|  |
| --- |
| Referees 1 and 2 - Point 1  Ref1:  It is important to provide a fair estimate for the likelihood that a random gene pair can be efficiently targeted by a multi-target synthetic miRNA. If only a small fraction of gene pairs share efficient synthetic miRNA target sites this means that the proposed approach has little general utility. In the manuscript the authors provide a rather optimistic (upper bound) estimate of this likelihood by randomly sampling 10,000 gene pairs and counting the fraction of pairs that share a 7mer sequence. The main problem with this approach is that only a fraction of 7mers are efficient target sites in a given 3'UTR sequence. Indeed, this is exactly the concept of their proposed method, that information such as AU-context and RNA accessibility can predict efficient target sites. The authors should therefore filter week 7mer target sites (eg. sites with low AU-context and low accessible) before computing the fraction of gene pairs that share a 7mer. Additionally, it is not clear if the authors use all possible 7mers in this analysis, as it is obvious that certain 7mer sequences may strongly bias this analysis. For example, the polyA-signal motifs (eg. AAUAAA and AUUAAA) are ubiquitous in mRNA 3'UTRs and including 7mers with this motif in the analysis will greatly inflate the estimate. Furthermore, in an unbiased analysis the authors should also exclude homopolymer motifs (i.e. AAAAA) and known target seed sequences for evolutionarily conserved miRNAs (also an option in their web service).  Ref 2:  Applicability. One drawback of the approach used in the paper is the difficulty of finding microRNA seeds present in several 3’ UTRs. This is discussed in the paper, where 64% of 3’ UTR pairs are estimated to share a 7mer (or a longer sequence). The corresponding number for 3’ UTR triples is 16%. This limits the practical usefulness of the method. |
| Answer:  Following the comments and suggestions of both referees, we refined our analysis of the applicability and usefulness of miR-Synth. The basic idea behind multiple targeting is that the target sequences share some similarities, which can become binding sites for a unique a-miR seed. Although given a random pair, triplet or, more in general, n-tuple of genes, is not always possible to create an a-miR which is able to bind them all, there are many other cases in which many a-miRs can be created for a given set of targets. We collected gene expression data associated to different disease conditions from the Gene Expression Atlas (<http://www.ebi.ac.uk/gxa>) and focused on the up-regulated genes, which represent typical plausible targets for an artificial miRNA. For each condition, we calculated all the possible combinations of two and three up-regulated genes and counted of how many of them share 7mer 3' UTR sites. We filtered out polyA-signal motifs, homopolymer motifs and sites matching the seeds of endogenous miRNAs. We were able to perform this analysis for all pair and triplets of up-regulated genes related to 83 different diseases and the results show that 97.3% of pairs and 81.32% of triplets share at least one 7mer site. On average, pairs and triplets shared about 136 and 24 7mer sites, respectively. For 96 additional diseases we were able to perform the pairs analysis and the cumulative results confirmed the previous tests, with 97.13% pairs sharing on average 132 7mer sites. We couldn't perform the triplets analysis on these additional diseases due to the high number of up-regulated genes which generated a huge number of triplets which were too computationally expensive to analyze. The results of these tests are briefly reported in the manuscript (changes are highlighted) and details are provided as supplementary material.  miR-Synth is the only tool available on-line which allows the design of multi-site and multi-target artificial miRNAs. In light of this and of the results of the above analysis, we believe that it represents a resource which is potentially applicable to a significant number of cases where two or more genes must be simultaneously repressed. |
| Referee 1 - Point 2  It is confusing that the authors use the term 'miRNAs' for their computationally derived synthetic miRNA sequences. I strongly suggest the use of an appropriate acronym, e.g. synMir or aMir |
| Answer  We now use the term a-miR to refer the artificially designed miRNAs, as suggested by the referee. We changed the manuscript accordingly (changes are highlighted). |
| Referee 1 - Point 3  It seems odd that the algorithm predicts 125 synthetic miRNAs that could potentially target both c-MET and EGFR (p. 8). What fraction of these 125 synthetic miRNAs target distinct/unique site? And of the 6 high scoring miRNAs, what fraction of these correspond to unique sites? |
| Answer  miR-Synth uses shared 6mer/7mer sequences as binding sites for the a-miR seed sequences, thus for each shared seed site only one amiR will be created. The 125 a-miRs designed for both c-MET and EGFR target all distinct sites and are based on 125 distinct shared 7mers. Some sites are slightly overlapping, but their sequences are all different, being different the seed regions that they are designed for. |
| Referee 2 -  Point 2  Off target effects. These have not been analyzed, but such analysis could improve the manuscript. One way to show the benefits of using a single microRNAs to repress multiple targets, would be to demonstrate that this reduces off target effects compared to using siRNAs. This could be done using a genome wide method, i.e. RNA-seq. |
| Answer  Off-target effects represent an important issue related to artificial gene silencing and we agree with the referee that such analysis would improve the manuscript. Although genome-wide methods, such as microarray and RNA-seq, could be a valid way to evaluate the impact of small interfering RNAs, they don't allow to distinguish between direct and indirect effects. Moreover, since a-miRs mimic the way of action of endogenous miRNAs, there could be cases in which the mRNA levels don't reflect the protein outcome, whereas siRNAs almost always promote degradation of mRNAs. [Dennis doesn’t understand this sentence. Could there be a typo?]  Nevertheless we wanted to investigate the potential differences of the impact of a-miRs and siRNAs on the transcriptome, thus we performed an *in-silico* analysis of off-target effects. Evidence shows that siRNAs may act like endogenous miRNAs and repress unintended targets by perfect complementarity of their seed region. Since siRNAs are designed to target a single gene, the simultaneous repression of c-MET and EGFR would take at least two different siRNAs. Thus, we designed siRNAs for c-MET and EGFR by using three different tools based on different algorithm and features. For each tool and each target, we chose the top 6 siRNAs that didn't share the seed sequence with any endogenous miRNA. We considered all the possible pairs of c-MET/EGFR siRNAs from each tool separately and, for each pair, calculated the number of potential off-target genes by matching their seed sequences with the whole 3' UTRome (34619 sequences). Although seed match is not always a reliable indicator of an actual interaction, most functional interactions occur through perfect seed pairing. We considered both 6mer and 7mer matches and compared these off-targets numbers with those predicted for the top 6 a-miRs designed by miR-Synth. On average, a-miRs had about 8911 off-target 6mer and 3108 7mer matches as opposed to 13190 off-target 6mer and 6067 off-target 7mer matches for siRNAs. Figure xx included in supplementary material shows the distribution of the off-targets, and clearly demonstrates that double-targeting a-miRs are likely to have less off-target effects than pairs of single-targeting siRNAs. This test is also briefly described in the manuscript (changes are highlighted). |
| Referee 2 - Point 3  Finding seed sequences that are present in a set of 3’ UTRs and absent from another set is a trivial exercise. The miR-Synth method refines this approach by taking other features into account (nucleotide composition etc.), but it is not clear how much this refinement improves the designed microRNAs. The validation using the miRTarBase showed a statistically significant increase in performance, but the manuscript should also show how large this increase was on average (a 1% increase might be statistically significant, but not practically meaningful). Ideally, experimentally testing the effect of lower scoring microRNA candidates (that still have the same length seed sequences) would demonstrate the additional benefits of the miR-Synth method. Although I don’t think this is required, it would give strong evidence that miR-Synth works well for design of synthetic microRNAs. |
| Answer  The evaluation of the miR-Synth scoring function performed on validated miRNA-target pairs retrieved from miRTarBase showed that, on average, 65% of the top 10 interactions as ranked by miR-Synth were true direct interactions (+15% compared to the 50% of random samples), when considering two classes (direct evidence of interaction, no interaction). Considering three classes (direct evidence of interaction, indirect evidence, no interaction), 58% of the top 10 interactions as ranked by miR-Synth were true direct interactions (+25% compared to the 33% of random samples).  We also followed the suggestion of the referee and synthesized and tested the bottom 6 a-miRs for MET and EGFR according to miR-Synth ranking (120th to 125th ranking a-miRs). We found that 3 of these a-miRs yielded a mild repression of EGFR, lower than observed with the best top 6 a-miRs, and that none of them was able to significantly repress c-MET. It is also important to mention that all the tested a-miRs have 7mer-m8/8mer seed matches, as for all the a-miRs in the list. This test is briefly described in the manuscript (changes are highlighted) and more details are reported as supplementary material. |
| Refeee 2 - Point 4  Some validation experiments are done in HeLa cells (qt-RT-PCR and westen blot), other in HEK-293A cells (luciferase reporter assays). What is the reason for this? |
| Answer  In our laboratory we have an optimized protocol for performing luciferase assays on HEK-293A cells. These cells, indeed, constitute an optimal system for co-transfection of miRNAs and luciferase vectors and yield high efficiency at 24h. The analysis of the endogenous proteins through Western Blot require longer time points (> 48h). The fast growth rate of HEK-293A cells, together with their tendency to easily detach from the culture plates, make them a less reliable system for such analysis, for which we prefer to use HeLa cells, which are more stable at longer time points. |
| Referee 2 - Point 5  page 5, lines 42-43: “...considering the fact that groups with more than 5 well conserved sequences might share at most one common 7nt pattern or one or more 6nt sites...” I don’t understand this statement. |
| Answer  We replaced that part with a new section describing the new tests performed on pairs and triples, so that statement is no longer present in the manuscript. |
| Referee 2 - Point 6  page 8, line 22: Should “Fig. 4e,f” be “Fig. 4e,f,g”? |
| Answer  We made the correction on the manuscript. |