

Discussion

In this work, we developed SEARCH-MaP and SEISMIC and applied them jointly to detect structural ensembles involving long-range RNA:RNA interactions in SARS-CoV-2 and other coronaviruses. This study is certainly not the first to perturb RNA structure with ASOs, nor even the first to use DMS-MaPseq to quantify the structural changes upon binding ASOs to SARS-CoV-2 RNA [1]. But while this previous study examined local structural perturbations caused by binding an ASO, we show that we can detect changes in the structure at more distant locations in an RNA molecule that interact with the nucleotides bound by an ASO.

SEARCH-MaP bears conceptual similarity to another method, mutate-and-map read out through next-generation sequencing (M2-seq) [2]. Both involve perturbing one region of an RNA molecule (in the case of M2-seq, by pre-installing mutations through error-prone PCR) and measuring the effects on other bases in the RNA using chemical probing. The major differences are the precision and scale of the interactions identified, as well as the throughput. M2-seq can pinpoint interactions down to the resolution of a single base pair, and is thus more precise than SEARCH-MaP. However, DMS-guided RNA structure prediction can propose structure models at single-base-pair resolution, which SEARCH-MaP can validate, and in this way achieve single-base-pair resolution. SEARCH-MaP is also capable of finding interactions over a much longer range because M2-seq requires the interacting bases to be in the same Illumina sequencing read. Within this length limit, one M2-seq experiment can theoretically find all pairwise interactions between bases, while one SEARCH-MaP experiment can find only interactions that involve the region to which the ASOs were hybridized. M2-seq is also limited by the formation of alternative structures. Some methods, such as [CITE something by Rhiju, maybe REEFIT] and DANCE-MaP [CITE], have been designed to work around this limitation SEARCH-MaP; however, [something by Rhiju] has [this problem], and DANCE-MaP requires extremely high sequencing depth of several million reads [MORE PRECISE]. SEARCH-MaP, by contrast, assumes from the start that the RNA may

form alternative structures; for simply detecting long-range interactions, even a 5,000 read depth is sufficient coverage; and for clustering, we have found [SOME LIMIT].

Another limitation of SEARCH-MaP as presented here is that it cannot distinguish between direct and indirect interactions. If RNA segment A interacts with segment B, while B interacts with both segment A and C, then hybridizing an ASO to segment A would perturb the structure of B, which could consequentially perturb the structure of C. Hence, C would appear to interact with A, even though this interaction is indirect, through B. One possible workaround (not shown in this study) would be to mutate or hybridize an ASO to segment B, and then repeat the experiment with hybridizing an ASO to segment A. If the interaction between A and C is direct, then C should still be perturbed even when segment B is incapable of interacting with A or C. But if B mediates an indirect interaction between A and C, then disrupting B should eliminate the apparent interaction between A and C.

Functional long-range interactions up to four kilobases involving an FSE have been found previously in two plant viruses ???. In both cases, frameshifting required the long-range interaction, suggesting that this interaction enables negative feedback on synthesis of viral RNA polymerase ?. When polymerase levels are low, the interaction would form and stimulate frameshifting, which is needed to synthesize RNA polymerase. Once the polymerase had accumulated, it would begin to replicate the genomic RNA; in its passage from the genomic 3' end to the 5' end, it would disrupt the 3' side of the long-range interaction, attenuating frameshifting and reducing synthesis of more polymerase.

However, this strategy cannot be the role, if any, of the long-range interactions in coronaviruses. Unlike in the two plant viruses, a long-range interaction is not required to stimulate frameshifting in coronaviruses: numerous studies have shown that even the isolated FSE can cause 15 - 40% of ribosomes to frameshift ???????. In coronaviruses, the long-range interaction is not only unnecessary for frameshifting but also may even attenuate it, given that in SARS-CoV-2, the FSE-arch and the frameshift-stimulating pseudoknot seem to be mutually exclusive. Moreover, coronaviruses partition translation and RNA synthesis

into two different cellular compartments (the cytosol and the double-membrane vesicles, respectively) ?, so structural changes induced by RNA polymerases would not be seen by ribosomes.

The functions of these long-range interactions involving the FSE in coronaviruses remain mysterious. However, given that they occur in multiple coronaviruses across at least two genera, it seems reasonable that they could play a role in the viral life cycle, possibly by affecting the rate of frameshifting. Further research may reveal new mechanisms of translational regulation in coronaviruses via long-range RNA:RNA interactions.