCATC | ATCACCTTCTTCTTCATC | 703    |
| 35     | ENSG00000198793  | 0.78  | ATCACCTTCTTCTTCATC | ATCACCTTCTTCTTCATC | 3000   |
|   | ENSG00000128973 | 0.78 | ATCCCCCTTCTTCCTCATC ATCACCTTCTTCCTCATC |
|---|----------------|------|----------------------------------------|
| 37 | ENSG00000118596 | 0.78 | ATCTCCTGCTTCTTCATC ATCACCTTCTTCCTCATC |
| 38 | ENSG00000164778 | 0.78 | ATCACCAACTTCTTCATC ATCACCTTCTTCCTCATC |
| 39 | ENSG0000012660  | 0.78 | A-CACTTTCTTCTTCATC ATCACCTTCTTCCTCATC |
| 40 | ENSG00000138161 | 0.78 | ATCACCTT-TCCTGCATC ATCACCTTCTTCCTCATC |
| 41 | ENSG00000116329 | 0.78 | ATC-TCTTCTTCTTCATC ATCACCTTCTTCCTCATC |
| 42 | ENSG00000237693 | 0.78 | ATCACCTTCTTCTTCATC ATCACCTTCTTCCTCATC |
| 43 | ENSG00000150712 | 0.78 | AGCACCTTCTTCTTCATC ATCACCTTCTTCCTCATC |
| 44 | ENSG00000152932 | 0.78 | ATACAACCTTCTTCTTCATC AT-C-ACCTTCTTCCTCATC |
| 45 | ENSG00000229373 | 0.78 | ATTCAGCCTTCTTCTTCATC A-TCA-C-CTTCTTCCTCATC |
| 46 | ENSG00000146574 | 0.78 | AT-ACCTTCTACTCTTCATC ATCACCTTCTTCCTCATC |
| 47 | ENSG00000198589 | 0.78 | ATCACGTCTTCTTCTTAATC ATCACCTTCTTCCTCATC |
| 48 | ENSG00000107611 | 0.78 | ATCAACKCTTCTACTTCATC ATCACCTTCTTCCTCATC |
| 49 | ENSG00000109501 | 0.78 | CTCACTCTTCTTCTTCATC ATCACCTTCTTCCTCATC |
| 50 | ENSG00000016402 | 0.78 | ATCACCTTCTTATCCATC ATCACCTTCTTCCTCATC |
| 51 | ENSG00000094963 | 0.78 | ATCACCTTCTTCTTCATC ATCACCTTCTTCCTCATC |
| 52 | ENSG00000151067 | 0.78 | ATCTCCATCTTCTTCATC ATCACCTTCTTCCTCATC |
| 53 | ENSG00000183853 | 0.78 | CTCAATCTTCCTTCCTCATC ATCACCTTCTTCCTCATC |
ATCACCTTCTTCTTCATC

54  ENSG00000114841  0.78  ATCACCTTCTTCTTCATC
      ATCACCTTCTTCTTCATC  8495

55  ENSG00000197077  0.78  ATCACCTTCTTCTTCTTGGC
      ATCACCTTCTTCTTCATC  5266

56  ENSG00000083544  0.78  ATCACCTTCTTCTTCTTGC
      ATCACCTTCTTCTTCATC  5611

57  ENSG00000187533  0.78  ATCACCTTCTTCTTCTTCATC
      ATCACCTTCTTCTTCATC  4084

58  ENSG00000147724  0.78  ATC-CCTGTCATTCTTTCATC
      ATCACCTTCTTCTTCATC  719

59  ENSG00000143126  0.78  CTCACCTTCTTCTTCTTCATC
      ATCACCTTCTTCTTCATC  7471

60  ENSG00000185950  0.78  ATCACCTTCTTCTTCTTCATC
      ATCACCTTCTTCTTCATC  1270

61  ENSG00000266885  0.78  CTCACCTTCTTCTTCATC
      ATCACCTTCTTCTTCATC  1963

62  ENSG00000198796  0.78  ATCACCTGCTGTCTTTCATC
      ATCACCTTCTTCTTCATC  9006

63  ENSG00000081059  0.78  ATCACCTGCTGTCTTTCATC
      ATCACCTTCTTCTTCATC  3422

64  ENSG00000091986  0.78  ATCACCTGCTGTCTTTCATC
      ATCACCTTCTTCTTCATC  5768

65  ENSG00000100142  0.78  ATCACCTTCTTCTTCATC
      ATCACCTTCTTCTTCATC  68

66  ENSG00000237921  0.78  ATCACCTTCTTCTTCATC
      ATCACCTTCTTCTTCATC  272

67  ENSG00000074706  0.78  ATCACCTTCTTCTTCATC
      ATCACCTTCTTCTTCATC  3106

68  ENSG00000164418  0.78  ATCACCTTCTTCTTCATC
      ATCACCTTCTTCTTCATC  5587

69  ENSG00000167216  0.78  ATCACCTTCTTCTTCATC
      ATCACCTTCTTCTTCATC  4520

70  ENSG00000257230  0.78  CTACCTTCTTCTTCATC
      ATCACCTTCTTCTTCATC  214
|   | ENSG00000168329  | 0.78 | ACCGCTTCTTCTTCATC  |
|   |                 |     | ATCACCTTCTTCTTCATC  |
| 71 |                 |     | 955                |
|   | ENSG00000160818 | 0.78 | ATCACCTTCTTCTTCATC  |
|   |                 |     | ATCACCTTCTTCTTCATC  |
| 72 |                 |     | 399                |
|   | ENSG00000089199 | 0.78 | ATCACCTTCTTCTTCATC  |
|   |                 |     | ATCACCTTCTTCTTCATC  |
| 73 |                 |     | 1384               |
|   | ENSG00000166847 | 0.78 | ATGACCTT-TTCTTCATC  |
|   |                 |     | ATCACCTTCTTCTTCATC  |
| 74 |                 |     | 4221               |
|   | ENSG00000259030 | 0.78 | CTCACCTTCTTCTTCATC  |
|   |                 |     | ATCACCTTCTTCTTCATC  |
| 75 |                 |     | 9070               |
|   | ENSG00000231728 | 0.78 | AGTACCTTCTTCTTCATC  |
|   |                 |     | A-TCACCTTCTTCTTCATC |
| 76 |                 |     | 1153               |
|   | ENSG00000095564 | 0.78 | ATCAGTTCTTCCTTCTTCATC |
|   |                 |     | ATCA-CCTT-CTTCTTCATC  |
| 77 |                 |     | 2590               |
|   | ENSG00000100714 | 0.78 | ATCACGGATTT-TTCTTCATC |
|   |                 |     | ATCACCT-TTCTTCTTCATC |
| 78 |                 |     | 130                |
|   | ENSG00000231216 | 0.78 | CTCATCTCTCTTCTTCATC  |
|   |                 |     | ATCA-CCTT-CTTCTTCATC  |
| 79 |                 |     | 2111               |
|   | ENSG00000231216 | 0.78 | ATCACCTTCTTCTTCATC  |
|   |                 |     | 3157               |
| 80 |                 |     | 5028               |
|   | ENSG00000234224 | 0.78 | ATACAACTTCCTTCTTCATC |
|   |                 |     | AT-CACCTT-CTTCTTCATC |
| 81 |                 |     | 8585               |
|   | ENSG0000017209  | 0.78 | ATGACCTTC-TTCTTCATC  |
|   |                 |     | ATCACCTTCTTCTTCATC  |
| 82 |                 |     | 15860              |
|   | ENSG00000177200 | 0.78 | ATCATCTTCTTCTTCATC  |
|   |                 |     | ATCACCTTCTTCTTCATC  |
| 83 |                 |     | 9797               |
|   | ENSG00000234224 | 0.78 | TTCACCTTCTTCTTCATC  |
|   |                 |     | ATCACCTTCTTCTTCATC  |
| 84 |                 |     | 792                |
|   | ENSG00000182568 | 0.78 | ATC-TCTTCTTCTTCATC  |
|   |                 |     | ATCACCTTCTTCTTCATC  |
| 85 |                 |     | 3896               |
|   | ENSG00000257964 | 0.78 | ATCACCTTCTTCTTCATC  |
|   |                 |     | 321                |
|
|    | ENSG00000251615 | 0.78 | TTCACCTTCTTTCTTCATC | 1388 |
|----|----------------|------|---------------------|------|
|    | ATCACCTTCTTTCTTCATC |      |                     |      |
|    | ENSG00000164185 | 0.78 | ATCATCTTCTTTGTTTCATC | 1070 |
|    | ATCACCTTCTTTGTTTCATC |      |                     |      |
|    | ENSG00000164185 | 0.78 | A-CACCTTCTTTGTTTCATC | 3526 |
|    | ATCACCTTCTTTGTTTCATC |      |                     |      |
|    | ENSG00000176783 | 0.72 | ATCACCTTCTTTCTTCATC | 4534 |
|    | ATCACCTTCTTTCTTCATC |      |                     |      |
|    | ENSG00000226145 | 0.72 | AGCACCTTCTTTCTTCATC | 32  |
|    | ATCACCTTCTTTCTTCATC |      |                     |      |
|    | ENSG0000000151413 | 0.72 | ATGAACCTCTTTCTTTTCATC | 6002 |
|    | ATGAACCTCTTTCTTTTCATC |      |                     |      |
|    | ENSG000000184014 | 0.72 | ATCATC-TCTTCTTTTCATC | 7058 |
|    | ATCACCTTCTTTCTTCATC |      |                     |      |
|    | ENSG000000244242 | 0.72 | ATCATC-TCTTCTTTTCATC | 1867 |
|    | ATCACCTTCTTTCTTCATC |      |                     |      |
|    | ENSG000000184014 | 0.72 | ATGAACCTCTTTCTTTTCATC | 6002 |
|    | ATGAACCTCTTTCTTTTCATC |      |                     |      |
|    | ENSG0000000279675 | 0.72 | ATCCCCCTACTTCTTTTCAGC | 3155 |
|    | ATCCACCTTCTTTCTTCATC |      |                     |      |
|    | ENSG000000165424 | 0.72 | TTGCACCTTCTTTCTTTATC | 4332 |
|    | TTGCACCTTCTTTCTTTATC |      |                     |      |
|    | ENSG0000000279675 | 0.72 | ATCCCCCTACTTCTTTTCAGC | 3155 |
|    | ATCCACCTTCTTTCTTCATC |      |                     |      |
|    | ENSG000000205086 | 0.72 | ATTCACTCATTCTTTCTTTTCATC | 87  |
|    | A-TCAC-C-TTC-TTC-TTCATC |      |                     |      |
|    | ENSG000000177058 | 0.72 | CTCACCTTCTTTTGATGC | 936  |
|    | ATCCACCTTCTTTCTTCATC |      |                     |      |
|    | ENSG000000238062 | 0.72 | ATCAACCTCTCTTCTTCTTCATC | 2188 |
|    | ATCAACCTCTCTTCTTCTTCATC |      |                     |      |
|    | ENSG000000135824 | 0.72 | AGCACCTTCTTTCTTTGATC | 2871 |
|    | ATCACCTTCTTTCTTCATC |      |                     |      |
|    | ENSG000000261371 | 0.72 | ATCACCTTCTTTCTTCATC | 1118 |
|    | ATCACCTTCTTTCTTCATC |      |                     |      |
|    | ENSG000000274333 | 0.72 | ATCACCTTCTTTCTTCATC | 6919 |
|    | ATCACCTTCTTTCTTCATC |      |                     |      |
|    | ENSG00000248966 | 0.72 | ATCACCTTCTTTCTCCCCAGTTC  
|    |                  |      | ATCACCTTCTTTCT-C-A-C   |
| 107 | ENSG00000179218 | 0.72 | CTCACCTTTCTTTCTCTTC  
|    |                  |      | ATCACCTTCTTTCT-CATC   |
| 108 | ENSG00000123243 | 0.72 | ATCACCTTTTTTTTCTCTTC  
|    |                  |      | ATCACCTTCTTTCTCATC   |
| 109 | ENSG00000104321 | 0.72 | A-CACCTTCTTTCTTCATC  
|    |                  |      | ATCACCTTCTTTCT-CATC   |
| 110 | ENSG00000100150 | 0.72 | AGTC-CCTTTCTTTCTCTTCATC  
|    |                  |      | A-TCACCTTCTTTCTCATC   |
| 111 | ENSG00000072041 | 0.72 | AGCACCATTCTTTCTTCATC  
|    |                  |      | ATCACCTTCTTTCTCATC   |
| 112 | ENSG00000230894 | 0.72 | AACACCGTTCTCTTTCTCATC  
|    |                  |      | ATCACCTTCTTTCTCATC   |
| 113 | ENSG00000086570 | 0.72 | ATCAGACCTTTGACTTTCTCATC  
|    |                  |      | ATC-A-CCTTTCTTTCTCATC   |
| 114 | ENSG00000162511 | 0.72 | ATCACCATGCTTTCTTCTCATC  
|    |                  |      | ATCACCTTCTTTCTCATC   |
| 115 | ENSG00000099977 | 0.72 | ATAATCC-TCTTTCTTCTCATC  
|    |                  |      | ATCA-CCTTTCTTTCTCATC   |
| 116 | ENSG00000270104 | 0.72 | ATCTACCTATCTTGCTTTCTCATC  
|    |                  |      | ATC-ACCT-TCTTTCT-CATC   |
| 117 | ENSG00000153822 | 0.72 | A-CACCTTTCTTTCTTCATC  
|    |                  |      | ATCACCTTCTTTCTCATC   |
| 118 | ENSG00000236449 | 0.72 | ATCACCATCTTTCTCCACTC  
|    |                  |      | ATCACCTTCTTTCTTCATC   |
| 119 | ENSG00000278903 | 0.72 | ATCACCTT-TGTGTCTTCTCATC  
|    |                  |      | ATCACCTTCTTTCTCATC   |
| 120 | ENSG00000173705 | 0.72 | ATCGCC-TCTTTCTTCTCATC  
|    |                  |      | ATCACCTTCTTTCTCATC   |
| 121 | ENSG00000172215 | 0.72 | AACACCTGCTTTCTTCTCATAC  
|    |                  |      | ATCACCT-TCTTTCTTCATC-C   |
| 122 | ENSG00000055609 | 0.72 | ATC-CCTACTTTCTTCAGTC  
|    |                  |      | ATCACCTTCTTTCTTCATC   |
| 123 | ENSG00000139174 | 0.72 | ATCTCCTTTCCATTTCTTCTCATC  
|    |                  |      | ATCTCCTTTCCATTTCTTCTCATC   |
| ID   | Ensembl Gene ID | Log2FoldChange | Sequence 1 | Sequence 2 | Length |
|------|----------------|----------------|------------|------------|--------|
| 124  | ENSG00000178607| 0.72           | ATCACC-TT-C-TTCTTCATC | ATCAACCTCTTCTTCTGT-ATC | 2866   |
| 125  | ENSG00000109163| 0.72           | ATC-ACCT-TCTTCTTCATC  | ATCA-CTTCTTTCTTTCTGATC | 939    |
| 126  | ENSG00000179912| 0.72           | ATCACCTTCTTC-ACTTCATC | ATCACTCTTCTTCTGATC    | 5117   |
| 127  | ENSG00000250220| 0.72           | ATC-ACCT-TCTTCTTCATC | ATCACTCTTCTTCTGATC    | 592    |
| 128  | ENSG00000279128| 0.72           | ATC-ACCT-TCTTCTTCATC | ATC-CTTCTTCTTCTGATC   | 685    |
| 129  | ENSG00000088726| 0.72           | ATCACCTTCTTC-TCTTCTTCATC | ATCCTCCTCTTCTTCTTCATC | 741    |
| 130  | ENSG00000047579| 0.72           | ATCACCTTCTTC-TCTTCTTCATC | ATCACCTTCTTC-TCTTCTTCATC | 2703   |
| 131  | ENSG00000183908| 0.72           | ATCAACCCTCCTTCTTCTTCATC | ATCATTCTTTCTTCTTCATC  | 2729   |
| 132  | ENSG00000035862| 0.72           | ATCACCTTCTTC-TCTTCTTCATC | ATCACCTTCTTC-TCTTCTTCATC | 1092   |
| 133  | ENSG00000133247| 0.72           | ATCACCTTCTTC-TCTTCTTCATC | ATCACCTTCTTC-TCTTCTTCATC | 2098   |
| 134  | ENSG00000154065| 0.72           | ATCACCTTCTTC-TCTTCTTCATC | ATCACCTTCTTC-TCTTCTTCATC | 1881   |
| 135  | ENSG00000254541| 0.72           | ATCACCTTCTTC-TCTTCTTCATC | ATCACCTTCTTC-TCTTCTTCATC | 184    |
| 136  | ENSG00000157765| 0.72           | ATCACCTTCTTC-TCTTCTTCATC | ATCACCTTCTTC-TCTTCTTCATC | 2227   |
| 137  | ENSG00000143669| 0.72           | ATCACCTTCTTC-TCTTCTTCATC | ATCACCTTCTTC-TCTTCTTCATC | 2947   |
| 138  | ENSG00000179841| 0.72           | ATCACCTTCTTC-TCTTCTTCATC | ATCACCTTCTTC-TCTTCTTCATC | 792    |
| 139  | ENSG000000213760| 0.72          | ATCACCTTCTTC-TCTTCTTCATC | ATCACCTTCTTC-TCTTCTTCATC | 1280   |
| 140  | ENSG00000137210| 0.72           | ATCACCTTCTTC-TCTTCTTCATC | ATCACCTTCTTC-TCTTCTTCATC | 3948   |
|       | ENSG00000065135 | 0.72 | ATCACCTTTTGTTCCTTCAGC | ATCAC-CTTCTTCTTCATC |
|-------|----------------|------|----------------------|---------------------|
| 141   | ENSG00000181847 | 0.72 | ATCTACCTTTTTCTCAGTTC | ATCAC-CTTCTTCTTCATC |
| 142   | ENSG00000250383 | 0.72 | ATCCATCTTCTTCTCCATC  | AT-CACCTTCTTCT-T-CATC|
| 143   | ENSG00000197852 | 0.72 | TTCACCCATTTCTCTCACGT  | ATCA-CC-TCTTCTTCCTC-TC |
| 144   | ENSG00000170374 | 0.72 | ATCACCTTCTCTCTTCTACC | ATCACCTTCTCTTCATC    |
| 145   | ENSG00000106344 | 0.72 | ATC-CACTCTTCTTCATTC  | ATCAC-CTTCTCTTCATC   |
| 146   | ENSG00000185811 | 0.72 | ACTCACACTTCTTCTTCATC | A-TCAC-CTTCTTC-TCTTCATC|
| 147   | ENSG00000130649 | 0.72 | AGTCATCTTCTCTTACTC   | A-TCA-CCCCTCTCTTCTTCATC|
| 148   | ENSG00000196935 | 0.72 | ATTCAGCCTTCTATCTGCTAC | A-TCA-CCCCTCTCTTCATC |
| 149   | ENSG00000260073 | 0.72 | ATCACCTTCTGCTCTCCTC  | ATCACCTTCTTCTTCATC   |
| 150   | ENSG00000249915 | 0.72 | ATGACCTTTTTTCTCATTCT | AT-CACCTTCTTCTTCATC   |
| 151   | ENSG00000170275 | 0.72 | ATGACCTTCTTCTTCTTCATC| AT-CACCTTCTTCTTCATC   |
| 152   | ENSG00000145348 | 0.72 | ATCACCTGTCTCTCCTCCTC | ATCACCTTCTTCTTCATC   |
| 153   | ENSG00000165629 | 0.72 | TTCTACCTTTTGCTCTTCATC | ATC-ACC-TTCTTCTTCATC |
| 154   | ENSG00000281357 | 0.72 | ATCACCTTCTTCTCTTCATC | ATCACCTTCTTCTTCATC   |
| 155   | ENSG00000117472 | 0.72 | ACTCA-GTTCTTCTTCATC  | A-TCACCTTCTTCTTCATC  |
| 156   | ENSG0000099341  | 0.72 | A-CACCTTCTTCTGACATC  | 1453                |
|    | ENSG00000213853 | 0.72 | ATCACCTTCTTCTTCTTCATC | AT-CACC-TTC-TTC-TTCA-TC | 5343 |
|----|----------------|------|-----------------------|--------------------------|------|
| 159| ENSG00000148600| 0.72 | ATCATCC-TCTTCTTCATC   | ATCA-CCTTCTTCTTCATC      | 3955 |
| 160| ENSG00000280548| 0.72 | ATCACCTATCTTCTCTCA-C   | ATCACCT-TCTTCTTCATC      | 2047 |
| 161| ENSG00000258754| 0.72 | ACCACCTTCGTTCCCTTCATC  | ATCACCTTC-TC-CTTCATC     | 2201 |
| 162| ENSG00000125538| 0.72 | AGCACCTTCTTCTTCATC     | ATCACCTTC-TC-CTTCATC     | 1330 |
| 163| ENSG00000235100| 0.72 | AT-ACCTTCTTACTT-ATC    | ATCACCTTCTTCTTCATC       | 463  |
| 164| ENSG00000184154| 0.72 | ATCATCC-TCTTCTTCATC    | ATCA-CCTTCTTCTTCATC      | 5558 |
| 165| ENSG00000267767| 0.72 | ATCCACATTCTTCTCTCATC   | AT-CAC-CTCTTCTTCATC      | 2191 |
| 166| ENSG00000276077| 0.72 | ATCACCTT-TTGTTCATC     | ATCACCTTCCTTCTTCATC      | 1680 |
| 167| ENSG00000280156| 0.72 | TTCACCTTCTTCTACCTC     | ATCACCTTCTTCTTCATC       | 78199|
| 168| ENSG00000175697| 0.72 | A-CAGCTTCTTCTATCATC    | ATCACCTTCTTCTTCATC       | 5324 |
| 169| ENSG000000259485| 0.72 | ATCACCTGCCTTCTGCAGCTAC | ATCACCT-TCTCTCTACTC      | 363  |
| 170| ENSG000000259485| 0.72 | ATCACCTTGCTTCTGCAGCTAC | ATCACCT-TCTCTCTACTC      | 363  |
| 171| ENSG00000129467| 0.72 | ATCACCTTCTCCTCTCCTTC   | ATCACCTTCT-CT-TCTTCATC   | 4859 |
| 172| ENSG00000129467| 0.72 | ATCA-CTTCCATTCCT-ATC   | ATCACCTTC-TCCTTCATC      | 2118 |
| 173| ENSG00000158528| 0.72 | ATC-CCTGCTTCTTCTCTCAATC | ATCACCT-TCTCTTCATC-ATC   | 4429 |
| 174| ENSG000000197157| 0.72 | ATCACCTTTGGCTTGAGCTAGTT | ATCACCTTCTT-CTT-CA-TC    | 3279 |
| 175| ENSG000000139219| 0.72 | ATTTCTACCTTCTCTTCATC   | A-TC-ACCTTCTTCTTCATC     | 1599 |

23
| Gene ID       | Score | Region Sequence                       | Score |
|--------------|-------|---------------------------------------|-------|
| ENSG00000153395 | 0.72  | ATCACCTGTTTGCTTTCCCATC                | 1211  |
| ENSG00000140943 | 0.72  | ATCAACCTCTCTGCTTTTCATC                | 8498  |
| ENSG00000213676 | 0.72  | TTCACCTCTCTCTTCTTCATC                 | 3305  |
| ENSG00000087460 | 0.72  | ATC-CCTTCTCTCTTGC-TC                 | 155   |
| ENSG00000204160 | 0.72  | ATC-CTCTTCTCTCTCTGTC                 | 552   |
| ENSG00000279006 | 0.72  | ATCACCTTCTCTC-TCATC                   | 1031  |
| ENSG000000185272 | 0.72  | ATTA-CTTCTCTTCTTCATC                 | 720   |
| ENSG00000134444 | 0.72  | ATGCCACCTTCTC-TCATC                   | 1818  |
| ENSG00000196083 | 0.72  | ATCACCTGCTCTCAGC                     | 206   |
| ENSG00000149972 | 0.72  | ATCAACCTTCTTACACTC                   | 8455  |
| ENSG00000161999 | 0.72  | ACTCACCTTCTTCTTCATC                  | 2216  |
| ENSG00000144596 | 0.72  | ATTC-CCTGCTCTTCATC                   | 3428  |
| ENSG00000134243 | 0.72  | ACTCACCTTCTTCTTCATC                  | 7525  |
| ENSG0000007312  | 0.72  | ATCATCC-TCTTCATC                     | 1409  |
| ENSG00000235750 | 0.72  | ATCACACCTCTTCTTTCATC                 | 877   |
| ENSG00000112394 | 0.72  | ATCACCTTCTTTACATC                    | 6666  |
| ENSG00000143294 | 0.72  | ATCACATATCTTATTTCATC                 | 3644  |

24
| ID            | ENSG Identifier | Similarity Score | Sequence 1                          | Sequence 2                          | Sequence 3                          |
|--------------|----------------|------------------|-------------------------------------|-------------------------------------|-------------------------------------|
| 194          | ENSG00000105202| 0.72             | ATCACCTATCTTCTCTCA-C                | ATCACCT-TCTT-CT-TCATC              | 2047                                |
| 195          | ENSG00000115183| 0.72             | AGCACCGTTCTTC-TCATC                | ATCACCT-TCTTCTCCTCATC              | 2647                                |
| 196          | ENSG00000188283| 0.72             | ATCATTTCTTTTCTTTCAGCTC             | ATCA-CC-TCCTTCTTCA-TCT            | 684                                 |
| 197          | ENSG00000226312| 0.72             | ATTTACCAACCTTCTTCTTCCCATC          | AT-TC-ACCTTCTTTCTTTCATC           | 886                                 |
| 198          | ENSG00000091536| 0.72             | CTCCTCCTTCTTCTTCTCATC              | AT-ACCTT-TCTTCTTCTCTCATC          | 17565                               |
| 199          | ENSG00000154265| 0.72             | AT-ACCTT-TCTTCTTCTCATC             | ATCACCCTTCTT-CTTCTATC            | 10718                               |
| 200          | ENSG00000147118| 0.72             | ATTCAGCCTTTTCTTCTTCTATC            | AT-TCACCTTCTTTCTTTCATC            | 2376                                |
| 201          | ENSG00000232021| 0.72             | ATCACAACCTTCTTCTTACATC             | AT-AC-CTT-CTTCTTCTTCTATC          | 402                                 |
| 202          | ENSG00000169087| 0.72             | ATCACACTTCTTCTTAC                 | ATCAC-CTTCTTCTTCTTCTATC           | 3353                                |
| 203          | ENSG00000154237| 0.72             | A-CAGCCTTCTTTCTA-TC               | ATCA-CCTTCTTCTTCTTCTATC           | 5809                                |
| 204          | ENSG00000132182| 0.72             | ATC-CCTCCTTTCTTACAT                | ATCACCCTTCTTCTTCTTATC             | 645                                 |
| 205          | ENSG00000260880| 0.72             | ATC-CCTCATTCTTCAGC                | ATCACCCTTCTTCTTCTATC              | 640                                 |
| 206          | ENSG00000153347| 0.72             | AGTCATCTTCTTTCTTTCATC             | ATCACCCTTCTTCTTCTTCATC           | 2332                                |
| 207          | ENSG00000044446| 0.72             | ATCAGCCTTTTCTTCCACC               | ATCA-CCTTTCTTCTTATC              | 6802                                |
| 208          | ENSG00000177663| 0.72             | ATCACCTTAATT-TCACAT                | ATCACCTTCTTCTTCTATC              | 8539                                |
| 209          | ENSG00000144029| 0.72             | ATCACCTGCTTTCTGCTGCAGTC           | ATCACCCTTCTTCTTCTATC              | 3912                                |
| 210          | ENSG00000131591| 0.72             | CTCACCTGCTTTCTTCTCATC             | ATCACCTTCTTCTTCTATC              | 1166                                |
| ID       | ENSG Identifier    | Score | Sequence 1                  | Sequence 2                      |
|----------|--------------------|-------|-----------------------------|---------------------------------|
| 211      | ENSG00000116641    | 0.72  | ATCACCTTCTGTTTCTTCATC       | ATCACCTTCTTCTTCATC              |
| 212      | ENSG00000282961    | 0.72  | ATCACCATGTTCTATC-TCTATC     | ATCCAC-T-TCT-TCTTCTTCATC        |
| 213      | ENSG00000277991    | 0.72  | ATCACCTTTGTTCATC            | ATCACCTTTCTTCTTCATC             |
| 214      | ENSG00000231049    | 0.72  | ATCTCTCTTTTCTTCCTCACTC      | A-TCACCTTTCTTCTTCATC            |
| 215      | ENSG00000130038    | 0.72  | AGTCA-CTTCTTCTTCAGC         | A-TCACCTTTTTCTTCATC             |
| 216      | ENSG00000259134    | 0.72  | ATCAGCTATCTTCTATGCTATC      | ATCCAC-T-TCTTCTTCATC            |
| 217      | ENSG00000253894    | 0.72  | ATC-CCTTCATTCTTCATC         | ATCACCTTTCTTCTTCATC             |
| 218      | ENSG00000174255    | 0.72  | ATCA-C-TCTGTCTTCATC         | ATCACCTTCTTCTTCATC              |
| 219      | ENSG00000088387    | 0.72  | ATCATAGCCTGTGCCTTCATC       | ATC-A-CCTTCTTCTTCATC            |
| 220      | ENSG00000162804    | 0.72  | AT-ACCTTTCTATGCTGCTATC      | ATCACCTTCTTCTTCATC              |
| 221      | ENSG00000080815    | 0.72  | ATCACCTTTCTTCATCAC          | ATCACCTTTCTTCTTCATC             |
| 222      | ENSG00000109182    | 0.72  | ATCACCTTTCTTCGTCACC         | ATCACCTTTCTTCTTCATC             |
| 223      | ENSG00000143322    | 0.72  | AGTGCA-CTTACCTTCTTCATC      | A-T-CACCTTTCTTCTTCATC           |
| 224      | ENSG00000152705    | 0.72  | ATTCACCATCATCTTCATC         | A-TACACCTTTCTTCTTCATC           |
| 225      | ENSG00000242715    | 0.72  | ATTCCTACCTT-TCTTCTTCATC     | A-T-C-ACCTTTCTTCTTCATC          |
| 226      | ENSG00000186847    | 0.72  | AGCACCTTC-TCTTCACTC         | ATCACCTTTCTTCTTCATC             |
| 227      | ENSG00000035499    | 0.72  | ATCACCTTTTCTTCTTCTTCATC     | ATCACCC-T-TCT-TCTTCTTCATC       |
| 228      | ENSG00000245937    | 0.72  | ATCACCTTCCCACCTTCATC        | ATCACCTTCTTCTTCATC              |
| Gene ID | Score | Sequence 1 | Sequence 2 |
|--------|-------|------------|------------|
| ENSG00000227110 | 0.72 | ATCACCCGTTCTTCTGCACTC | ATCA-CTTCTTTCTTCATC |
| ENSG00000165181 | 0.72 | ATCACCTTCTCTTCAATA | ATCACCTTCTTTCTTCATC |
| ENSG00000188984 | 0.72 | ATGCCACCTTCTCTCTTCATAC | AT-ACCTTCTTCTTCATC |
| ENSG00000126777 | 0.72 | ATACCTTCTCTTCTTCATAC | ATCACCTTCTTTCTTCATC |
| ENSG00000155886 | 0.72 | ATGCCACCTTCTCTCTTCATAC | ATACCTTCTTTCTTCATC |
| ENSG00000254465 | 0.72 | ACTCACACTTCTTTTCTATCATC | A-TCAC-CTTC-CTTCATC |
| ENSG00000175841 | 0.72 | ATCACCTTTCTCTTTCATAC | ATCACCTTCTTTCTTCATC |
| ENSG00000112096 | 0.72 | ATCACCTTCTTCTTCATC | ATCACCTTCTTTCTTCATC |
| ENSG00000157680 | 0.72 | ATCA-CTTCATT-TTCATC | ATCACCTTCTTTCTTCATC |
| ENSG00000081014 | 0.72 | ATCTCTCCTTTCTTCATC | A-TCACCTTCTTTCTTCATC |
| ENSG00000169679 | 0.72 | ATCA-TCTTTCTTTCTTCATC | ATCA-C-CTTCTTTCTTCATC |
| ENSG00000101096 | 0.72 | ATCACACTTTTCTTTGC-TC | ATCACCTTCTTTCTTCATC |
| ENSG00000280145 | 0.72 | ATCACCTT-TGTTCATC | ATCACCTTCTTTCTTCATC |
246  ENSG00000111481  0.72  AT-ACCTCTTCTCTCTTCATC
           ATCA-CTTCTCTTTCATC

247  ENSG00000152061  0.72  ATCACCTGTCTTCTGCGTC
           ATCACCTTCTTCTTCATC

248  ENSG00000154783  0.72  ACTCCCTTTTCTTCTTCTC
           A-TCACCTTTTCTTTCATC

249  ENSG00000235522  0.72  AT-AACTTCTTCTGTCATC
           ATCACCTTCTTCTTCATC

250  ENSG00000174844  0.72  AT-A-CTTCTTCTTCTTCATC
           ATCACCTTCTTCTTCATC

251  ENSG00000153046  0.72  ATCACCTGCTTTCTTCTTCATC
           ATCACCTTCTTCTTTCATC

252  ENSG00000234072  0.72  ATCACCTGCCCTTCTTCAGTC
           ATCACCTTCTTCTTCATC

253  ENSG00000136854  0.72  ATCATCCTTC-TCTACATC
           ATCA-CCTTCTCTTCATC

254  ENSG00000153551  0.72  ATC-CCTTCTACTTCACTC
           ATCACCTTCTTCTTCATC

255  ENSG00000136449  0.72  ATTTACCTTTCTTCTTCAAGTC
           A-TCACCTTTTCTTTCATC

256  ENSG00000231918  0.72  ATC-CCTTCTTTCTTCTTCATC
           ATCACCTTCTTCTTCATC

257  ENSG00000164587  0.72  ACTCCAGCCTTCTCTTCTTCATC
           A-T-CA-CCTTCTTCTTCATC

258  ENSG00000127507  0.72  CTCACC-TCTTCTTCCATC
           ATCACCTTCTTCTTCCATC

259  ENSG00000253355  0.72  ATCCACCTTGCT-TTCCATC
           AT-CACCTTCTTCTTCCATC

260  ENSG00000174197  0.72  ATCTCCTGTCTTCTTTCATC
           ATCACCT-CTTCTTCATC

261  ENSG00000116095  0.72  ATC-CC-TCTATCTCCTCATC
           ATCACCTTCTTCTTCATC

262  ENSG00000179604  0.72  AACAGCCCTTTCTTCTTCATC
           ATCA-CCTTCTTCTTCCATC

263  ENSG00000024862  0.72  AT-ACCCCTTCTTCTTCAGGTC
| Gene ID          | Value   | Sequence 1                      | Sequence 2                      | Length |
|-----------------|---------|--------------------------------|---------------------------------|--------|
| ENSG00000198569 | 0.72    | ATCA-CCTTCTTCTTCATC            | TC                             | 1710   |
| ENSG00000144792 | 0.72    | ATCACACTTCTCTCTAATC            | TC                             | 3422   |
| ENSG00000091262 | 0.72    | ATCTACC-TCCTCTCATC            | ATC-ACCTCTCTCTCATC             | 236    |
| ENSG00000149948 | 0.72    | ATCACCTTCTT-TGATATC           | TC                             | 12587  |
| ENSG00000170180 | 0.72    | TTCAACCTTCTTCTCTGTC           | ATCACCTCTTTCTCTCA-TC           | 1575   |
| ENSG00000160588 | 0.72    | ATCATTTTCTTTCTTCTCATC         | ATCACCTTCTTCTTCTCATC           | 3133   |
| ENSG00000171827 | 0.72    | ATCACCCTTCTTCTTCATC           | ATCACCTTCTTCTTCTCATC           | 1923   |
| ENSG00000171827 | 0.72    | AT-ACCTTCTTCTTGCTAT           | ATCACCTTCTTCTCTCAT             | 9217   |
| ENSG00000138778 | 0.72    | GTCAACCTACGTCTTCTTCTCATC      | ATCACCTTCTTCTTCTCAT           | 1413   |
| ENSG00000198246 | 0.72    | ATCATCTTTTTTCTTCATC           | ATCACCTTCTTCTTCTCAT            | 474    |
| ENSG00000224885 | 0.72    | ATCCACCTTCTTCTTCTATGC         | ATACCTTCTTCTTCTCATC           | 2328   |
| ENSG00000127920 | 0.72    | ATCACCTGCTTCTTTCTGAT         | ATCACCTTCTTCTTCTC             | 2544   |
| ENSG00000102053 | 0.72    | ACCACCTATCTCTTCTCATC          | ATCACCTTCTTCTTCTCAT            | 91     |
| ENSG00000154122 | 0.72    | ATCACCT-TCATTCTTCTTATTC      | ATCACCTTCTTCTTCTCATC          | 10325  |
| ENSG00000115419 | 0.72    | ATTCACACTTCTTCTTCTTC         | A-TCAC-CTTCTTCTTCTCAT          | 5668   |
| ID          | ENSG ID           | Score | Matched Sequence                      | Length |
|-------------|-------------------|-------|---------------------------------------|--------|
| 281         | ENSG00000105722   | 0.72  | ATC-CTTCTTTCTTTCTTCTTC               | 1853   |
| 282         | ENSG00000135636   | 0.72  | ATCATCC-TCTTTCTCATCATC               | 8006   |
| 283         | ENSG00000267784   | 0.72  | ATCA-CCTTTCTTCTTC-TC                | 2064   |
| 284         | ENSG00000099958   | 0.72  | ATTCAGCTTCTTCTTCAAC                 | 536    |
| 285         | ENSG00000080644   | 0.72  | AT-ACCTTGCTTTCTTACACC               | 4287   |
| 286         | ENSG00000135966   | 0.72  | ATACACCCTC-TCTTCTCATC               | 4443   |
| 287         | ENSG00000147874   | 0.72  | ATCACCTTTCTTTCTTTATGC               | 7807   |
| 288         | ENSG00000215182   | 0.72  | CTCAACCTTTCTTCTCATCATC              | 1637   |
| 289         | ENSG00000140092   | 0.72  | ATCGGCC-TCTTTCTTCTTC                | 1294   |
| 290         | ENSG00000187555   | 0.72  | ATCACTTTATTTTTTCTCATC               | 7155   |
| 291         | ENSG00000103335   | 0.72  | ATCATCC-TCTTTCTTCTCATC              | 9423   |
| 292         | ENSG00000170836   | 0.72  | ATTC-CCTTTCTTTCTTCTTCATC            | 3027   |
| 293         | ENSG00000188649   | 0.72  | AT-ACCTTCCTTTCTTCTTACC              | 70     |
| 294         | ENSG00000141837   | 0.72  | ATCGGCATGCTTTCTTCTCATC              | 33937  |
| 295         | ENSG00000205730   | 0.72  | ATCACCTTTCTCGCTTTCA-C               | 2027   |
| 296         | ENSG00000184277   | 0.72  | AGTCACCTTTCT-TTC-TC                | 8527   |
| 297         | ENSG00000204697   | 0.72  | ATCA-CCTTTCTTCTTCTCATC              | 636    |
| 298         | ENSG00000112137   | 0.72  | ACCACCTT-TTCTTTCCATC               | 6578   |
|    | ENSG00000169083 | 0.72 | ATCACATGCITTC-TCTTCATC | ATCAC-CTTCTTCTTCATC | 4499 |
|---|----------------|------|------------------------|---------------------|------|
| 300 | ENSG00000198093 | 0.72 | ATCTTCCCAITTCCTTCCATC | ATC-ACC-CTTCTTCTTCATC | 1594 |
| 301 | ENSG00000280224 | 0.72 | ATCACCATCATCATCATC | ATCACCTTCTTCTCTCATC | 883  |
| 302 | ENSG00000099998 | 0.72 | ATTCACCTTT-TTCATTCATTC | ATC-AC-CTTCTTCTTCATC | 1645 |
| 303 | ENSG00000172269 | 0.72 | ATCATCTCTTTCTGCTTCATC | ATCA-C-CTTCTTCTTCATC | 2459 |
| 304 | ENSG00000275079 | 0.72 | ATCACCTGTTTCTGCTTCATC | ATCACCTTCTTCTTCATC | 809  |
| 305 | ENSG00000171105 | 0.72 | ATCACTTTTTCTCCTGCTTCATC | ATCACCTTCTTCTTCATC | 9192 |
| 306 | ENSG00000131732 | 0.72 | ATCACTCTTT-TCTTCTTCATC | ATCACCTTCTTCTTCATC | 1923 |
| 307 | ENSG00000171094 | 0.72 | ATCACCGTTTTTTCTTCCTCCATC | ATCAC-CTTCTTCTTCATC | 2844 |
| 308 | ENSG00000072657 | 0.72 | ATCTTTCTTTACTCTTCTTCATC | ATC-ACC-CTTCTTCTTCATC | 3711 |
| 309 | ENSG00000248905 | 0.72 | ATCACGTTT TTCTTCTTCATC | ATCAC-CTTCTTCTTCATC | 169  |
| 310 | ENSG00000162337 | 0.72 | ATCATCTCTCTCTCTCGTC | ATCA-CCT-TCTTCTTCATC | 4690 |
| 311 | ENSG00000108387 | 0.72 | ATGACCTTCTCTTTCTCTTCATC | AT-CACCTTCTTCTTCATC | 2142 |
| 312 | ENSG00000088808 | 0.72 | ATGATCTTTCTTCCCTTCATC | ATCA-CTTCTTCTTCATC | 8092 |
| 313 | ENSG00000123643 | 0.72 | ATCA-CTTCTTCTTCTGTA | ATCACCTTCTTCTTCATC | 3610 |
| 314 | ENSG00000117408 | 0.72 | ATCCACCCCTTT-TCTTCTTCATC | AT-CA-CCTTCTTCTTCATC | 4760 |
| 315 | ENSG00000165970 | 0.72 | ATCACCTCTCTTCTTACAAC | ATCAC-CT-TCTTCTTCATC | 1729 |
|  Number | Gene Identifier | Score | Sequence 1 | Sequence 2 |
|---------|----------------|-------|------------|------------|
| 316 ENSG00000205060 | 0.72 | ATCACCTTTTCCTTCTTCAGC ATCAC-TC-TTCTTCTCATC | 2028 |
| 317 ENSG00000172671 | 0.72 | A-CAGCCTTTCTCTTCTTCATC ATCA-CCTTCTTCTTCTCATC | 4096 |
| 318 ENSG00000274286 | 0.72 | ATCACCTTCCTCATTC-TC ATCACCTCTTTC-TTCATC | 151 |
| 319 ENSG00000101337 | 0.72 | ATCGAGC-TCTTCTTCATC ATC-ACCTTCTTCTTCTCATC | 2186 |
| 320 ENSG00000105409 | 0.72 | ATCACCTTCTT-CTCCACC ATCACCTTCTTCTTCTCATC | 1654 |
| 321 ENSG00000029363 | 0.72 | ATCACCTTCTTCTTCCAGTTC ATCACCTTCTTCTC-TCA-TC | 4555 |
| 322 ENSG00000160716 | 0.72 | ATCCATCCTT-TTGCTTCTCATC AT-CA-CCTTCTTCTTCTCATC | 1909 |
| 323 ENSG00000113073 | 0.72 | ATCAGCC-TCATCTTCTCATC ATCA-CCTTCTTCTTCTCATC | 1638 |
| 324 ENSG00000154727 | 0.72 | ATGCACCTTCTCTTCTTATT AT-CA-CCTTCTTCTTCTCATC | 3569 |
| 325 ENSG00000080189 | 0.72 | AT-ACCTTCTCATCTCACTC ATCACCTTCTTCTTCTCA-TC | 5934 |
| 326 ENSG00000242550 | 0.72 | ATCACCCCATTTCTCTTCTCTCATC ATCA-CC-TCTCTTCTTCTCATC | 1755 |
| 327 ENSG00000103222 | 0.72 | AGTCACCTT-TTCTTCCATC AT-TCACCTTCTTCTTCTCATC | 3333 |
| 328 ENSG00000146005 | 0.72 | ATC-CACTTCTT-TCATC ATCAC-CTTCTTCTTCTCATC | 3102 |
| 329 ENSG00000001626 | 0.72 | CTCAGCCTTCTCTTCTTCTCATC ATCA-CCTTCTTCTTCTCATC | 1494 |
| 330 ENSG00000182021 | 0.72 | CTACCTTCTTTGCTTCTTCTCATC ATCACCTTCTTCTTCTCATC | 1782 |
| 331 ENSG00000197893 | 0.72 | A-CACTTCTTCTTCTC-TC ATCACCTTCTTCTTCTCATC | 1988 |
| 332 ENSG00000177932 | 0.72 | ATCATCCTTCTT-TTCTCATC ATCA-CCTTCTTCTTCTCATC | 1074 |
| 333 ENSG00000123007 | 0.72 | ATCACCC-TCTTTTCTTCTCATC | 1520 |
| Gene ID     | Score | Sequence 1                   | Sequence 2                   | Length |
|------------|-------|------------------------------|------------------------------|--------|
| ENSG00000212743 | 0.72  | ATCACCTTCTTCTTCATC           | ATCACCTTCTTCTTCATC           | 37     |
| ENSG00000196923 | 0.72  | ACCACCTTC-TCTTCCATC         | A-TCACCTTCTTCTTCATC          | 408    |
| ENSG00000179630 | 0.72  | ATCA-CTTCTTCTTGATC          | ATCACCTTCTTCTTCATC          | 1313   |
| ENSG00000111012 | 0.72  | ATCACCTT-TTCATTTCATC        | ATCACCTT-TTCATTTCATC        | 2656   |
| ENSG00000132915 | 0.72  | ATCAACCTTCCTTCCATTCA-TC    | ATCAACCTTCCTTCCATTCA-TC    | 6001   |
| ENSG00000185958 | 0.72  | ATCACCAATCTTCTCTCACTAC     | ATCACCAATCTTCTCTCACTAC     | 6273   |
| ENSG00000258943 | 0.72  | ATCAAACCTTTATTTCTGTACATC   | ATCA-ACTTCTTCTTCTCATC       | 827    |
| ENSG00000123066 | 0.72  | ACTCAACCTTCTGTTTACTTCATC   | A-TCACCTTCTTCTTCATC         | 13350  |
| ENSG00000073803 | 0.72  | ATCTATCCATT-TTCTTCATC      | ATC-A-CC-TTCTTCTTCATC       | 5564   |
| ENSG00000100395 | 0.72  | AT-ACCCTTC-TCTTCATC        | AT-ACCCTTC-TCTTCATC         | 4962   |
| ENSG00000100379 | 0.72  | ATCACCTGCTTGGCTTACCTC      | ATCACCTTCT-CTTCATC          | 1323   |
| ENSG00000118777 | 0.72  | ACCACCTCCTTCTGTCATC        | ATCACCTTCTTCTTCATC          | 2018   |
| ENSG00000100379 | 0.72  | ATCCACC-TCTTCTTCTTCATC     | AT-CACCTTCTTCTTCATC         | 1613   |
| ENSG00000066322 | 0.72  | ATGGCACCATCCTTCTTCATGTC    | AT-GACACCTTCTTCTTCAT-C      | 1864   |
| ENSG00000134909 | 0.72  | ATCGACCCCTCTTCTTCATC       | AT-CCACCTTCTTCTTCATC        | 5069   |
| ENSG00000061676 | 0.72  | ATCACCTTCTTCTCAGTAAC       | ATCACCTTCTTCTTCATC          | 20776  |
| ENSG00000280162 | 0.72  | ATC-CCTTCTTGTCTGTACATC     | ATCACCTTCT-TCT-TCTCATC      | 744    |
| Gene ID       | Score | Sequence 1                          | Sequence 2                  |
|--------------|-------|------------------------------------|----------------------------|
| ENSG00000101204 | 0.72  | ATCACCTATGACTCTGATACTC             | ATCACCT-T-CTTCTTCATC       |
| ENSG00000116299 | 0.72  | ATTACCTTCTTTCTTCTTC               | ATCACCTTCTTC-TCTCA-TCTC    |
| ENSG00000165240 | 0.72  | ATCACTCTTGTTATCTCATC              | ATCAC-CTTCT-TCTTCA-TC      |
| ENSG00000221813 | 0.72  | ATCACCATGTCTACTTCATC              | ATCACT-T-TCTTCTTC-ATC      |
| ENSG00000279800 | 0.72  | ATCACCTTCTTCTCCAGTTC              | ATCACCTTCTTCTC-TCA-TCTC    |
| ENSG00000106692 | 0.72  | ATCACTTCTCTCTTCATC                | ATCACCTTCTTCTTCA-TCTC      |
| ENSG00000153317 | 0.72  | CTCACCTTCTTCTTCATC                | ATCACCTTCTTCTTCA-TCTC      |
| ENSG00000114251 | 0.72  | ATCCACCTTCTTTCTTCA-C              | AT-ACCTTCTTTCTTCATC        |
| ENSG00000130939 | 0.72  | ATCATCTTCTTTCTATC                 | ATCACC-T-TCTTTCTCATC       |
| ENSG00000146278 | 0.72  | ATCCACCTTCTTCTTCATC               | A-TACCTTCTTCTTCATC         |
| ENSG00000116871 | 0.72  | ATCATCTTCTTCTGATCCATC             | ATCACCTTCTTCTTCATC         |
| ENSG00000153750 | 0.72  | ATCTAGACCTCCTTCTTCATC             | ATC---ACCTTCTTCTTCATC      |
| ENSG00000172493 | 0.72  | ATCTAGACCTCCTTCTTCATC             | ATC---ACCTTCTTCTTCATC      |
| ENSG00000174130 | 0.72  | ATCTAGACCTCCTTCTTCATC             | ATC---ACCTTCTTCTTCATC      |
| ENSG00000160753 | 0.72  | ATCTAGACCTCCTTCTTCATC             | ATC---ACCTTCTTCTTCATC      |
| ENSG00000173331 | 0.72  | ATCTAGACCTCCTTCTTCATC             | ATC---ACCTTCTTCTTCATC      |
| ENSG00000230040 | 0.72  | ATCTAGACCTCCTTCTTCATC             | ATC---ACCTTCTTCTTCATC      |
|   | ENSG00000118113 | 0.72 | ATCACCTCTCATCTTACCACC  
|   |               |      | ATCACCTTCTCTCCATC   |
|---|----------------|------|----------------------
|   | ENSG00000159842 | 0.72 | ATCACCTTCTCTTCCTGCTC  
|   |               |      | ATCACCTTCTCTTCATC   |
|   | ENSG00000164106 | 0.72 | ATTCATCTTTTCCTCATC  
|   |               |      | A-TACACTTCTCTTCATC   |
|   | ENSG00000267586 | 0.72 | ATCACCTCTTCTCTTCATC  
|   |               |      | ATCACCTTCTCTTCATC   |
|   | ENSG00000185532 | 0.72 | A-CACCTTCTTATATTTATC  
|   |               |      | ATCACCTTCTCTTCATC   |
|   | ENSG00000215580 | 0.72 | ATCACCTTGCTATCTTCAGTC  
|   |               |      | ATCACCTCTCTCTTCATC   |
|   | ENSG00000085999 | 0.72 | ATCACCTTCGTCTTCCACC  
|   |               |      | ATCACCTTCTCTTCATC   |
|   | ENSG00000105339 | 0.72 | ATCACCTTCTCTTCATC  
|   |               |      | ATCACCTTCTCTTCATC   |
|   | ENSG00000267665 | 0.72 | ATCACCTTTCTTCTTCATC  
|   |               |      | ATCACCTTTCTTCTTCATC   |
|   | ENSG00000231367 | 0.72 | ATCACCTTTCTTCTTCATC  
|   |               |      | ATCACCTTTCTTCTTCATC   |

Table 2: Potential target RNAs for svRNA1.
| Gene ID     | Score | Matched Sequence | Loci  |
|------------|-------|------------------|-------|
| ENSG00000196628 | 0.86  | AGATCTTCTCTTTGCTACTTCTTC | 10282 |
| ENSG00000183496  | 0.82  | TATCTTCTCCTGCTCTTCTTC | 2447  |
| ENSG00000102290  | 0.77  | AATTCCTTCTTCCTCTTCTTC | 1349  |
| ENSG00000174175  | 0.77  | AATCATTCATTCCTTCTCTTCTTC | 2614  |
| ENSG00000186104  | 0.77  | ATATCTTGCTCTGCTCTTCTTC | 4378  |
| ENSG00000255767  | 0.77  | AATCTTGCTCTTCTCTTCTTC | 355   |
| ENSG00000137648  | 0.73  | ACTTCTTCTTCCTTCTCTTCTTC | 9396  |
| ENSG00000151532  | 0.73  | AATCTTTCTCTCTTCCTTC | 3316  |
| ENSG00000235824  | 0.73  | ATCTTCTTCTTCTTCTTCTTCTTC | 43    |
| ENSG00000145012  | 0.73  | AATCGTTACTATCCTTTGACCTTTTCTTC | 21971 |
| ENSG00000169484  | 0.73  | AATCTTCTTCTTTGACTT-TAC | 497   |
| ENSG00000154262  | 0.73  | AAT-TCCTCTCTCTTCTTC | 6729  |
| ENSG00000106588  | 0.73  | AAT-TCTCTCTCTCTTTCTTC | 2481  |
| ENSG00000184068  | 0.73  | AATCTTCTGCTCTTTTCTTTCTTC | 76    |
| ENSG00000185052  | 0.73  | AGAGCTTCTCTGCTCTCTTCTTC | 2676  |
| ENSG00000143669  | 0.73  | AACTCATTTCTCTTCTCTCTTC | 15026 |
| ENSG00000189046  | 0.73  | AAT-TCTTTCTTCTCTTTCTTC | 1801  |
| ENSG00000230461  | 0.73  | AATCTTTCTTCTTGCTCTTCTTC | 2627  |
AATCTTCTTCTTGCTCTTCTTC

19 ENSG00000185745 0.73 AATC-TCTCTTCTCTCTCGTCCTCTCTTC
   AATCTCTCT-TCT-TGC-CTTTCTCTTC

20 ENSG00000126016 0.73 AAT-TTCTTGCTTTGCTCTCCTCCTTC
   AATCTTCTT-TCT-TGTGCTCTCTTTCTTC

21 ENSG00000144840 0.73 AATCTTCTT-TCTCTCTCCTTC
   AATCTTCTT-TCTCTCTCCTTC

22 ENSG00000170871 0.73 AAGTCTGTTCTTACTCTAGCTCTTCTTC
   AA-TTCTCTTCTCITGCTCTTCTTC

23 ENSG00000188655 0.73 AATC-TCTCTCCTCTCTCTCTTCATTTCTTC
   AATCTTCTTCTCTCTCTCTTCTTC

24 ENSG00000113645 0.73 AATCTCTT-TATCTGCTCTTCTCCTTCTTC
   AATCTTCTT-TATCTGCTCTTCTCCTTCTTC

25 ENSG00000226419 0.73 AA-CTCTGTTCTTGCTCTTCTTC
   AATCTTCTT-TATCTGCTCTTCTCCTTCTTC

26 ENSG00000168314 0.73 AA-CTCTGTTCTTGCTCTTCTTC
   AATCTTCTT-TATCTGCTCTTCTCCTTCTTC

27 ENSG00000221866 0.73 AATATT-TTCTTGCTCTCTTTCTTC
   AATCTTCTTCTTCTCTCTCTCTTCTTC

28 ENSG00000033867 0.73 ATCTCTTCTTTTGCTCTTTCTTCTTC
   A-ATCTTCTTTGCTCTTTCTTCTTC

29 ENSG000000152763 0.73 CTCTTTCTTCTTTCTTCTTCTTCTTC
   AATCTTCTTCTTCTTCTTCTTCTTCTTC

30 ENSG00000114473 0.73 ATTCATTCTTCTTGCTCTTCTTC
   AATCTTCTTCTTCTTCTTCTTCTTCTTC

31 ENSG00000181852 0.73 AGATCTCTTTCTTTGCTTGCTT-TTC
   A-ATCTT-C-TTCTTGTGCTCTTTCTTC

32 ENSG00000117174 0.73 AATCTCTACTCTTTCTTTGCTTGCTT-TTC
   AATCTTCTT-TC-TTGCTCCTTTCTTC

33 ENSG00000247271 0.73 AATCTACCTTCTTCTTTCTTCTTC
   AATCTTCTTCTTCTTCTTCTTCTTCTTC

34 ENSG00000070985 0.73 AGCTCTTCTTCTTTCTTCTTCTTCTTCTTC
   A-ATCTTCTTCTTCTTCTTCTTCTTCTTC

35 ENSG00000133641 0.73 AATATTCTTCTTCTTCTTCTTCTTCTTC
   AATCTTCTTCTTCTTCTTCTTCTTCTTC

37
|   | ENSG*00000118946 |   | ACTCTGTCCTCCTCCTCCTGCTCTTCTTCATC |   |
|---|----------------|---|----------------------------------|---|
|   | ENSG*00000181222 |   | AATC-TC-TCTTGCTTTTCTTC          |   |
|   | ENSG*00000141503 |   | AACCTCTCTGCTGCTTTTCCTTC       |   |
|   | ENSG*00000089818 |   | AATCTTCTCTTGCTCTTTCTTCTTC |   |
|   | ENSG*00000214029 |   | ACATCTTCTCTGGCTTTTCTTC       |   |
|   | ENSG*00000145362 |   | AATCTTCTTCTGCTCTTTCTTC|   |
|   | ENSG*00000235823 |   | AATCTTCTTCTCTGCTCTTTCTTC |   |
|   | ENSG*00000175224 |   | TATCTTCTTC-TCTCTTCTTC          |   |
|   | ENSG*00000151617 |   | AATC-TCTTCTCTGCTCTTCTTC        |   |
|   | ENSG*00000066427 |   | AATCTTTCTTTCTCTCCTTCCTTC       |   |
|   | ENSG*00000143437 |   | AATCTTATCTTTTCTCTCTTCCTTC      |   |
|   | ENSG*00000226688 |   | AATCTCTCTCTCTCTCCCTCCTTC       |   |
|   | ENSG*00000156738 |   | AATCTTGTCTCTGCTCTTCCTTC       |   |
|   | ENSG*00000049618 |   | AATCTTCTACTCT-CT-TCTCTTC       |   |

Table 3: Potential target RNAs for svRNA2.
| Gene ID          | Score | Matched Sequence                        | Loci |
|------------------|-------|----------------------------------------|------|
| ENSG00000166908  | 0.84  | ACGCTTCTGTGCCCCAGTTCCCT                | 759  |
|                  |       | A-GCTTCTGGCCCCAGTTCCCT                |      |
| ENSG00000174243  | 0.84  | AGCTTCTTGCCCCAGTTCCCT                | 4934 |
|                  |       | AGCTTCTGGCCCCAGTTCCCT                |      |
| ENSG00000280011  | 0.84  | AGCATCTGGCCCCAGTTCCCT                | 7953 |
|                  |       | AGCTTCTGGCCCCAGTTCCCT                |      |
| ENSG000000145685 | 0.84  | AG-TTCTGGCCCCAGTTCCCT                | 5046 |
|                  |       | AGCTTCTGGCCCCAGTTCCCT                |      |
| ENSG00000008710  | 0.79  | AGCTTCTGGCCCCAGTTCCCT                | 14664|
|                  |       | AGCTTCTGGCCCCAGTTCCCT                |      |
| ENSG00000179698  | 0.79  | AGCTTCTGGCCCCAGTTCCCT                | 3729 |
|                  |       | AGCTTCTGGCCCCAGTTCCCT                |      |
| ENSG00000198722  | 0.79  | ACGCTTCTGTGCCCCAGTTCCCT              | 1093 |
|                  |       | A-GCTTCTGTGCCCCAGTTCCCT              |      |
| ENSG00000131979  | 0.79  | AGCTTCTGGCCCCAGTTCCCT                | 2034 |
|                  |       | AGCTTCTGGCCCCAGTTCCCT                |      |
| ENSG00000254401  | 0.79  | AGCTTCTGGCCCCAGTTCCCT                | 1539 |
|                  |       | AGCTTCTGGCCCCAGTTCCCT                |      |
| ENSG00000071246  | 0.79  | AGCTTCTGGCCCCAGTTCCCT                | 260  |
|                  |       | AGCTTCTGGCCCCAGTTCCCT                |      |
| ENSG00000142698  | 0.79  | AGCTTCTGTGCCCCAGTTCCCT               | 1122 |
|                  |       | AGCTTCTGGCCCCAGTTCCCT               |      |
| ENSG00000237945  | 0.79  | AGCTTCTGGCCCCAGTTCCCT                | 7297 |
|                  |       | AGCTTCTGGCCCCAGTTCCCT                |      |
| ENSG00000246582  | 0.79  | AGCTTCTGGCCCCAGTTCCCT                | 528  |
|                  |       | AGCTTCTGGCCCCAGTTCCCT                |      |
| ENSG00000223651  | 0.79  | AGCTTCTGTGCCCCAGTTCCCT               | 1184 |
|                  |       | AGCTTCTGTGCCCCAGTTCCCT               |      |
| ENSG00000136002  | 0.79  | AGCTTCTGGCCCCAGTTCCCT                | 1439 |
|                  |       | AGCTTCTGGCCCCAGTTCCCT                |      |
| ENSG00000268926  | 0.79  | AGCTTCTGTGCCCCAGTTCCCT                | 1670 |
|                  |       | AGCTTCTGTGCCCCAGTTCCCT                |      |
| ENSG00000173846  | 0.79  | AGCTTCTGTGCCCCAGTTCCCT                | 1765 |
|                  |       | AGCTTCTGTGCCCCAGTTCCCT                |      |
| ENSG00000167792  | 0.79  | AGCTTCTGTGCCCCAGTTCCCT                | 871  |
|                  |       | AGCTTCTGTGCCCCAGTTCCCT                |      |
AGCTTCTGGCCCAGTTCCT
19 ENSG00000162779 0.79 AGCTTCT-GCCCCAGTCCT
AGCTTCTGGCCCAGTTCCT
20 ENSG00000173020 0.79 AGCTTTGTAGGCCTCACTTTCCT
AGCTTCT-GGCC-CAGTTCCT
21 ENSG00000178425 0.79 AGCTT-T-GCCCAGTTCCT
AGCTTCTGGCCCAGTTCCT
22 ENSG00000165757 0.79 AGACTTCTGGTGGCCCAGTTCCT
AGCTTCTGGTGGCCCAGTTCCT
23 ENSG00000077942 0.79 AGCTTCTGGTGGCCCAGTTCCT
AGCTTCTGGTGGCCCAGTTCCT
24 ENSG00000115539 0.79 ACGCTTCTGTGCCTCAGTTCCT
AGCTTCTGGTGGCCCAGTTCCT
25 ENSG00000139637 0.79 AGCTGCTGGCCCAGTTCCT
AGCTTCTGGTGGCCCAGTTCCT
26 ENSG00000127481 0.79 GGCTTCTGGCCCAGTTCCT
AGCTTCTGGTGGCCCAGTTCCT
27 ENSG00000142230 0.79 AGCTTCTGGCAGTTCCT
AGCTTCTGGTGGCCCAGTTCCT
28 ENSG00000118729 0.79 AGCTTCTGGCAGTTCCT
AGCTTCTGGTGGCCCAGTTCCT
29 ENSG00000171658 0.74 AGCTTCTGGCAGTTCCT
AGCTTCTGGCAGTTCCT
30 ENSG00000171658 0.74 AGCTTCTGGCAGTTCCT
AGCTTCTGGCAGTTCCT
31 ENSG0000015802 0.74 AGCTTCTGGCAGTTCCT
AGCTTCTGGCAGTTCCT
32 ENSG00000134184 0.74 AGCTTCTGGCAGTTCCT
AGCTTCTGGCAGTTCCT
33 ENSG00000239213 0.74 AGCTTCTGGCAGTTCCT
AGCTTCTGGCAGTTCCT
34 ENSG00000176473 0.74 AGCTTCTGGCAGTTCCT
AGCTTCTGGCAGTTCCT
35 ENSG00000084693 0.74 AGCTTCTGGCAGTTCCT
AGCTTCTGGCAGTTCCT
|   | ENSG00000107537 | 0.74 | AGCTTACTGCGGCCTCAAGTTCCT 1294 |
|---|----------------|------|-----------------------------|
|   | ENSG00000104613 | 0.74 | AGCTTCT-C-CCCCAGTCTCCT 6185 |
|   | ENSG00000072134 | 0.74 | AGCTTGGCTAGCCAAAGTTCCT 631 |
|   | ENSG00000164591 | 0.74 | AGCCTCTGGCTCCAGCTCCT 317 |
|   | ENSG00000256166 | 0.74 | AGCTTCCC-TAGCCCAAGTTCCT 179 |
|   | ENSG00000185829 | 0.74 | AGCCT-TGGCCCAGTGCCT 4968 |
|   | ENSG00000007312 | 0.74 | AGCCTGTGGCCCAAGTTCCT 344 |
|   | ENSG00000137496 | 0.74 | AGCTTGTGTGGCCCAAGTTCCT 304 |
|   | ENSG00000111678 | 0.74 | ATGCTTCTGGGCAGTGCCT 934 |
|   | ENSG00000127561 | 0.74 | GGCTTCC-TGGCCCAGTCCCT 3053 |
|   | ENSG00000197563 | 0.74 | TGCTGTGTGGCCCAGTTCCT 12631 |
|   | ENSG00000158292 | 0.74 | AGCCTGCTGGCCAGTTCCCT 3407 |
|   | ENSG00000184640 | 0.74 | AGC-TCTG.TGGCCAGTTCCCT 1977 |
|   | ENSG00000185038 | 0.74 | AGCTTCTGGCCCAAGTTCCT 2163 |
|   | ENSG00000240891 | 0.74 | AGCTTCTGGCCCAAGTTCCT 6415 |
|   | ENSG00000131871 | 0.74 | AGCTTACTGCGGCCTCAAGTTCCT 1261 |
|    | Protein ID          | Confidence | Sequence 1          | Sequence 2          |
|----|--------------------|------------|---------------------|---------------------|
| 54 | ENSG00000138073    | 0.74       | AGCTTCTGGCCAGTTTCCCT | AG-CTTCTGGCCAGTTTCCCT |
| 55 | ENSG00000078808    | 0.74       | AGCTCTCTGTCGCGCCGAGTTCATCT | AGCT-CCTG-GCCCAGTTTCCCT |
| 56 | ENSG00000197056    | 0.74       | AGCTATCTGAGCCCGAGTTCATCT | AGCTTCTG-GCCCAGTTTCCCT |
| 57 | ENSG00000168350    | 0.74       | ATG-CTTCTGGCCAGTTTCCCT | AGCTTCTG-GCCCAGTTTCCCT |
| 58 | ENSG0000004777     | 0.74       | AGCTCTCTGTCGCGCCGAGTTCATCT | AGCTTCTG-GCCCAGTTTCCCT |
| 59 | ENSG0000010803     | 0.74       | CGTCTTCTTGGGCAAGTTCATCT | AGCTTCTG-GCCCAGTTTCCCT |
| 60 | ENSG00000158786    | 0.74       | ACGTCTTTTAAGCCCGAGTTCATCT | AGCTTCTG-GCCCAGTTTCCCT |
| 61 | ENSG00000279072    | 0.74       | AGCTTTTTTTGCCCCAGTTTCCCT | AGCTTCTG-GCCCAGTTTCCCT |
| 62 | ENSG0000017920     | 0.74       | AGCTTCTTGTCGCGCCGAGTTCATCT | AGCTTCTG-GCCCAGTTTCCCT |
| 63 | ENSG00000175215    | 0.74       | AGCTTCTTGTCGCGCCGAGTTCATCT | AGCTTCTG-GCCCAGTTTCCCT |
| 64 | ENSG00000229847    | 0.74       | AGCTTTTTTTGCCCCAGTTTCCCT | AGCTTCTG-GCCCAGTTTCCCT |
| 65 | ENSG00000072501    | 0.74       | ACCTTCTGTGCGCAGTTTCCCT | AGCTTCTG-GCCCAGTTTCCCT |
| 66 | ENSG00000276805    | 0.74       | AGCTTCTGTGCGCAGTTTCCCT | AGCTTCTG-GCCCAGTTTCCCT |
| 67 | ENSG00000203685    | 0.74       | AGCTTCTGTGCGCAGTTTCCCT | AGCTTCTG-GCCCAGTTTCCCT |
| 68 | ENSG00000112541    | 0.74       | TGGCTTCTGAGGCCCAAGTTTCCCT | AGCTTCTG-GCCCAGTTTCCCT |
| 69 | ENSG00000167972    | 0.74       | AGCTTCTGTGCGCAGTTTCCCT | AGCTTCTG-GCCCAGTTTCCCT |
| 70 | ENSG00000133961    | 0.74       | AGCTTCTGTGCGCAGTTTCCCT | AGCTTCTG-GCCCAGTTTCCCT |
71  ENSG00000257431  0.74  AGC-T-TGCCACCCAGTTTCACT
AGCTTTCTGGC-CCAGTTTCTT

72  ENSG00000185361  0.74  AGCTTCTCTGCTCAGTTTCCC
AGCTTCT-TGCCAGTTTCTT

73  ENSG00000166340  0.74  AGC-CTGTGCCCAGTTCCT
AGCTTCTCTGCCCAGTTCCT

74  ENSG00000120053  0.74  AGTTCCTCTGTGTGCCAGTTTCT
AG-CTCTCTGCCCAGTTCCT

75  ENSG00000242852  0.74  AAGGCTTTCAGTTGGCCCAGTTTCT
A-GC-TTCATGCCAGTTTCTT

76  ENSG00000168795  0.74  AGCTTTGTGCGCCAGTTTCT
AGCTTCTCTGCCCAGTTCCT

77  ENSG00000257060  0.74  AGCTTCTCTGCCCAGTTCCT
AGCT-TCTGCCCAGTTCCT

78  ENSG00000107021  0.74  AGCTGTATGCTGCCCAGTTCCT
AGCT-TCTGCCCAGTTCCT

79  ENSG00000230724  0.74  AGC-CTTCTTGGCCCAGTTTCT
AGC-TCTTCTGCCCAGTTCCT

80  ENSG00000197046  0.74  AAGGCTTTCCTGTCCCAGTTTCT
A-GCTTCTCTGCCCAGTTCCT

81  ENSG00000116698  0.74  AGCTCATCTGTTCCCAGTTTCT
AG-CTCTCTGCCCAGTTCCT

82  ENSG00000150961  0.74  AAGGCTTTCCTGTCCCAGTTTCT
A-GCTTCTCTGCCCAGTTCCT

83  ENSG00000137504  0.74  AGGCTTCTGCCCAGTTTCT
AGCTTCTCTGCCCAGTTCCT

84  ENSG00000123938  0.74  AGCTTCTGCCCAGTTTCT
AGCTTCTCTGCCCAGTTCCT

85  ENSG00000137504  0.74  AGGCTTCTGCCCAGTTTCT
AGCTTCTCTGCCCAGTTCCT

86  ENSG00000137504  0.74  AGGCTTCTGCCCAGTTTCT
AGCTTCTCTGCCCAGTTCCT

87  ENSG00000137504  0.74  AGGCTTCTGCCCAGTTTCT
AGCTTCTCTGCCCAGTTCCT

88  ENSG00000137504  0.74  AG-CTTCTGCCCAGTTTCT
AGCTTCTCTGCCCAGTTCCT

43
| Gene ID      | Score | Sequence 1                        | Sequence 2                        | Length |
|-------------|-------|-----------------------------------|-----------------------------------|--------|
| ENSG0000015592 | 0.74  | AGCTTCTGGCCCA-G-TTC-CT            | AG-CT-TCTCTGTGCCTTCAGTTTCCT       | 1774   |
| ENSG00000186469 | 0.74  | AGC-CTTTGGCCCACTCAGTTTCCT         | AGCTTC-TGGCCCACTCAGTTTCCT         | 2335   |
| ENSG00000103148 | 0.74  | AGCTTCTGCTGAGCCCACTCAGTTTCCT     | AG-C-TCTCTGGCCCACTCAGTTTCCT       | 5486   |
| ENSG00000272556 | 0.74  | AGCCTTTTTGGCCCACTCAGTTTCCT       | AGC-TCTCTGGCCCACTCAGTTTCCT        | 3440   |
| ENSG00000253326 | 0.74  | AGCCTTTTTTTGGCCCACTCAGTTTCCT     | AGC-TCTCTGGCCCACTCAGTTTCCT        | 4      |
| ENSG00000165816 | 0.74  | AGCTTTTTGGCCCACTCAGTTTCCT       | AGCTTTTTTTGGCCCACTCAGTTTCCT       | 690    |
| ENSG00000183242 | 0.74  | AGCTTTTTGGCCCACTCAGTTTCCT       | AGCTTTTTTTGGCCCACTCAGTTTCCT       | 1576   |
| ENSG00000178187 | 0.74  | AGCTTTTTGGCCCACTCAGTTTCCT       | AGCTTTTTTTGGCCCACTCAGTTTCCT       | 484    |
| ENSG000000108219 | 0.74  | AGCTTTTTGGCCCACTCAGTTTCCT       | AGCTTTTTTTGGCCCACTCAGTTTCCT       | 12861  |
| ENSG00000263072 | 0.74  | AGCTTTTTGGCCCACTCAGTTTCCT       | AGCTTTTTTTGGCCCACTCAGTTTCCT       | 2448   |
| ENSG00000237424 | 0.74  | AGCTTTTTGGCCCACTCAGTTTCCT       | AGCTTTTTTTGGCCCACTCAGTTTCCT       | 2295   |
| ENSG000000263072 | 0.74  | AGCTTTTTGGCCCACTCAGTTTCCT       | AGCTTTTTTTGGCCCACTCAGTTTCCT       | 3482   |
| ENSG000000108219 | 0.74  | AGCTTTTTGGCCCACTCAGTTTCCT       | AGCTTTTTTTGGCCCACTCAGTTTCCT       | 1058   |
| ENSG00000143355 | 0.74  | AGCTTTTTGGCCCACTCAGTTTCCT       | AGCTTTTTTTGGCCCACTCAGTTTCCT       | 1271   |
| ENSG00000198911 | 0.74  | AGCTTTTTGGCCCACTCAGTTTCCT       | AGCTTTTTTTGGCCCACTCAGTTTCCT       | 1937   |
| ENSG00000125637 | 0.74  | AGCTTTTTGGCCCACTCAGTTTCCT       | AGCTTTTTTTGGCCCACTCAGTTTCCT       | 8161   |
| ENSG00000241973 | 0.74  | AGCTTTTTGGCCCACTCAGTTTCCT       | AGCTTTTTTTGGCCCACTCAGTTTCCT       | 7916   |
| ENSG ID  | Score | Sequence 1 | Sequence 2 |
|---------|-------|------------|------------|
| ENSG00000092068 | 0.74 | AGCATTGGCTGAAGCCCAGTTCCT | AGC-TT-CTG-GCCCAGTTCCT |
| ENSG00000251143 | 0.74 | AGCTTTCTGGCCCTAGCTTCCT | AGCTTCTGGCCCAGTTCCT |
| ENSG00000188612 | 0.74 | AACTTCCTGAGCTTCCCT | AGCTTCTGGCCAGTTCCT |
| ENSG00000188039 | 0.74 | AGCCTTCTGTCCCA-TTCCT | AGCTTCTGGCCCAGTTCCT |
| ENSG00000225465 | 0.74 | AGCTT-GGGCTCCAGTTCCT | AGCTTCTGGCCCAGTTCCT |
| ENSG00000064999 | 0.74 | AGC-TGTGTGCCCAGTTCCT | AGCTTCTGGCCCAGTTCCT |
| ENSG00000150995 | 0.74 | ACCTTCTGGCCCTGAGCTTCCT | AGCTTCTGGCCCAGTTCCT |
| ENSG00000226995 | 0.74 | AGTTTCTGGCCTCAGGTCCT | AGCTTCTGGCCCAGTTCCT |
| ENSG00000279141 | 0.74 | AGCATCTGGGACCCCAGTTCCT | AGCTTCTGGCCCAGTTCCT |
| ENSG00000279440 | 0.74 | AGTCCCTCTGGCCCTAGTTCCT | AGCTTCTGGCCCAGTTCCT |
| ENSG00000233837 | 0.74 | A-CTTCTGGCCCCAAGTTCCT | AGCTTCTGGCCCAGTTCCT |
| ENSG00000230590 | 0.74 | AGACTTCTGGCCTCAAGTAATCCT | AGCTTCTGGCCCAGTTCCT |
| ENSG00000060718 | 0.74 | ATGCTTTCTGGCCTACAGTTCCT | AGCTTCTGGCCCAGTTCCT |
| ENSG00000170322 | 0.74 | AGTTTTCTGGCCCCAGTTCCT | AGCTTCTGGCCCAGTTCCT |
| ENSG00000182010 | 0.74 | AGCTTTCTGGGGTGCTTCAT | AGCTTCTGGCCCAGTTCCT |
| ENSG00000148187 | 0.74 | AGCTTCTGTGCCCAGTTCCT | AGCTTCTGGCCCAGTTCCT |
| ENSG00000130830 | 0.74 | AGCCTTCTGGCCCAGGTTTCCT | AGCTTCTGGCCCAGTTCCT |
| ENSG00000228327 | 0.74 | AGCCTTTTTTGCCAGTTCCT | AGCTTCTGGCCCAGTTCCT |

45
AG-C-TTCTGGCCCAGTTCCT

124 ENSG0000029433 0.74 AGCCTCTCTGCGCCGCAGTTCCCT
AG-CT-TCTG-GCC-CAGTT-CCT 748

AG-CT-TCTG-GCC-CAGTT-CCT

125 ENSG00000101203 0.74 AGGCTTCTGAGCCCCAAGTTGCCT
A-GCTTCTG-G-CCC-AGTT-CCT 601

AG-C-TTCTGGCCCAGTTCCT

126 ENSG00000149809 0.74 AGCCTGAGGCCCAGTTCCCT
AG-CTTCTGGCCCAGTTCCCT 2273

AG-CTTCTGGCCCAGTTCCT

127 ENSG00000283050 0.74 AGCCTTTTTTGCCCCAGTTCCCT
AG-C-TTCTGGCCCAGTTCCCT 4237

AGCTTCTGG–CCC-AGTTCCT

128 ENSG00000256028 0.74 AGC-TCTGGCCCGAGTTCCCT
AGCTTCTGG–CCC-AGTTCCT 241

AG-CTTCTGGCCCAGTTCCT

129 ENSG00000164073 0.74 ATTCATCTGGCCCAGTTCCT
A-GCTTCTGGCCCAGTTCCCT 5201

AG-CTTCTGGCCCAGTTCCT

130 ENSG00000237438 0.74 AGCCTTCTGGCCCCAGTTCCCT
AG-CTTCTGGCCCAGTTCCCT 2347

AGCTTCTGGCCCAGTTCCT

131 ENSG00000119772 0.74 AGCTTCTGGCCCCAGTTCCCT
AG-CTTCTGGCCCAGTTCCCT 6769

AGCTTCTGGCCCAGTTCCT

132 ENSG00000256028 0.74 AGC-TCTGGCCCGAGTTCCCT
AGCTTCTGG–CCC-AGTTCCT 241

AGCTTCTGGCCCAGTTCCT

133 ENSG00000164073 0.74 ATTCATCTGGCCCAGTTCCT
A-GCTTCTGGCCCAGTTCCCT 5201

AGCTTCTGGCCCAGTTCCT

134 ENSG00000168765 0.74 AGTCTTCTG-CCCAATTCCT
AG-CTTCTGGCCCAGTTCCCT 1005

AGCTTCTGGCCCAGTTCCT

135 ENSG00000172819 0.74 AGCTAC-GGCCCAGTTCCCT
AGCTTCTGGCCCAGTTCCCT 915

AGCTTCTGGCCCAGTTCCT

136 ENSG00000168765 0.74 AGTCTTCT-GCCCAATTCCT
AG-CTTCTGGCCCAGTTCCCT 1005

AGCTTCTGGCCCAGTTCCT

137 ENSG00000149091 0.74 AGCTAC-GGCCCAGTTCCCT
AGCTTCTGGCCCAGTTCCCT 915

AGCTTCTGGCCCAGTTCCT

138 ENSG00000214837 0.74 AGCCTTTTTGCCCCAGTTCCCT
AG-C-TTCTGGCCCAGTTCCCT 3757

AGCTTCTGGCCCAGTTCCCT

139 ENSG00000168939 0.74 AGC-TCT-GCCCCAGTTCT
AGCTTCTGGCCCAGTTCCCT 3687

AGCTTCTGGCCCAGTTCCCT

140 ENSG00000178188 0.74 AGCTGCTCTGGCCCCAG-TTCT
AGCT-TCTGG–CCCAGTTCCCT 2194

46
|   | ENSG00000164855 | 0.74 | AGCCTCTCCTGGCCTCAGTATCCT | AG-CT-T-CTGGCC-CAGT-TCCT | 5633 |
|---|-----------------|------|--------------------------|--------------------------|------|
|   | ENSG00000260404 | 0.74 | AGCCTTTTTTGGCCCCAGTTCCT | AG-C-TTCTGGCCCAGTTCCT    | 2012 |
|   | ENSG00000228696 | 0.74 | AGCTT-GGCCCCAGTGCCCT    | AGCTTCTGGCCCAGTTCCT      | 3341 |
|   | ENSG00000205176 | 0.74 | AGACTGCTGGCCCCAGTTCCT   | AG-CTTCTGGCCCAGTTCCT     | 6432 |
|   | ENSG00000111325 | 0.74 | AGC-TCTGGCCGCCCAGTTCCT  | AGCTTCTGG-CCC-AGTTCCT    | 1749 |
|   | ENSG00000174405 | 0.74 | AGTCCTTCTGCGCCAGCTCCG   | AG-CTTCTGGCCAGTTCCT      | 206  |
|   | ENSG00000177084 | 0.74 | CGCTCTCTGGCCCCAGTCAGCT  | AGCT-TCTGGCCCAGTTC-CT    | 9130 |
|   | ENSG00000171700 | 0.74 | AGCTGGCTGGCCCCAGTGCCTC  | AGCT-TCTGGCCCAGTTC-CT    | 1605 |

Table 4: Potential target RNAs for svRNA3 (truncated version).
| Gene ID       | Score | Matched Sequence                        | Loci     |
|--------------|-------|-----------------------------------------|----------|
| NC000002.12  | 0.89  | GAGGAAGAAGAAGGTGAT                     | 134222364|
| chr2         |       | GATGAAGAAGAAGGTGAT                     |          |
| NC000003.12  | 0.89  | GATAAAGAAGAAGGTGAT                     | 62430918 |
| chr3         |       | GATGAAGAAGAAGGTGAT                     |          |
| NC000012.12  | 0.89  | GATGCAAGATAGAAGGTGAT                   | 43685381 |
| chr12        |       | GATG-AAGA-AGAAGGTGAT                   |          |
| NC000012.12  | 0.89  | GAT-AAGAAGAAGGTGAT                     | 22739819 |
| chr12        |       | GATGAAGAAGAAGGTGAT                     |          |
| NC000007.14  | 0.89  | GATGAATGAAGAAGGTGAT                    | 690091   |
| chr7         |       | GATGAA-GAAGAAGGTGAT                    |          |
| NC000007.14  | 0.89  | GATGAAGAAGAAGGTGAT                     | 5993096  |
| chr7         |       | GATGAA-GAAGAAGGTGAT                    |          |
| NC000011.10  | 0.89  | GATG-AGAAGAAGGTGAT                     | 33698282 |
| chr11        |       | GATGAAGAAGAAGGTGAT                     |          |
| NC000015.10  | 0.89  | GAAGAAGAAGAAGGTGAT                    | 13363820 |
| chr15        |       | GATGAAGAAGAAGGTGAT                     |          |
| NC000015.10  | 0.89  | GAAGAAGAAGAAGGTGAT                    | 5648009  |
| chr15        |       | GATGAAGAAGAAGGTGAT                     |          |
| NC000006.12  | 0.89  | GAAGAAGAAGAAGGTGAT                    | 39349582 |
| chr6         |       | GATGAAGAAGAAGGTGAT                     |          |
| NC000003.12  | 0.89  | GATGATAGAAGACAGGTGAT                   | 38020582 |
| chr3         |       | GATGA-AGAAGA-AGGTGAT                   |          |
| NC000003.12  | 0.89  | GATGAAGAAAGAAGGTGAT                   | 52630276 |
| chr3         |       | GATGAA-GAAGAAGGTGAT                    |          |
| NC000004.12  | 0.89  | GATGAAAGAAGG-GAT                      | 88689121 |
| chr4         |       | GATGAAGAAGAAGGTGAT                     |          |
| NC000003.12  | 0.89  | GATGAAG-AGAAGGTGAT                    | 96774073 |
| chr3         |       | GATGAAGAAGAAGGTGAT                     |          |
| NC000006.12  | 0.89  | GATCAAGAAGAAGGTGAT                    | 68689908 |
| chr6         |       | GATGAAGAAGAAGGTGAT                     |          |
| NW021160000.1| 0.83  | GATGTAGAAGACAGGTGAT                   | 376177   |
| chr10        | patch| GATGAAGAAGAAGGTGAT                     |          |
| FIX          |       | GATGAAGAAGAAGGTGAT                     |          |
| NC000009.12  | 0.83  | GATGCAAGAAGACAGGTGAT                   | 5316996  |
| chr9         |       | GATGAAGAAGAAGGTGAT                     |          |
|   | Accession | Chromosome | Position | % Similarity | Sequence A | Sequence B |
|---|-----------|------------|----------|--------------|------------|------------|
| 18 | NC000009.12 chr9 | 0.83 | GATGAAGAAGGAAGGTGGT GATGAAGAA-GAAGGTGAT | 17792287 |
| 19 | NC000009.12 chr9 | 0.83 | GATGCAGAAGACAGGTGAT GATGAAGAAGA-AGGTGAT | 39264386 |
| 20 | NW012132919.1 chr7 patch FIX | 0.83 | AATGAAGAAGACAGGTGAT GATGAAGAAGA-AGGTGAT | 106833 |
| 21 | NW018654717.1 chr8 patch FIX | 0.83 | GATGAAGAAGAGGGTGAT GATG-AAGAAGAAGGTGAT | 3239611 |
| 22 | NC000008.11 chr8 | 0.83 | GATGCAGAAGACAGGTGAT GATGAAGAAGA-AGGTGAT | 20240737 |
| 23 | NC000008.11 chr8 | 0.83 | GAGTGAAGAATGAAGGTGATCAT GA-AGAAGAA-GAAGGTGAT | 23109525 |
| 24 | NC000008.11 chr8 | 0.83 | GAT-AAGAAGAGGTGAT GATGAAGAAGA-AGGTGAT | 29218614 |
| 25 | NC000008.11 chr8 | 0.83 | GA-GAAGAAGAGCCGTGAT GATGAAGAAGAAGGTGAT | 29390573 |
| 26 | NC000008.11 chr8 | 0.83 | GATGCAGAAGACAGGTGAT GATGAAGAAGA-AGGTGAT | 36287121 |
| 27 | NC000008.11 chr8 | 0.83 | GATGAAGCCAAGAAGGTGAT GATGAAGAAGAAAGGTGAT | 48086402 |
| 28 | NC000008.11 chr8 | 0.83 | GATGAAGAAGGAAGGTACAGAT GATGAAGAAGAAGGTGAT | 49093880 |
| 29 | NC000002.12 chr2 | 0.83 | GATGAAGAAGAAATGGTGAT GATGAAGAAGAAGGTGAT | 132283207 |
| 30 | NC000006.12 chr6 | 0.83 | GATGAAGAATGAAGTTGAT GATGAAGAAGAAGGTGAT | 11193208 |
| 31 | NC000006.12 chr6 | 0.83 | GATGGGAGGAGGAAGGTGAT GAT-GA-AGAAGAAGGTGAT | 12861402 |
| 32 | NC000006.12 chr6 | 0.83 | GATGCAGAAGACAGGTGAT GATGAAGAAGA-AGGTGAT | 20151370 |
| 33 | NC000006.12 chr6 | 0.83 | GATGATAGAAGAGGTTGAT GATGA-AG-AAGAAGG-TGAT | 20163463 |
| 34 | NC000006.12 chr6 | 0.83 | GATGCAGAAGACAGGTGAT GATGAAGAAGA-AGGTGAT | 21271973 |
| ID     | Chromosome | Start | End   | Sequence                  |
|--------|------------|-------|-------|---------------------------|
| 35     | NC000005.10 chr5 | 0.83  | GAAAGAAGAAGGTGAAT GATGAAGAAGAAGGTGAT | 22996338 |
| 36     | NC000006.12 chr6 | 0.83  | GATGAAGAATGAAAGTTTGAT GATGAAGAAGAAGGTGAT | 23465955 |
| 37     | NC000005.10 chr5 | 0.83  | GATGAAGAAGAAGGTGAT GATGAAGAAGAAGGTGAT | 8462537  |
| 38     | NC000005.10 chr5 | 0.83  | GATGAAGAAGAAGGTGAT GATGAAGAAGAAGGTGAT | 7721853  |
| 39     | NC000006.12 chr6 | 0.83  | GATGAAGAAGAAGGTGAT GATGAAGAAGAAGGTGAT | 24455160 |
| 40     | NC000006.12 chr6 | 0.83  | GATGAAGAAGAAGGTGAT GATGAAGAAGAAGGTGAT | 28251421 |
| 41     | NC000005.10 chr5 | 0.83  | GATGAAGAAGAAGGTGAT GATGAAGAAGAAGGTGAT | 6331339  |
| 42     | NC000009.12 chr9 | 0.83  | GATGAAGAAGAAGGTGAT GATGAAGAAGAAGGTGAT | 4019042  |
| 43     | NC000011.10 chr11 | 0.83  | GATGAAGAAGAAGGTGAT GATGAAGAAGAAGGTGAT | 7591884  |
| 44     | NC000011.10 chr11 | 0.83  | GATGAAGAAGAAGGTGAT GATGAAGAAGAAGGTGAT | 32927191 |
| 45     | NC000014.9 chr14 | 0.83  | GATGAAGAAGAAGGTGAT GATGAAGAAGAAGGTGAT | 19059994 |
| 46     | NC000002.12 chr2 | 0.83  | GATGAAGAAGAAGGTGAT GATGAAGAAGAAGGTGAT | 10699700 |
| 47     | NC000002.12 chr2 | 0.83  | GATGAAGAAGAAGGTGAT GATGAAGAAGAAGGTGAT | 23878256 |
| 48     | NC000002.12 chr2 | 0.83  | GATGAAGAAGAAGGTGAT GATGAAGAAGAAGGTGAT | 30292923 |
| 49     | NC000002.12 chr2 | 0.83  | GATGAAGAAGAAGGTGAT GATGAAGAAGAAGGTGAT | 33544254 |
| 50     | NC000002.12 chr2 | 0.83  | GATGAAGAAGAAGGTGAT GATGAAGAAGAAGGTGAT | 36195723 |
| 51     | NC000002.12 chr2 | 0.83  | GATGAAGAAGAAGGTGAT GATGAAGAAGAAGGTGAT | 41776674 |
| 52     | NC000002.12 chr2 | 0.83  | GATGAAGAAGAAGGTGAT GATGAAGAAGAAGGTGAT | 44826227 |
| Position | Accession | Chromosome | Identity | Sequence 1 | Sequence 2 |
|----------|-----------|------------|----------|------------|------------|
| 53       | NC000002.12 chr2 | chr2 | 0.83     | GATGAAGAAGA-AGGTGAT | GATG-AAG-AAGA-AGGTGAT |
| 54       | NC000002.12 chr2 | chr2 | 0.83     | GATGAAGAAG-AGGTGAGT | GATGAAGAAGAAGGTGAGT-T |
| 55       | NC000001.11 chr1 | chr1 | 0.83     | GATGAAGAAGGACAGGTGAGT | GATGAAGAA-GA-AGGTGAGT-T |
| 56       | NC000001.11 chr1 | chr1 | 0.83     | GAGTGAAGGAAGAAAGGTGAT | GA-TGAA-GAAG-AAGGTGAT |
| 57       | NC000001.11 chr1 | chr1 | 0.83     | GAATGAAGGAAGAAAGGTGAT | G-ATGAA-GAA-GAAGGTGAT |
| 58       | NC000014.9 chr14 | chr14 | 0.83    | GATAGAAGATAGAAAGGTGAGT | GAT-GAAGAAGG-AAGGTGAT |
| 59       | NC000014.9 chr14 | chr14 | 0.83    | GATGAAGAAGGAGGTGAT | GATGAAGAA-GAAGGTGAT |
| 60       | NC000014.9 chr14 | chr14 | 0.83    | GATGCAGAAGACAGGTGAT | GATGAAGAAGAAGGTGAT |
| 61       | NC000011.10 chr11 | chr11 | 0.83    | GATGAAGAATAAGGTGTAT | GATGAAGAAGAAGGTGAT-AT |
| 62       | NC000014.9 chr14 | chr14 | 0.83    | GATGATAGATAGATAGGTGAT | GATGA-AGA-AGA-AGGTGAT |
| 63       | NC000014.9 chr14 | chr14 | 0.83    | GATGTAGAAGCAGGTGAT | GATGAAGAAGAAGGTGAT |
| 64       | NC000014.9 chr14 | chr14 | 0.83    | GAATGAAGAAGTAGGTGAT | G-ATGAA-GAA-GAAGGTGAT |
| 65       | NC000014.9 chr14 | chr14 | 0.83    | GATGAAGTAGGAAGGTGAT | GATGAAGAAGAAGGTAAGGTGAT |
| 66       | NC000014.9 chr14 | chr14 | 0.83    | GATGAAGATAGGAGGTGAT | GATGAAGAAGAAGGTGAT |
| 67       | NC000014.9 chr14 | chr14 | 0.83    | GATGAAGATAGGAGGTGAT | GATGAAGAAGAAGGTGAT |
| 68       | NC000012.12 chr12 | chr12 | 0.83    | GATGCAGAAGACAGGTGAT | GATGAAGAAGAAGGTGAT |
| 69       | NC000012.12 chr12 | chr12 | 0.83    | GATGCAGAAGACAGGTGAT | GATGAAGAAGAAGGTGAT |
| ID          | Accession   | Chromosome | Similarity | Sequence 1               | Sequence 2               | Length     |
|------------|-------------|------------|------------|-------------------------|-------------------------|------------|
| 70         | NC000012.12 | chr12      | 0.83       | GATGCAGAAGACAGGTGAT     | GATGAAGAAGA-AGGTGAT     | 4602928    |
| 71         | NC000012.12 | chr12      | 0.83       | GATAGAAGAAG-AGGTGAT     | GAT-GAAGAAGAAGGTGAT     | 4904999    |
| 72         | NC000012.12 | chr12      | 0.83       | GATGAGAAGAAGAAAGTGAT    | GATGA-AGAAGAAGGTGAT     | 8026984    |
| 73         | NW015495299 | chr2       | 0.83       | GATGAAGAAGAA-GTGCAT     | GATGAAGAAGAAGGTGAT-AT   | 465928     |
| 74         | NC000011.10 | chr11      | 0.83       | GATGCAGAAGACAGGTGAT     | GATGAAGAAGA-AGGTGAT     | 11790646   |
| 75         | NC000011.10 | chr11      | 0.83       | GATGAAGAAGTTAAGGCTGAT   | GATGAAGAAGG-AGG-TGAT    | 13950942   |
| 76         | NW019805490 | chr3       | 0.83       | GATTTGAAGAGTGATG        | GA-TGAAGAAGAAGGTGAT     | 212653     |
| 77         | NW019805503 | chr18      | 0.83       | GCATTAAAGAAGGTGAT       | G-ATGAAGAAGAAGGTGAT     | 148518     |
| 78         | NC000002.12 | chr2       | 0.83       | GAT-AAGAAGACAGGTGAT     | GATGAAGAAGA-AGGTGAT     | 10335731   |
| 79         | NC000003.12 | chr3       | 0.83       | GATGAAGAAGGAAAGTGAG    | GATGAAGAA-GAAGGTGAT     | 10434488   |
| 80         | NC000003.12 | chr3       | 0.83       | GATGCAGAAGACAGGTGAT     | GATGAAGAAGA-AGGTGAT     | 6804109    |
| 81         | NC000018.10 | chr18      | 0.83       | GATGCAGAAGACAGGTGAT     | GATGAAGAAGA-AGGTGAT     | 16198420   |
| 82         | NC000004.12 | chr4       | 0.83       | GATGATGAAGAAGAGTGAT     | GATGAAGAAGAAG-AGTGAT    | 84784338   |
| 83         | NC000004.12 | chr4       | 0.83       | GATGCAGAAGACAGGTGAT     | GATGAAGAAGAAGGTGAT     | 83332038   |
| 84         | NC000004.12 | chr4       | 0.83       | GATGAAGAAGAAGGAAGAT     | GATGAAGAAGAAGG-TGAT    | 78956485   |
| 85         | NC000004.12 | chr4       | 0.83       | GAATGAAGAAGAAGGTGCT    | G-ATGAAGAAGAAGGTGAT     | 78780579   |
| 86         | NC000004.12 | chr4       | 0.83       | GTATGAAGAAG-AGGTGAT     | GATGAAGAAGAAGGTGAT     | 77419436   |

52
| ID     | Accession  | Chromosome | Distance | Sequence 1       | Sequence 2       | Coordinate  |
|--------|------------|------------|----------|-----------------|-----------------|-------------|
| 87     | NC000004.12 | chr4       | 0.83     | GAGGAAGAAGAAGAGTGAT | GATGAAGAAGAAGAGTGAT | 76308292    |
| 88     | NC000004.12 | chr4       | 0.83     | GATGAAGAAGAAGAGTGAT | GATGAAGAAGAAGAGTGAT | 73534765    |
| 89     | NC000004.12 | chr4       | 0.83     | GATGAAGAAGAAGAGTGAT | GATGAAGAAGAAGAGTGAT | 62458987    |
| 90     | NC000020.11 | chr20      | 0.83     | GATGAAGCAGAAGAAGGTGAT | GATGAAGCAGAAGAAGGTGAT | 37225026    |
| 91     | NC000016.10 | chr16      | 0.83     | GATGAAGAAGAAGAAGGTGAT | GATGAAGAAGAAGAAGGTGAT | 34440755    |
| 92     | NC000016.10 | chr16      | 0.83     | GATGAAGAAGAAGAAGGTGAT | GATGAAGAAGAAGAAGGTGAT | 34440752    |
| 93     | NC000016.10 | chr16      | 0.83     | GATGAAGAAGAAGAAGGTGAT | GATGAAGAAGAAGAAGGTGAT | 4949206     |
| 94     | NC000016.10 | chr16      | 0.83     | GATGAAGAAGAAGAAGGTGAT | GATGAAGAAGAAGAAGGTGAT | 17439087    |
| 95     | NC000016.10 | chr16      | 0.83     | GATGAAGAAGAAGAAGGTGAT | GATGAAGAAGAAGAAGGTGAT | 16213349    |
| 96     | NC000016.10 | chr16      | 0.83     | GATGAAGAAGAAGAAGGTGAT | GATGAAGAAGAAGAAGGTGAT | 4949206     |
| 97     | NC000019.10 | chr19      | 0.83     | GATGAAGAAGAAGAAGGTGAT | GATGAAGAAGAAGAAGGTGAT | 18521261    |
| 98     | NC000019.10 | chr19      | 0.83     | GCATGAAGAAGAAGAAGGTGACT | GATGAAGAAGAAGAAGGTGACT | 19625401    |
| 99     | NC000018.10 | chr18      | 0.83     | GATGAAGAAGAAGAAGGTGAT | GATGAAGAAGAAGAAGGTGAT | 2511062     |
| 100    | NC000018.10 | chr18      | 0.83     | GATGAAGAAGAAGAAGGTGAT | GATGAAGAAGAAGAAGGTGAT | 3244108     |
| 101    | NC000018.10 | chr18      | 0.83     | GATGAAGAAGAAGAAGGTGAT | GATGAAGAAGAAGAAGGTGAT | 12177624    |
| 102    | NC000018.10 | chr18      | 0.83     | GATGAAGAAGAAGGATCGTAT | GATGAAGAAGAAGGATCGTAT | 14560185    |
| 103    | NC000003.12 | chr3       | 0.83     | GATGAAGAAGAAGGATCGTAT | GATGAAGAAGAAGGATCGTAT | 10434487    |
|       | Accession   | Chromosome | jaccard  | 1st sequence                            | 2nd sequence                            | Position   |
|-------|-------------|------------|----------|-----------------------------------------|-----------------------------------------|------------|
| 104   | NC000004.12 | chr4       | 0.83     | GAATGAGGAAGAAGGGTGAT                    | G-ATGAAGAAGAAGGGTGAT                    | 87353341   |
| 105   | NC000002.12 | chr2       | 0.83     | GATTGAAAGCAAGAAGGGTGAT                  | GA-TG-AAAG-AAAGAAGGGTGAT                | 90988927   |
| 106   | NC000003.12 | chr3       | 0.83     | GATTTGAAGATGAAGGGTGAT                   | GA-TGAAGAAGAAGGGTGAT                   | 34637960   |
| 107   | NC000002.12 | chr2       | 0.83     | GATGGAAGAAGAAGGGTGAT                   | GA-TGAAAGAAGAAGGGTGAT                 | 127837384  |
| 108   | NC000002.12 | chr2       | 0.83     | GATGAAATGAAGAAGGGTGAT                  | GATGAA-GAAGAAGGGTGAT                  | 122632234  |
| 109   | NC000002.12 | chr2       | 0.83     | GAGTGAAGAAAGAAGGGTGAT                  | GA-TGAAAGAAGAAGGGTGAT                 | 114885326  |
| 110   | NC000002.12 | chr2       | 0.83     | GATGTAAGAAGAAGGGTGAT                   | GATGAAAGAAGAAGGGTGAT                 | 108862672  |
| 111   | NC000002.12 | chr2       | 0.83     | GATGATGAAAGAAGGGTGAGT                  | GATGAAAGAAGAAGGGTGAGT               | 107221350  |
| 112   | NC000002.12 | chr2       | 0.83     | GATGCAGAAGCAAGGGTGAT                   | GATGAAAGAAGAAGGGTGAT                 | 103810398  |
| 113   | NC000002.12 | chr2       | 0.83     | GATGAAAGAAGAAGGGATAGTGAGT              | GATGAAAGAAGAAGGGATAGTGAGT             | 102943227  |
| 114   | NC00018.10  | chr18      | 0.83     | GATG-AGACAGAAGGGTGAT                   | GATGAAAGAAGAAGGGTGAT                 | 18128465   |
| 115   | NC000003.12 | chr3       | 0.83     | TATGAAGAAGAAGGGTGAT                   | GATGAAAGAAGAAGGGTGAT                 | 19152201   |
| 116   | NC000003.12 | chr3       | 0.83     | GATGCAGAAGATAGGGTGAT                   | GATGAAAGAAGAAGGGTGAT                 | 25424584   |

Table 5: Potential target regions on the human genome for svRNA1.
| Gene ID | Score | Matched Sequence | Loci       |
|---------|-------|------------------|------------|
| NC000004.12 chr4 | 0.86 | GAAAGAGAGGAAGAAGAAGAATT GAAGAAGAGCAAGAAGAGAATT | 117558370 |
| NC000014.9 chr14 | 0.86 | GAAAGAGGCAAATGAAGAAGAAGACTT GAAGAAGGCAA-AAG-AAG-AAGTT | 12719463 |
| NW018654725.1 chrY patch FIX | 0.86 | GAAAGAGGAAAGAAGAAGATTT GAAGAAGAGCAAGAAGAAGATT | 21110 |
| NC000004.12 chr4 | 0.86 | GAAAGAAGAAG-AAGAAGAAGAATT GAAGAAG-AGCAAGAAGAAGAAGATT | 130552436 |
| NC000014.9 chr14 | 0.86 | GAAAGAAGATGTAGCTCAAGAAGAGAGATAT GAAGAAGA-G-CAAGAAGAAGAT-T | 79840239 |
| NC000018.10 chr18 | 0.86 | GAAAGAGATGTAGCTCAAGAGAAGATCT GAAGAAGA-G-CAAGAAGAAGAT-T | 783992 |
| NC000003.12 chr3 | 0.86 | GAAAGAAGAGAAGAAGAAGAATT GAAGAAGAGCAAGAAGAAGATT | 60432276 |
| NC000002.12 chr2 | 0.86 | GAAAGATAGAGCAATGAAGAAGAAGATT GAAGAAGA-G-CAAGAAGAAGATT | 44201986 |
| NC000015.10 chr15 | 0.82 | GAAAGAAGAGAAGAAGAAGATAT GAAGAAG-AGCAAGAAGAAGAT-T | 5850233 |
| NC000006.12 chr6 | 0.82 | GAAAGAAGAGC-AGAAGAAGAATT GAAGAAGAGCAAGAAGAAGATT | 38233680 |
| NC000019.10 chr19 | 0.82 | GAAAGAAGAG-AGAAGAAGAAGAATT GAAGAAGAGCAAGAAGAAGATT | 16552894 |
| NC000020.11 chr20 | 0.82 | GAAAGAAGAGAAGAAGAAGACCTT GAAGAAGAGCAAGAAGAAGAATT | 2384918 |
| NC000021.9 chr21 | 0.82 | GAAAGAAGAGCAA-TAGAAGAAGATT GAAGAAGAGCAAGAAGAAGAATT | 3080346 |
| NW018654706.1 patch NOVEL chr1 | 0.82 | GAAAGAAGAG-AGAAGAAGAAGGATT GAAGAAGAGCAAGAAGAAGAATT | 21426 |
| NW017852928.1 patch NOVEL chr1 | 0.82 | GAAAGAAGAGCAAGAAGAAGAAGGATT GAAGAAGAGC-AAGAAGAAGAATT | 42684 |
| NW013171802.1 patch FIX chr6 | 0.82 | GAAG-AGAGCAAGAAGAAGAAGGATT GAAGAAGAGCAAGAAGAAGAATT | 64033 |
|   | Accession | Chromosome | Position | Sequence 1 | Sequence 2 |
|---|-----------|------------|----------|------------|------------|
| 17 | NC000002.12 chr2 | 0.82 | GAAAAAGAG-AAGAAGAAGATT<br>GAAAGAGAGCAAGAGAGAGATT | 25750788 |
| 18 | NC000002.12 chr2 | 0.82 | GAAGAAGAAG-AAGAAGAAGATT<br>GAAGAAGAGCAAGAAGAGATT | 80694044 |
| 19 | NC000005.10 chr5 | 0.82 | GAAGAAGAGCAAGAAGAGATT<br>GAAGAAGAGCAAGAAGAGATT | 14573711 |
| 20 | NC000006.12 chr6 | 0.82 | GAAGAAGAAG-AAGAAGAAGATT<br>GAAGAAGAGCAAGAAGAGATT | 69217989 |
| 21 | NC000011.10 chr11 | 0.82 | GAAGAAGAAGGAAGAAGATT<br>GAAGAAGAAGGAAGAAGATT | 25049429 |
| 22 | NC000008.11 chr8 | 0.82 | GAAGAAGAGAAAGAAGAAGATT<br>GAAGAAGAGCAAGAAGAAGATT | 34613950 |
| 23 | NC000015.10 chr15 | 0.82 | GCAGAAGAGCAAGCAAGAAGATT<br>GAAGAAGAGCAAGAAGAAGATT | 11738083 |
| 24 | NC000011.10 chr11 | 0.82 | GAAGAAGAGCAAGCAAGAAGATT<br>GAAGAAGAGCAAGAAGAAGATT | 16510998 |
| 25 | NC000012.12 chr12 | 0.82 | GAAGAAGAGCAAGAAGAAGATT<br>GAAGAAGAGCAAGAAGAAGATT | 60548156 |
| 26 | NC000018.10 chr18 | 0.82 | GAAGAAGATAGGCAAGAAGAAGATT<br>GAAGAAGAAGCAAGAAGAAGATT | 783990 |
| 27 | NC000002.12 chr2 | 0.82 | GAAGAAGAGGAAGAAGCAAGCATT<br>GAAGAAGAGCAAGAAGAAGATT | 42576014 |
| 28 | NC000012.12 chr12 | 0.82 | GAAGAAGAGAAAGAAGAAGAAGATT<br>GAAGAAGAGCAAGAAGAAGAAGATT | 58338875 |
| 29 | NC000012.12 chr12 | 0.82 | GAAGAAGAGAAAGAAGAAGAAGATT<br>GAAGAAGAGCAAGAAGAAGAAGATT | 57133351 |
| 30 | NC000012.12 chr12 | 0.82 | GAAGAAGAGAAAGAAGAAGAAGATT<br>GAAGAAGAGCAAGAAGAAGAAGATT | 23370439 |
| 31 | NC000014.9 chr14 | 0.82 | GAAGAAGAGAATGAAAGAAGAAGACT<br>GAAGAAGAGCAAGAAGAAGAAGATT | 12719462 |
| 32 | NC000014.9 chr14 | 0.82 | GAAGAAGAGAAAGAAGAAGAAGATT<br>GAAGAAGAGCAAGAAGAAGAAGATT | 3784429 |
| 33 | NC000002.12 chr2 | 0.82 | GAAGAAGAGAAAGAAGAAGAAGATT<br>GAAGAAGAGCAAGAAGAAGAAGATT | 135101854 |
| 34 | NC000005.10 chr5 | 0.82 | GAAGAAGAGCAAGAAGAAGGATT<br>GAAGAAGAGCAAGAAGAAGGATT | 17629750 |
| Table 6: Potential target regions in the human genome for svRNA2. |
|---------------------------------------------------------------|
| **Gene** | **Chromosome** | **Score** | **Target Sequence** | **Coordinates** |
|---------|----------------|-----------|---------------------|----------------|
| NC000018.10 | chr18 | 0.82 | GAAGAAGTAGGCAAGAAAGGATT | 783991 |
| NC000003.12 | chr3 | 0.82 | GAAGAAGAAG-AGGAAAGAAGATCT | 27102255 |
| NC000003.12 | chr3 | 0.82 | GAAGAAGAAGAAGAAAGAAGAT-T | 56345429 |
| NC000003.12 | chr3 | 0.82 | GAAGAAGAAGAAGAAAGAAGAT | 60432275 |
| NC000004.12 | chr4 | 0.82 | GAAGAAGAGGAAGAAGAAGAT | 117558369 |
| NC000016.10 | chr16 | 0.82 | GAAGAAGAAGAAGAAGAAGAT | 37420425 |
| NC000004.12 | chr4 | 0.82 | GAAGAAGAGCAAGAAAGAAGAAAG | 70599182 |
| NC000004.12 | chr4 | 0.82 | GAAGAAGAGCAAGAAAGAAGA-A | 70599181 |
| NC000004.12 | chr4 | 0.82 | GAAGAAGAGCAAGAAAGAAGA– | 70599180 |
| NC000002.12 | chr2 | 0.82 | GAAGAAGAGGAAGAAGAAGA-T | 1886808 |
| NC000002.12 | chr2 | 0.82 | GAAGAAGAGGAAGAAGAAGATG | 1886809 |
| NC000002.12 | chr2 | 0.82 | GAAGAAGAGGAAGAAGAAAGAT | 70555008 |
| Gene ID      | Score | Matched Sequence                                                                 | Loci     |
|-------------|-------|----------------------------------------------------------------------------------|----------|
| NC000002.12 | 0.89  | AGGACTACTGGGCGCCAGAGGCT  
              |       | AGGA–ACTGGGCGCCAGAGGCT             | 48975650 |
| NC000007.14 | 0.89  | AGGAACTGAGGGCCAGGAGGCT  
              |       | AGGAACTGAGGGCCAGGAGGCT             | 14547625 |
| NC000008.11 | 0.89  | AGGAACTGGGGCCAGAAGGCT  
              |       | AGGAACTGGGGCCAGAAGGCT             | 10614535 |
| NC000014.9  | 0.89  | AGGAACTGGGGCCAGAAGGCT  
              |       | AGGAACTGGGGCCAGAAGGCT             | 1419369  |
| NC000014.9  | 0.89  | AGGAACTGGGGCCAGAAGGCT  
              |       | AGGAACTGGGGCCAGAAGGCT             | 1419367  |
| NC000014.9  | 0.89  | AGGAACTGGGGCCAGAAGGCT  
              |       | AGGAACTGGGGCCAGAAGGCT             | 1419366  |
| NC000004.12 | 0.84  | AGGCAACTTGTCGCGCCAGAAGGCT  
              |       | AGGCAACTTGTCGCGCCAGAAGGCT         | 69436080 |
| NC000009.12 | 0.84  | AGGAACTGGGGCCAGAAGGCT  
              |       | AGGAACTGGGGCCAGAAGGCT             | 25416967 |
| NC000020.11 | 0.84  | AGGAGACCCTGGGGCCAGAAGGCT  
              |       | AGGAGACCCTGGGGCCAGAAGGCT          | 11374836 |
| NC000019.10 | 0.84  | AGGAACTGTCGCGAGAAGGCT  
              |       | AGGAACTGTCGCGAGAAGGCT             | 3570717  |
| NC000019.10 | 0.84  | AGGAACTGTCGCGAGAAGGCT  
              |       | AGGAACTGTCGCGAGAAGGCT             | 24050685 |
| NC000009.12 | 0.84  | AGGAACTGTCGCGAGAAGGCT  
              |       | AGGAACTGTCGCGAGAAGGCT             | 581979   |
| NC000011.10 | 0.84  | AGGAACTGTCGCGAGAAGGCT  
              |       | AGGAACTGTCGCGAGAAGGCT             | 22741205 |
| NC000011.10 | 0.84  | AGGAACTGTCGCGAGAAGGCT  
              |       | AGGAACTGTCGCGAGAAGGCT             | 21037552 |
| NC000012.12 | 0.84  | AGGAACTGTCGCGAGAAGGCT  
              |       | AGGAACTGTCGCGAGAAGGCT             | 11370121 |
| NC000012.12 | 0.84  | AGGAACTGTCGCGAGAAGGCT  
              |       | AGGAACTGTCGCGAGAAGGCT             | 486225   |
| NC000006.12 | 0.84  | AGGAACTGTCGCGAGAAGGCT  
              |       | AGGAACTGTCGCGAGAAGGCT             | 70506695 |
| NC000004.12 | 0.84  | AGGAACTGTCGCGAGAAGGCT  
              |       | AGGAACTGTCGCGAGAAGGCT             | 92640053 |
Table 7: Potential target regions in the human genome for svRNA3 (truncated version).

| # | NC000003.12 chr3 | 0.84 | AGGA ACTGAGGCTCAGAGAGCT | 33209680 |
|---|-----------------|------|--------------------------|----------|
|   | AGGA ACTG-GG-CAGA-AGCT |     |                          |          |
|---|-----------------|------|--------------------------|----------|
| 20 | NC000003.12 chr3 | 0.84 | AGTGAAGCATGGCCAGAAGCT    | 4730544  |
|   | AG-GA-A-TGGGCCAGAAGCT |     |                          |          |

AG-GA ACTG GGGCCAGAAGCT