NAR Genomics and Bioinformatics, 2025, **7**, lqaf064
https://doi.org/10.1093/nargab/lqaf064
Methods Article

dsRNAmax: a multi-target chimeric dsRNA designer for safe and effective crop protection

Stephen J. Fletcher[1,2,*], **Jai Lawrence**[1], **Anne Sawyer**[1], **Narelle Manzie**[1], **Donald M. Gardiner**[1], **Neena Mitter**[1,2], **Christopher A. Brosnan**[1,*]

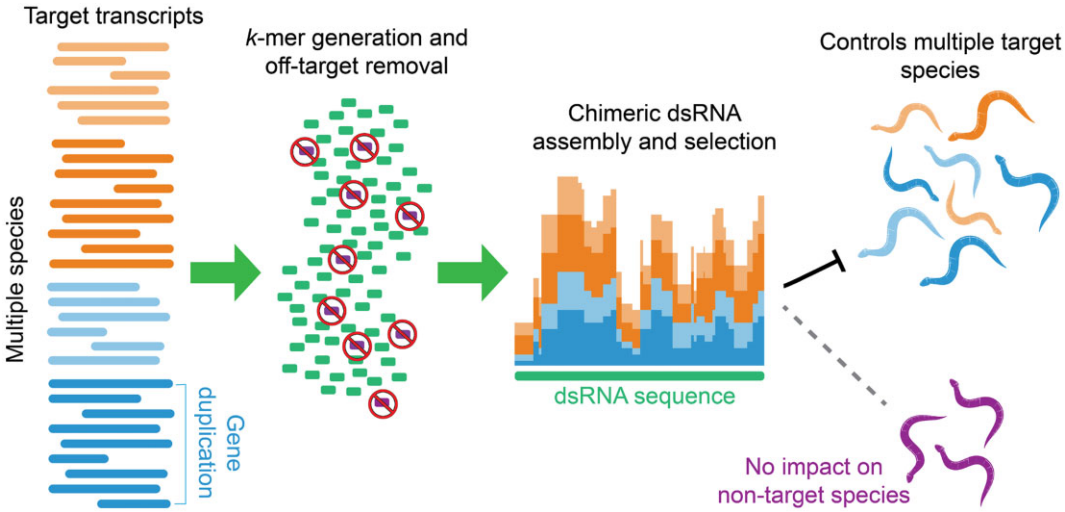
[1]Centre for Horticultural Science, Queensland Alliance for Food and Agriculture Innovation, University of Queensland, St Lucia 4072, Queensland, Australia
[2]Charles Sturt University, Wagga Wagga 2678, New South Wales, Australia

[*]To whom correspondence should be addressed. Email: s.fletcher@uq.edu.au
Correspondence may also be addressed to Christopher A. Brosnan. Email: c.brosnan@uq.edu.au

Abstract

Crop protection is undergoing significant evolution, transitioning towards sustainable approaches that minimize impacts on the environment and human health. Exogenous application of double-stranded RNA (dsRNA) that silences pest or pathogen genes via RNA interference (RNAi) has promise as a safe and effective next-generation crop protection platform without the need for genetic modification. However, exogenous dsRNA application at scale presents challenges. Specifically, a single dsRNA sequence needs to balance targeting the standing variation in a target pest or pathogen group against the potential for adverse impacts in a vast array of non-target and beneficial organisms at the application site and broader environment. To address these competing demands, we present dsRNAmax (https://github.com/sfletc/dsRNAmax), a software package that employs *k*-mer-based assembly of chimeric dsRNA sequences to target multiple related RNA sequences, to broaden the target spectrum. The package ensures that designed dsRNAs have no defined contiguous sequence homology with any off-target sequences, which can range from single transcriptomes through to metagenome sequence data and beyond. The efficacy of this package is demonstrated by a dsRNAmax-designed dsRNA that inhibits multiple root-knot nematode species but not a non-target nematode species, despite its susceptibility to environmental RNAi and high homology of the target gene.

Graphical abstract



Introduction

Pests and pathogens pose a critical threat to global food security, causing annual crop production losses of up to 40% and resulting in worldwide economic damages of ~US$220 billion [1]. Crop protection strategies have historically relied on a combination of host genetic resistance, agronomic management practices, and synthetic pesticides. However, the widespread deployment of synthetic pesticides has raised significant concerns due to their often adverse effects on human health and ecosystem stability, with negative impacts

Received: January 30, 2025. **Revised:** April 16, 2025. **Editorial Decision:** May 4, 2025. **Accepted:** May 7, 2025

throughout agricultural supply chains [2]. In response to these challenges, exogenous RNAi (RNA interference)-based crop protection strategies are emerging as promising alternatives. Topical application of double-stranded RNA (dsRNA) exploits the cellular mechanism of RNAi to target pest and pathogen species with high specificity, while avoiding many of the regulatory and public acceptance challenges associated with genetic modification [3, 4].

RNAi is a conserved mechanism in most eukaryotes, with roles in virus defence, gene regulation, and maintenance of genome stability. The RNAi pathway is activated by dsRNA, which is processed into 20–30 nt small interfering RNAs (siRNAs), depending on the organism [5]. When loaded onto ARGONAUTE (AGO) proteins, siRNAs direct cleavage and silencing of complementary RNAs. This pathway can be co-opted via exogenous application of dsRNA to specifically knock down expression of essential pest or pathogen genes, causing mortality in target organisms—a process that is effective against viral, fungal, oomycete, insect, and nematode targets [6]. A major advantage of this narrow-spectrum approach is that the sequence specificity of RNAi should mitigate unintended impacts on off-target species such as beneficial insects, even if they come into direct contact with the applied dsRNA.

To realize the safety benefits of RNAi-based pesticides, effective design of pest and pathogen targeting dsRNAs is critical. Multiple factors are at play. Firstly, there can be significant genetic variation among target organisms. Variation among viral isolates of the same species serves as a prominent example. Duplication and sequence divergence of functionally comparable target genes can also occur in the same organism. An effective dsRNA must account for this target sequence variation. On the other hand, topical dsRNA application via spraying means potential contact of the dsRNA with a considerable array of non-target organisms, both at the application site and in the surrounding environment. Mitigating unintended impacts, especially when a non-target species is related to the target pest or pathogen, is essential for reaping the environmental and safety benefits RNAi-based biopesticides offer. An example here is avoiding beneficial insects that have been deployed as an IPM (integrated pest management) strategy while targeting specifically a pest. To reduce barriers to adoption including dsRNA synthesis costs and regulatory hurdles relating to multiple active ingredients, application of a single highly efficient dsRNA sequence that targets multiple related pests or pathogens is strongly desired.

Although dsRNA design software packages have been available for decades, they do not adequately take into consideration the additional requirements that exogenous RNAi-based crop protection brings, including variation among sometimes large numbers of target sequences, and complex large-scale off-target requirements. Here, we present dsRNAmax (dsRNA maximizer), an assembly-based chimeric dsRNA designer that maximizes the number of dsRNA-derived siRNAs matching target RNAs, while eliminating all contiguous matches to any user-defined off-target sequences, whether they be transcriptomes, genomes, or unassembled environmental sample sequences. We used dsRNAmax to design a chimeric dsRNA targeting multiple plant parasitic nematode species but not a designated off-target nematode, and experimentally demonstrated multi-species targeting without off-target impacts.

Materials and methods

dsRNAmax development

dsRNAmax is written in Go (Golang), with binaries available for all major operating systems at https://github.com/sfletc/dsRNAmax. It makes use of Go's concurrency features, efficiently utilizing available CPU cores. Memory usage scales with input size, with most dsRNA design tasks able to be completed on commonly available consumer hardware such as laptops and desktops.

Data input format

dsRNAmax takes inputs for target sequences in FASTA format and off-target sequences in either FASTA or .kmer format, which is generated using the SeqToKmer package (https://github.com/sfletc/SeqToKmer). When very large off-target datasets are used (e.g. paired-end FASTQ files generated from environmental sequencing), the .kmer conversion approach allows for filtering of low-count *k*-mers prior to usage, faster execution, and more compact storage. Low-count *k*-mers can arise from sequencing errors rather than accurately reflecting underlying sequence variation. When one input sequence is a priority, a bias selection and amount can be specified.

dsRNAmax framework

For effective dsRNA design based on the ability of processed siRNAs to direct silencing, a *k*-mer approach was adopted (Fig. 1). Unique sense-orientation *k*-mers from all input sequences are generated along with a count of input sequences they appear in. The default *k*-mer size of 21 represents the most abundant class of siRNAs in many organisms, although this size is user-selectable. *k*-mers that have a contiguous match to any off-target sequence in either orientation are removed from this pool. One *k*-mer is randomly selected as an initiating seed, and then upstream and downstream extension takes place one nucleotide at a time by selecting the most abundant overlapping *k*-mer remaining in the pool, where abundance is measured by the number of input target sequences the *k*-mer exactly matches. Each time a *k*-mer is selected, it is removed from the pool. Where multiple overlapping *k*-mers have an equal abundance, one is randomly selected. Extension in both directions continues until no overlapping *k*-mers are present. If this sequence is longer than the selected dsRNA construct length (default = 300 nt), a dsRNA of the defined length is selected by maximizing the median of *k*-mer counts matching each target input sequence. Due to inherent random selections including the seed *k*-mer, multiple iterations are conducted to identify a more optimal solution (default = 100×). The user can increase the number of iterations for large, complex input datasets.

For each input target sequence, statistics including the number of matching *k*-mers, Smith–Waterman–Gotoh similarity, mean *k*-mer GC content, and matching *k*-mer percentages with 5′U, 5′A, and 5′C are presented. dsRNAmax includes the ability to bias the algorithm towards a priority input, which may be desirable where particular biotypes are more important in target population(s). This increases the number of *k*-mers matching the priority sequence but may reduce the median of *k*-mers matching the pool of target sequences. In all cases, no derived *k*-mer in any designed construct will perfectly match any off-target sequence, which is an important

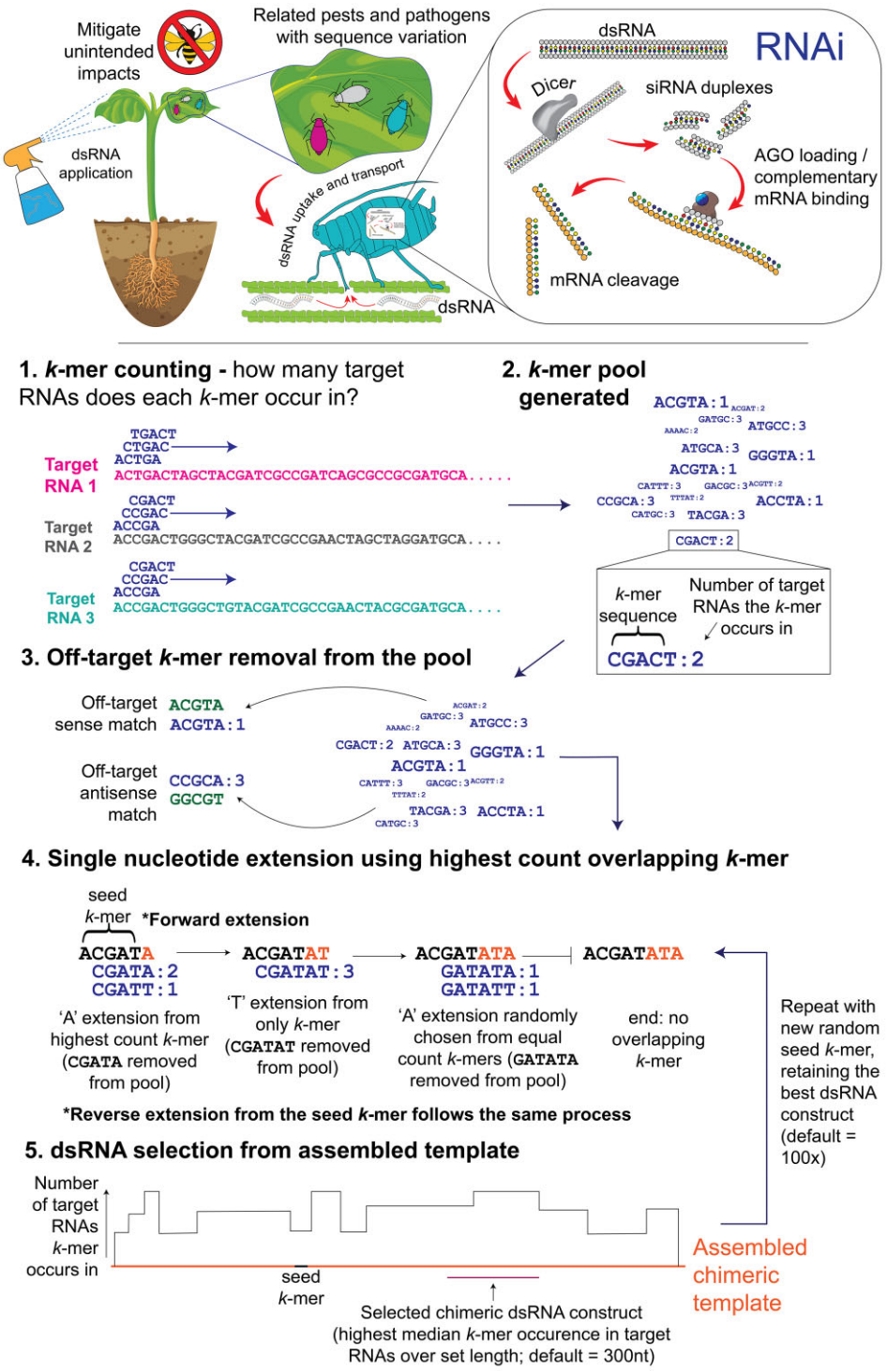


Figure 1. dsRNAmax generates a dsRNA construct targeting multiple sequences while mitigating off-target impacts. (Upper) Ingestion/uptake of dsRNAs leads to induction of RNAi, where the dsRNA is processed to siRNAs, which are loaded into ARGONAUTEs (AGOs) and direct cleavage of complementary host RNAs. (Lower) Example dsRNAmax schematic targeting three RNA targets. (1, 2) All unique *k*-mers from the target sequence are added to a pool, along with a count of the number of target RNAs they occur in (default = 21 nt; 5 nt shown here for illustration). Occurring more than once in the same target sequence does not add to the count. Bias can be provided by adding additional copies of a selected sequence at this step. (3) Any *k*-mer with a contiguous match (default = 21 nt; 5 nt shown here for illustration) to an off-target sequence in either orientation is removed from the pool. (4) A seed *k*-mer is selected randomly, followed by bidirectional extension a single nucleotide at a time. Forward extension is shown here. In each extension instance, three possibilities exist: (i) the overlapping *k*-mer with the highest count provides the next nucleotide (i.e. its 3′ nucleotide for 3′ extension) followed by removal from the pool; (ii) if multiple overlapping *k*-mers have an equal count, one is randomly selected to provide the extension nucleotide and then removed from the pool; or (iii) if no overlapping *k*-mer exists, extension ceases. Once extension has ceased in both directions, assembly of the chimeric template is complete. (5) The final dsRNA sense arm sequence is selected by scanning the chimeric template in a window (default = 300 nt) and calculating the median number of *k*-mers exactly matching each target sequence within the window. The sequence with the greatest median is then selected. The seed *k*-mer is not necessarily in the final dsRNA. To account for sub-optimal random selections, steps 4 and 5 are repeated by default 100 times, with the best dsRNA sequence retained each time.

consideration for mitigating unintended impacts [7]. In situations where this is not possible, no dsRNA will be designed.

dsRNA design for root-knot nematode (*Meloidogyne*) species

Genome assemblies for four root-knot nematode (RKN) species, *Meloidogyne arenaria*, *Meloidogyne enterolobii*, *Meloidogyne incognita*, and *Meloidogyne javanica*, were accessed from NCBI (GCA_017562155.1, GCA_903994135.1, GCA_014132215.1, and GCA_900003945.1, respectively) with structural annotations performed using Braker3 [8] and functional annotations via BLAST against UniRef90 [9] and HMMER (www.hmmer.org) against Pfam 37.1 [10].

To demonstrate the biological efficacy of a single construct against multiple species in *in vitro* bioassays, we added all *translation elongation factor 1a* (*TEF1a*) transcript copies annotated in the RKN genomes to a FASTA file as the target input. For the off-target nematode species, *Caenorhabditis elegans* was chosen. The latest genome assembly of this species in FASTA format (GCA_000002985.3) was used as the off-target input, with the genome-derived CDS (coding DNA sequence) annotation used for comparison. All other dsRNAmax settings except for the off-target setting were left at default. No bias towards any input sequence was used. Two dsRNA sequences were generated, one with the offTarget flag set to 17 (TEF-17; maximum contiguous off-target match allowed is 16 nt) and the other with the offTarget flag set to 21 (TEF-21; maximum contiguous off-target match allowed is 20 nt) (Supplementary Table S1 and Supplementary Fig. S1). Setting the offTarget flag lower than 17 nt resulted in no dsRNA being designed due to homology to the *C. elegans* genome.

In addition to generating the dsRNAs for performing on-target and off-target bioassays, we generated three additional dsRNA sense arm sequences to demonstrate the impact of not including any off-target sequences, using the *C. elegans* off-target genome, but biasing the dsRNA towards one *TEF* sequence, *M. enterolobii* g3046.t1, or using the *C. elegans* CDS dataset instead of the genome. These example dsRNAs were not used in *in vitro* assays.

dsRNA generation

Chimeric DNA templates were synthesized by Integrated DNA Technologies. dsRNA was generated using a HiScribe T7 transcription kit (New England Biolabs). For *C. elegans* targeting, dsRNA was synthesized from complementary DNA generated from N2 mixed stage nematodes. Oligonucleotides used to generate dsRNA are listed in Supplementary Table S2.

Meloidogyne spp. growth and maintenance

Meloidogyne spp. nematodes were provided by Queensland Department of Primary Industries, who maintained colonies on tomato roots (cv. Tiny Tim).

Meloidogyne spp. bioassays

Juvenile 2 (J2) stage *M. incognita*, *M. javanica*, or *M. arenaria* were used for all experiments. J2 stage nematodes were hatched from eggs harvested from tomato roots. Nematodes were treated with dsRNA of the indicated sequence and concentration in 50 or 100 µl supplemented with 10 mM octopamine hydrochloride (Sigma–Aldrich) in 1.5-ml Eppendorf tubes. Treated nematodes were incubated overnight at room temperature with survival assessed after 24 h under a dissecting microscope. To assess survival, each worm was observed for 5–10 s for any body movement with those that lacked any movement counted as dead. A minimum of 10 animals were counted per experimental condition.

Caenorhabditis elegans growth and maintenance

Caenorhabditis elegans culture and maintenance were performed using standard techniques [11]. Nematodes used were wild-type N2 and grown on NGM plates seeded with OP50 *Escherichia coli* as a food source.

Caenorhabditis elegans bioassays

Assessment of dsRNA targeting *C. elegans* was performed using a brood assay. Synchronized L4 stage worms were incubated in a solution containing 50 ng/µl of the indicated dsRNA (GFP, TEF-17, TEF-21, or ceTEF) for 24 h. Three nematodes per biological replicate and treatment were transferred to seeded NGM plates and allowed to mature to egg-laying adults. These were left for 4 h to lay eggs and then removed. The number of hatched L1/L2 nematodes was then counted.

Benchmarking dataset

To benchmark dsRNAmax with larger datasets, we downloaded 339 full-length cucumber mosaic virus (CMV) RNA3 genome sequences from NCBI GenBank. This dataset totalled 747 093 nt in length. For a benchmark metagenome off-target dataset, we downloaded paired-end FASTQ files from the NCBI Sequence Read Archive run SRR31016968 (Shotgun metagenome of non-amended cucumber rhizospheres:FORC-inoculated). This run had a total of 7.7G of bases. These read files were concatenated together and then converted to a .kmer file using SeqToKmer, with the *k*-mer length set to the default of 21 and the minKmerCount filter set to 2. This file was then used as an off-target input to dsRNAmax using the -offTargetKmers flag.

Statistical analysis

Statistical analysis of RKN kill curve data was performed using Fisher's exact tests on 2 × 2 matrices of alive/deceased counts for each pairwise comparison. Generated *P*-values were adjusted for multiple comparison using false discovery rate (FDR). *Caenorhabditis elegans* brood size tests were performed using pairwise *t*-tests followed by FDR multiple *P*-value correction (independent broods were derived from three egg-laying adults).

Results

Validation study: targeting four *Meloidogyne* spp. while avoiding *C. elegans*

To demonstrate the ability of dsRNAmax to design single chimeric dsRNA constructs against multiple targets while avoiding unintended impacts on non-target species, we selected four RKN species (*M. incognita*, *M. arenaria, M. javanica*, and *M. enterolobii*) as targets. RKNs are small thread-like roundworms ∼0.5 mm in length [12]. They invade the roots of many crops, including banana, tomato, carrot, and other horticultural and ornamental species, causing significant yield

loss. *Meloidogyne* species are ideal for use in validation studies due to their availability, likely effectiveness in *in vitro* dsRNA delivery assays, and importance as crop pests [12]. *Meloidogyne enterolobii* was used exclusively for *in silico* design as it is a biosecurity risk in Australia and could not be included in experiments.

We chose *C. elegans* as an off-target species for several reasons. Though not considered a beneficial organism, *C. elegans*' susceptibility to environmental RNAi and well-defined *in vitro* dsRNA response meant that a dsRNA too closely matching its own transcripts would likely have a negative phenotypic impact, making it an ideal control case for an RKN-targeting chimeric dsRNA.

We reasoned that targeting a constitutively expressed, highly conserved, and essential gene would be the best way to test dsRNAmax's design ability without other potential ectopic effects interfering with our interpretations. For this reason, we chose *TEF1a*. *TEF1a* is a highly conserved nucleotide-binding protein involved in the elongation of polypeptide chains. *TEF1a* was previously shown to be an effective exogenous dsRNA target in *Austropuccinia psidii* (myrtle rust) [13]. This target conservation increases the likelihood of generating a single chimeric dsRNA that targets multiple species but requires careful design to avoid negative impacts on related organisms.

Multiple copies of putative *TEF1a* transcripts were annotated in each RKN species: six copies in *M. arenaria* and *M. incognita* and nine copies in *M. javanica* and *M. enterolobii*. The functional redundancy of these transcripts is unknown; therefore, all 30 transcripts were used as inputs to the dsRNAmax design software. The percent identity between pairwise matches of these target RNAs ranged from 86.5% to 100% (Supplementary Fig. S2). As the degree of acceptable homology to *C. elegans* transcripts was also unknown, we designed two dsRNA constructs, selecting either 21 nt (TEF-21) or 17 nt (TEF-17) as the cut-offs for the maximum contiguous homology to the *C. elegans* genome. As expected, by reducing the allowable contiguous homology to the *C. elegans* genome from 20 nt (TEF-21) to 16 nt (TEF-17), the median number of 21 nt matches to all *TEF1a Meloidogyne* spp. transcripts was also reduced (from 200.5 to 185) (Fig. 2A and B). Pairwise alignments of each construct to the *C. elegans EF1-a* transcript are shown in Supplementary Fig. S3, with aligned region homologies of 79.3% and 85.1%, respectively.

To test the ability of our software to specifically target multiple RKN species, we treated three selected species, *M. arenaria*, *M. incognita*, and *M. javanica*, with varying concentrations of TEF-21 and TEF-17 dsRNAs, along with a non-target/negative green fluorescent protein (GFP) dsRNA control, in *in vitro* soaking assays. After 24 h of soaking, RKNs were scored for alive and deceased counts on plates. Kill curves for each RKN species using a two-fold dilution series of dsRNA concentrations from 50 to 6.25 ng/μl are shown in Fig. 2C–E. Example images of alive and deceased morphology are shown in Fig. 2F. A concentration of 50 ng/μl of either RKN targeting dsRNA was required for complete mortality in all RKN species (Fig. 2C–E and Supplementary Tables S3–S5). Concentrations of 25 and 12.5 ng/μl had partial levels of mortality, while concentrations of 6.25 ng/μl were ineffective. The *GFP*-targeting dsRNA had no negative impact on RKN health, demonstrating that TEF-21 and TEF-17 were acting in an RNAi-dependent sequence-specific manner.

An important design factor in this example is the avoidance of negative impacts on our selected off-target species, *C. elegans*. For the off-target assays, a positive control dsRNA exactly matching the *C. elegans EF1-a* transcript was used (CeTEF). None of the dsRNAs had negative impacts on *C. elegans* mortality upon soaking, but the impact on the subsequent generation as evidenced by brood size was clear (Fig. 2G). The CeTEF dsRNA resulted in near-zero brood sizes when compared with the negative control GFP dsRNA treated samples and TEF-17 and TEF-21 treatments. Interestingly, there was a significant difference between the *GFP* non-target control and the TEF-21 treatment, with a very slight decrease in brood size. This was not seen with the TEF-17 treated samples, matching the stringency of the targeting predictions. While there was no significant difference between TEF-21 and TEF-17 dsRNA impacts on *C. elegans* brood size, the individual comparisons to the negative GFP control trend with our *in silco* stringency predictions.

Impact of off-target and bias selections

In general, adding off-target sequences reduces the median number of *k*-mers matching each target input sequence if off-target *k*-mers are derived from related species and/or target genes are highly conserved. By reducing the pool of *k*-mers from which the pre-dsRNA sequences can be assembled, the likelihood increases that less abundant *k*-mers must be assembled. For example, if the *k*-mer that matches the most target input sequences is removed, the second most abundant overlapping *k*-mer is used for extension. Where the removal of a *k*-mer means that extension can no longer take place, a different region of the assembled sequence is used for the final chimeric dsRNA selection. Design against the *Meloidogyne* spp. *TEF1a* transcript set without the use of off-targets generated a median number of 21 nt matches to each transcript of 203 out of a maximum of 280 sense-orientation *k*-mers for a 300 nt construct (Supplementary Fig. S4). Addition of the *C. elegans* genome as an off-target dataset, and an off-target cut-off of 17 (i.e. there are no matches ≥17 nt between the dsRNA and the off-target dataset in either orientation), resulted in the median 21 nt matches to each target dropping to 185, with 1305 *k*-mers removed from the initial 8902 target pool (Supplementary Fig. S2). Using the *C. elegans* CDS dataset instead of the genome led to a very slight improvement with a 17 nt cut-off design, with median 21 nt matches increased to 186.5, though in this instance only 655 *k*-mers were removed from the pool (Supplementary Fig. S5). This off-target dataset also allowed for a 300 nt dsRNA design to complete with a 16 nt cut-off.

A further consideration of multi-target dsRNA design is the relative importance of individual targets within the target dataset. In some situations, a certain viral isolate or pathogen species may be considered of greatest importance for control. Given that a range of *k*-mer matches between the designed dsRNA construct and the target dataset may exist, biasing the design can ensure that a specified target sequence sits towards the top of this distribution. Biasing in this way can, however, reduce the median dsRNA-derived *k*-mer matches to the overall target dataset.

For chimeric dsRNA design against 30 *Meloidogyne* spp. *TEF1a* transcripts with a 17 nt *C. elegans* off-target selection, dsRNA-derived 21 nt matches to these targets ranged from 119 to 236, with the median 185 (Supplementary Fig. S1).

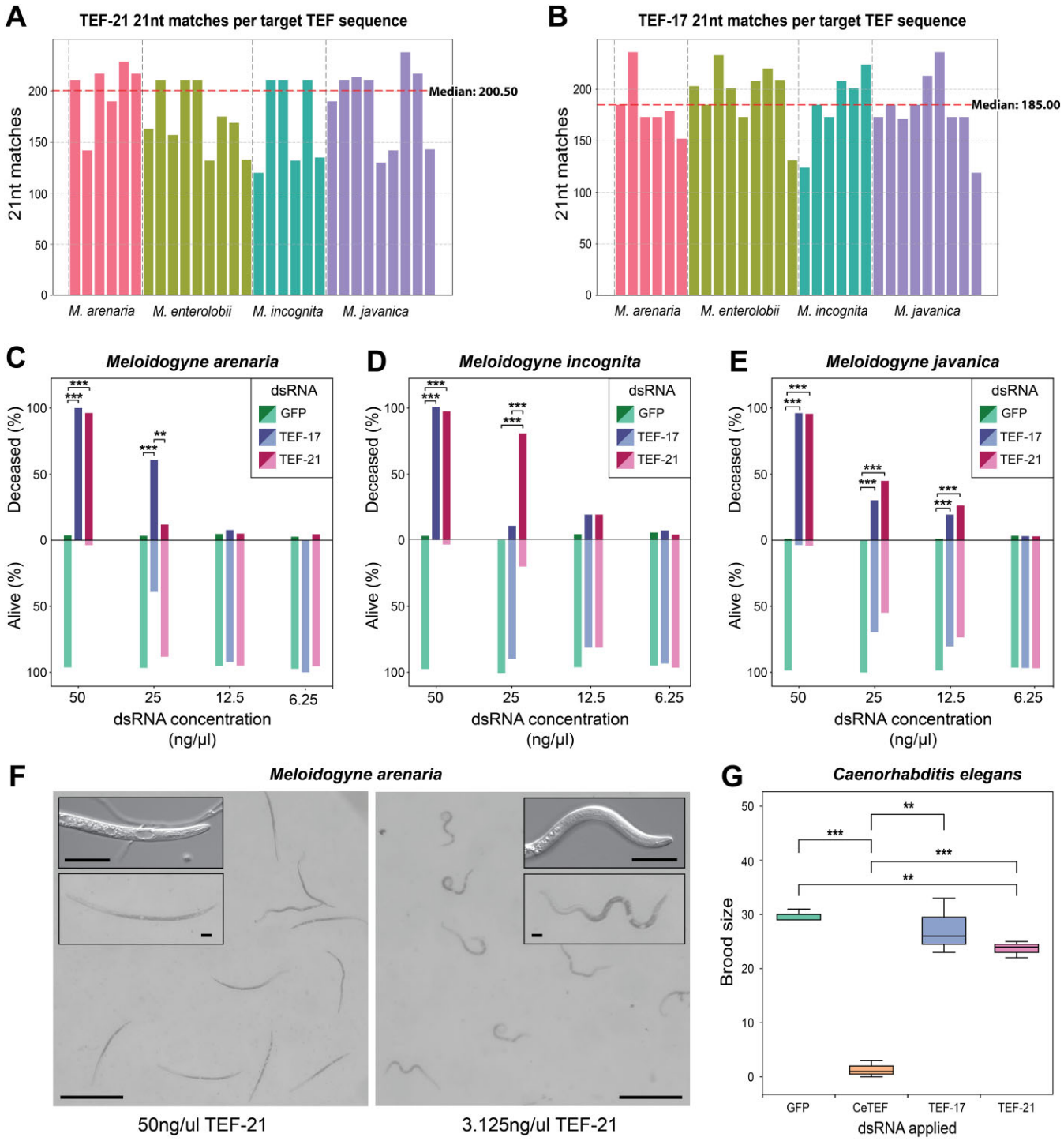


Figure 2. A single 300 nt dsRNA targeting *TEF1a* can control three RKN species without impacting the nematode *C. elegans*. (**A**) Number of 21 nt matches to each *TEF1a* transcript when the maximum contiguous match to the defined off-target *C. elegans* genome is 20 nt (median 21 nt matches of the dsRNA to all targets = 200.5). (**B**) Number of 21 nt matches to each *TEF1a* transcript when the maximum contiguous match to the defined off-target *C. elegans* genome is 16 nt (median 21 nt matches of the dsRNA to all targets = 185). (**C**) dsRNA concentration–kill curve for *M. arenaria*. (**D**) dsRNA concentration–kill curve for *M. incognita*. (**E**) dsRNA concentration–kill curve for *M. javanica*. (**F**) Example images of deceased (left) and alive *M. arenaria* (right) nematodes treated with 50 and 3.125 ng/µl of TEF-21 dsRNA, respectively. Deceased nematodes are largely straight and immobile, while live individuals exhibit a characteristic undulating movement. Scale bars on insets represent 50 µM and 0.5 mm for the full image. (**G**) *Caenorhabditis elegans* brood size in response to dsRNA application (50 ng/µl). Kill curve statistical tests were performed using Fisher's exact test followed by FDR multiple *P*-value correction. Brood size statistical tests were performed using pairwise *t*-tests followed by FDR multiple *P*-value correction (independent broods were derived from three egg-laying adults).

The number of 21 nt matches to the specific *M. enterolobii* transcript g3046.t1 in this situation was 203. By biasing towards this transcript by flagging it and setting the biasLvL flag to 10 (which adds 10 copies of this transcript to the design phase), the number of matching dsRNA-derived 21 nt *k*-mers rose to 236; however, the median *k*-mer hits reduced to 162 (Supplementary Fig. S6). That is, the increase in 21 nt matches to g3046.t1 comes at the cost of reduced matches to all transcripts.

dsRNAmax large dataset performance

Benchmarking on larger input and off-target datasets showed that dsRNAmax can rapidly design chimeric dsRNAs using desktop hardware (Supplementary Table S6). A larger target input set including 339 CMV RNA segments was used, along with an off-target .kmer dataset generated from cucumber soil rhizosphere metagenome sequencing. All design runs took <20 s. As FASTQ files cannot directly be used, the conversion to .kmer format provides a highly performant alternative in which off-target *k*-mers are stored as unsinged 64-bit integers. In this instance, singleton *k*-mers were removed to reduce the likelihood of sequencing errors impacting the off-target *k*-mer removal process. Similar to index generation for read alignment software, this process can be slow (over 10 min for a 7.7 GB read dataset), but only needs to be carried out once for each off-target dataset and is largely independent of on-target design concerns. The only requirement for the off-target contiguous match length is that it be less than or equal to the on-target *k*-mer size.

Discussion

Through validation experiments on major crop pests, RKNs, we have shown that the dsRNAmax software successfully achieves its primary objective: designing effective chimeric dsRNA constructs targeting multiple pest species while avoiding significant impacts on non-target organisms. By leveraging the conserved nature of the *TEF1a* gene across the four *Meloidogyne* species, despite multiple copies being present in each genome, dsRNAmax generated a single chimeric dsRNA construct that elicited high mortality in three tested species, confirming its control efficacy.

The use of the off-target *C. elegans* genome demonstrated the software's capacity to minimize unintended impacts. Constructs designed with stricter homology thresholds (e.g. TEF-17) demonstrated no significant difference in brood size to the non-target GFP control. In contrast, there was a significant but small magnitude difference between TEF-21 and GFP. If this were a real-world example and the off-target species were beneficial predatory nematodes present at application, the more conservative 17 nt off-target selection would be chosen to ensure no negative impacts on these beneficial species. Users may prefer to use transcriptome data for off-target species, as it may improve dsRNA *k*-mer matches to target sequences and, in some cases, allow for a lower off-target threshold, in contrast to genome data, which includes non-expressed or less relevant sequences.

This work highlights the flexibility of dsRNAmax in balancing specificity and efficacy through adjustable parameters, such as homology thresholds and target prioritization. The observed trade-off between maximizing coverage across all target transcripts and biasing the design towards specific sequences underscores the software's utility for tailoring dsRNA constructs to specific pest management needs.

Existing dsRNA design packages tend to focus on optimizing a selected region of a single RNA for targeting, based on factors such as AGO loading preference, siRNA strand selection, and accessibility of target sites (Supplementary Table S7). Where multi-gene or species targeting exists, e.g. dsRNAEngineer [14], a primary target gene is the template for dsRNA fragments that can be concatenated, with a selection process either maximizing matches to other related target species or minimizing matches to off-target species. In contrast, the chimeric design approach of dsRNAmax attempts to evenly target all selected genes (within and between species) and eliminates rather than minimizes any off-target matches of a set size. In addition, outside of local computational resources, no limits exist for the size and complexity of on- or off-target datasets.

dsRNAmax is designed for use against a wide range of targets, including plant viruses, fungi, oomycetes, nematodes, and insects, each with potentially differing siRNA biogenesis and AGO loading strategies. It is also intended for application in variable field conditions where other biotic and abiotic stresses may be present. Relatively little is known about dsRNA processing outside of model organisms in controlled laboratory conditions. Data concerning exogenous dsRNA uptake and processing are even more limited. To address these challenges, the primary design rule is to maximize the number of siRNAs that perfectly match each input sequence, rather than attempt to fit the preferences of a target species. Given that in some instances these preferences may be known [15–17], the software output includes information such as GC content and 5′ ribonucleotide occurrence for the matching *k*-mers for each target input, which may help guide design parameters.

Multi-target design and sequence conservation challenges

The development of dsRNAmax software for pest and pathogen control presents several critical considerations. Multi-target chimeric dsRNA design can be effective in scenarios involving conserved genetic targets, but limitations exist when functionally similar targets become too divergent at the transcript sequence level. This necessitates a nuanced approach to dsRNA design that recognizes the complexity of genetic variation. The *Meloidogyne* spp. *TEF1a* gene represents a case where, at a minimum, there is 86.5% sequence identity between two target sequences. As *TEF1a* is functionally a good target for exogenous dsRNA, dsRNAmax is capable of designing a single effective construct. A relevant example where greater divergence often exists is plant virus genomes of the same species. Accordingly, generating effective universal chimeric dsRNAs as target sequences diverge becomes increasingly difficult. In such situations, multiple or concatenated dsRNAs may be the best solution.

Risk assessment and practical implications

While off-target effects represent a complex consideration in dsRNA design, our experimental nematode data indicate that multiple species can be targeted with a single dsRNA and negative impacts in other dsRNA-susceptible species can be avoided. Bioinformatics-based design, however, should be understood as a preliminary step rather than a definitive solution. Rigorous empirical testing remains essential, particularly

when targeting novel or less-studied species, both for on-target efficacy and for off-target avoidance. The cost and time requirements of synthesizing chimeric dsRNAs should also be considered when screening potential target genes. In some instances, pre-screening targets with perfectly matching dsRNAs may be desirable.

Conclusion

Effective dsRNA design requires a multi-layered approach that balances computational prediction with empirical validation. The most successful strategies will integrate bioinformatics with experimental testing, particularly when addressing genetically varied targets in less-studied species and complex ecological systems. dsRNAmax offers a software solution that sits in the design phase of this workflow, generating safe and specific universal chimeric dsRNAs for use in crop pest and pathogen control applications.

Acknowledgements

We would like to thank Jennifer Cobon, Wayne O'Neill, Tim Shuey and Dylan Corner from Queensland Department of Primary Industries for their technical advice and for providing the RKNs.

Author contributions: Stephen Fletcher (Conceptualization, Software, Formal analysis, Methodology, Validation, Writing—original draft), Jai Lawrence (Investigation, Validation), Anne Sawyer (Investigation, Methodology, Writing—review & editing), Narelle Manzie (Project administration, Writing—review & editing), Donald Gardiner (Project administration, Writing—review & editing), Neena Mitter (Funding acquisition, Project administration, Writing—review & editing), and Christopher Brosnan (Conceptualization, Methodology, Validation, Writing—original draft)

Supplementary data

Supplementary data is available at NAR Genomics & Bioinformatics online.

Conflict of interest

None declared.

Funding

This work was supported by Australian Research Council [IH190100022]. A.S. was supported by an Advance Queensland Industry Research Fellowship. Funding to pay the Open Access publication charges for this article was provided by Australian Research Council.

Data availability

All data are incorporated into the article and its online supplementary material.

References

1. FAO. *FAO's Plant Production and Protection Division*. 2022. http://www.fao.org/documents/card/en/c/cc2447en (28 November 2024, date last accessed).
2. Smith J (ed.), *Environmental and Health Impacts of Pesticides and Fertilizers and Ways of Minimizing Them: Synthesis Report*. United Nations Environment Programme, Nairobi. 2023. https://www.fao.org/fileadmin/user_upload/soils/publications/pesticides.pdf (28 November 2024, date last accessed).
3. Chen C, Imran M, Feng X *et al.* Spray-induced gene silencing for crop protection: recent advances and emerging trends. *Science* 2025;**16**:544–50. https://doi.org/10.3389/fpls.2025.1527944
4. Willow J, Smagghe G. RNAi applications towards environmentally sustainable food security. *Curr Opin Environ Sci Health* 2025;**45**:100612. https://doi.org/10.1016/j.coesh.2025.100612
5. Svoboda P. Key mechanistic principles and considerations concerning RNA interference. *Front Plant Sci* 2020;**11**:1237. https://doi.org/10.3389/fpls.2020.01237
6. Niu D, Hamby R, Sanchez JN *et al.* RNAs—a new frontier in crop protection. *Curr Opin Biotechnol* 2021;**70**:204–12. https://doi.org/10.1016/j.copbio.2021.06.005
7. Bachman P, Fridley J, Mueller G *et al.* Sequence–activity relationships for the Snf7 insecticidal dsRNA in Chrysomelidae. *Front Plant Sci* 2020;**11**:1303. https://doi.org/10.3389/fpls.2020.01303
8. Gabriel L, Brůna T, Hoff KJ *et al.* BRAKER3: fully automated genome annotation using RNA-seq and protein evidence with GeneMark-ETP, AUGUSTUS and TSEBRA. *Genome Res* 2024;**34**:769–77. https://doi.org/10.1101/2023.06.10.544449
9. Suzek BE, Wang Y, Huang H *et al.* UniRef clusters: a comprehensive and scalable alternative for improving sequence similarity searches. *Bioinformatics* 2015;**31**:926–32. https://doi.org/10.1093/bioinformatics/btu739
10. Mistry J, Chuguransky S, Williams L *et al.* Pfam: the protein families database in 2021. *Nucleic Acids Res* 2021;**49**:D412–9. https://doi.org/10.1093/nar/gkaa913
11. Brenner S. The genetics of *Caenorhabditis elegans*. *Genetics* 1974;**77**:71–94. https://doi.org/10.1093/genetics/77.1.71
12. Perry RN, Moens M, Starr JL (eds.), *Root-Knot Nematodes*, 1st edn. UK: CABI, 2009.
13. Degnan RM, McTaggart AR, Shuey LS *et al.* Exogenous double-stranded RNA inhibits the infection physiology of rust fungi to reduce symptoms in planta. *Mol Plant Pathol* 2023;**24**:191–207. https://doi.org/10.1111/mpp.13286
14. Chen Y, Shi Y, Wang Z *et al.* dsRNAEngineer: a web-based tool of comprehensive dsRNA design for pest control. *Trends Biotechnol* 2025;**43**:969–83. https://doi.org/10.1016/j.tibtech.2025.01.002
15. Takeda A, Iwasaki S, Watanabe T *et al.* The mechanism selecting the guide strand from small RNA duplexes is different among Argonaute proteins. *Plant Cell Physiol* 2008;**49**:493–500. https://doi.org/10.1093/pcp/pcn043
16. Czech B, Zhou R, Erlich Y *et al.* Hierarchical rules for Argonaute loading in *Drosophila*. *Mol Cell* 2009;**36**:445–56. https://doi.org/10.1016/j.molcel.2009.09.028
17. Drinnenberg IA, Weinberg DE, Xie KT *et al.* RNAi in budding yeast. *Science* 2009;**326**:544–50. https://doi.org/10.1126/science.1176945

Received: January 30, 2025. **Revised:** April 16, 2025. **Editorial Decision:** May 4, 2025. **Accepted:** May 7, 2025