Bioinformatics, 32(8), 2016, 1232–1234
doi: 10.1093/bioinformatics/btv747
Advance Access Publication Date: 24 December 2015
Applications Note

OXFORD

Sequence analysis

OfftargetFinder: a web tool for species-specific RNAi design

R.T. Good[1,2,†], T. Varghese[3,†], J.F. Golz[2], D.A. Russell[4], A. Papanicolaou[5], O. Edwards[5] and C. Robin[1,2,*]

[1]The Bio21 Institute [2]School of Biosciences, The University of Melbourne, Melbourne 3010, Australia, [3]CSIRO National Facilities and Collections, Canberra, ACT 2601, Australia, [4]Faculty of Veterinary and Agricultural Sciences, The University of Melbourne, Melbourne 3010, Australia and [5]CSIRO Land and Water Flagship, Canberra, ACT 2601, Australia

[†]The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.
*To whom correspondence should be addressed.
Associate Editor: Ivo Hofacker

Received on 19 May 2015; revised on 22 October 2015; accepted on 16 December 2015

Abstract

Motivation: RNA interference (RNAi) technology is being developed as a weapon for pest insect control. To maximize the specificity that such an approach affords we have developed a bioinformatic web tool that searches the ever-growing arthropod transcriptome databases so that pest-specific RNAi sequences can be identified. This will help technology developers finesse the design of RNAi sequences and suggests which non-target species should be assessed in the risk assessment process.
Availability and implementation: http://rnai.specifly.org.
Contact: crobin@unimelb.edu.au

1 Introduction

RNA interference (RNAi) is a powerful reverse-genetics tool routinely used to reduce or eliminate gene activity within an organism. RNAi is most commonly applied in 'functional genomics' screens of laboratory model organisms and cell lines (Crane and Gelvin, 2007; Dietzl *et al.*, 2007; Kamath, 2003). RNAi technology is also being developed for so called 'environmental RNAi' where it can be applied to free-living organisms such as pest insects (Smagghe and Swevers, 2014). The key attribute of RNAi for all these applications is its specificity, which is determined by sequence identity between the RNAi molecule and the target gene. However the nature of specificity-requirements differs between applications. In reverse genetic screens of model organisms a premium is placed on targeting the specified target gene and no other 'offtarget' genes. There are many web-based tools to help with the design of such RNAi molecules (Arziman *et al.*, 2005; Mathews, 2010; http://www.flyrnai.org/snapdragon). However for environmental applications species-specificity is just as important as gene-specificity.

The application of RNAi to pest control is an attractive alternative to traditional control chemicals because it has the potential to only target pest species. Among the pests for which this technology is being developed are beetles, moths, locusts and various phloem feeders including aphids (Christiaens and Smagghe, 2014; Li *et al.*, 2013; McHale *et al.*, 2013; Mutti *et al.*, 2008; Tian *et al.*, 2009). A challenge of such an approach is in the delivery of RNAi-inducing molecules to pest organisms. This might be achieved by expressing dsRNA in transgenic crop plants that are hosts for the pests being targeted (Whyard, 2015), or by using sprays containing either naked dsRNA molecules or dsRNA molecules packaged within micro-organisms (Palli, 2014). The former approach has been taken by Monsanto, which has developed transgenic crop plants that produce dsRNA against the western corn rootworm (*Diabrotica virgifera virgifera*). Bolognesi *et al.* (2012) showed that ingestion of a 240 nucleotide RNAi sequence directed to the rootworm *snf7* gene is lethal. Experiments surveying ten insect families found that only closely related beetles, those sharing at least three 21-mers with the dsRNA sequence, were affected by consuming the *snf7* dsRNA molecule (Bachman *et al.*, 2013). This not only demonstrates the potential specificity of the technology but also illustrates the fact that only a small number of 21mers (siRNA's) may elicit a response.

Given the likelihood that RNAi-based technologies may be entering our agricultural landscapes in the near future it is critically important to establish the species-specificity of such RNAi-inducing molecules. We have developed a bioinformatic tool that searches transcriptomic databases that will (i) assist RNAi-technology developers in maximizing the pest specificity of RNAi molecules and (ii) provide information for regulatory authorities and the public on the relative environmental risks that these molecules have on non-target organisms. This web-based tool is called 'OffTargetFinder' and the various outputs it provides, enables users to refine the regions within gene of interest, so as to eliminate potential off-target effects in other arthropod species. It can also be used to identify which species should be tested experimentally in the ecological risk assessment process (Romeis *et al.*, 2013).

2 Methods

The user interface is designed using ExtJS3 (http://www.sencha.com/products/extjs/). The bioinformatics search is driven by custom PERL script that uses the Bowtie software (Langmead *et al.*, 2008) to rapidly search transcriptomes for 21mers with or without mismatches. It is then visualized using the CanvasXpress genome browser package (http://canvasxpress.org/). At the time of submission of this manuscript OffTargetFinder screens 101 Arthropod transcriptomes, including representatives from the 32 currently recognized orders within the subphylum Hexapoda, and from three other Arthropod subphyla; Crustacea, Myriapoda and Chelicerata. The database of transcriptomes includes sequences from the 1000 Insect Transcriptome Evolution (1KITE) dataset (http://www.1kite.org; Misof *et al.*, 2014). Various other transcriptomes are also included such as those from humans, the model plant *Arabidopsis thaliana*, and our annotation of the *Myzus persicae* genome using Exonerate (Slater and Birney, 2005) queries with the pea aphid transcriptome.

3 Usage

3.1 Enter query sequence

OffTargetFinder is accessed freely over the web at http://rnai.specifly.org.

The user pastes a DNA sequence in the FASTA format (with no white spaces in the definition line) that they believe may be a good RNAi target (e.g. a gene that is essential for pest insect viability). Users then specify how many mismatches (0, 1, 2, 3) will be tolerated in each window of 21 nucleotides. Once the 'Find Off Targets' button is chosen the query appears in the History table and the status of the job is reported. For multiple queries, a separate window can be opened for each query sequence that is submitted. The user can check for changes in query status by reloading their browser. They can view results when query status is completed. The green colored text indicates query has returned hits and red text indicates no hits were found. The program takes about 15 minutes to run, and when completed results can be viewed by clicking on the motif in the Show Result column.

3.2 Results page

A view of three panels is seen in the results tab. The top panel shows the query sequence and the panel on the left (*Organism classification panel*) shows a cladogram of taxonomic ranks in which hits are observed. The cladogram splits 'Arthropods' (most of the species considered) from 'Others' (which includes unclassified) and shows Order, Family, Genus and Species names that are based on 1KITE species classifications. The colour of the box next to the taxa names reflects the number of 'hits' observed in that taxa (purple = 1 hit, dark blue= 2–5, light blue = 6–10, green = 11–20, yellow = 21–50, orange = 51–100, red > 101). The plus and minus signs on the nodes of the cladogram can be clicked to collapse and expand those nodes. If the coloured boxes are clicked the analysis in the Query Result panel is limited to the species of that clade.

The Query Result panel houses a variety of different views that can be toggled between by clicking on view button.

3.3 View: browser

Among these views is 'Browser' which lists each species that has at least one 'hit' with the input sequence and indicates the location of the hits in a cartoon, much like the first figure of an NCBI Blast search (Fig. 1). A feature of this view is that the user can zoom in to nucleotide resolution to see the exact 21mer sequences responsible for the off-target hit/s.

3.4 View: query hit regions

The view provided by the 'Query Hit Regions' paints the query sequence in colours according to the number of 21mers in the database (or the portion of the database delimited by the Organism Classification panel) that match that particular 21 bp window. With this tool the user will identify stretches of nucleotides that are highly specific to the targeted species. This output will be useful in ensuring that a particular dsRNA molecule has minimal off-target hits.

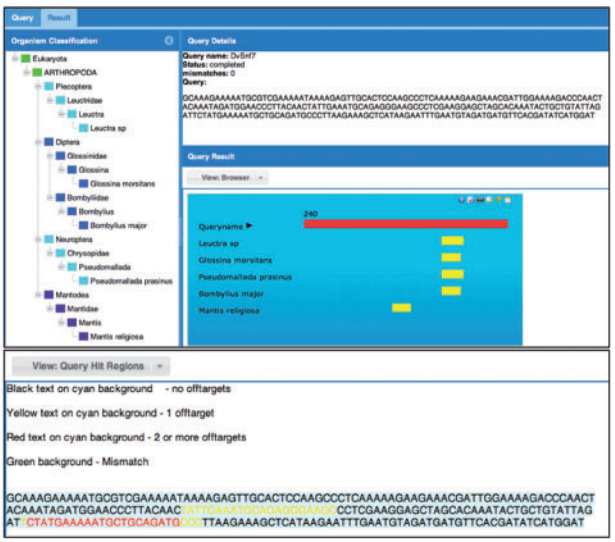

Fig. 1. Offtarget hits using the snf7 gene. The 240nt sequence of the *Snf7* gene from the beetle *Diabrotica virgifera* that was experimentally tested for cross-species hits by Bachman *et al.* (2013) was used as the query sequence in OffTargetFinder. Five hits were observed (see cladogram on the left panel), and they are distributed across the orders Plecoptera (a stonefly), Diptera (specifically a tsetse fly and a bee-fly), Neuroptera (a lacewing) and Mantodea (a praying mantis). This could motivate risk assessments, like the Bachman *et al.* (2013) experiments, on these particular 'off target' species. The large panel to the right of the cladogram shows which parts of the query contains at least one 21-mer hit across the five species. Note that in each of the cases here the hits are limited to one contiguous hit region. Note also that one region in the query sequence hits four distantly related non-target species. The specific nucleotides that are hitting the off target species are observed in the bottom panel. Together these views demark sequences that should be avoided in RNAi designs to ensure taxonomically limited RNAi pesticides

3.5 View: raw data

It is also possible to obtain a table of the number of 21mer 'hits' observed in each species ('Raw data'). In this, as in all other views, clicking on the 'Eukaryote' box in the left panel will display all the results, whereas clicking on any subordinate box restricts the data in the view box to the subset of taxa selected.

3.6 View: bar graph

The number of hits in each species can also be observed as bar charts with the *x*-axis having the species and the *y*-axis the hit count.

4 Conclusions

The web tool OfftargetFinder can be used to identify stretches of nucleotides that do not contain 21mer matches with a library of over 100 arthropod transcriptomes. It can therefore be used to design RNAi pesticides likely to have a narrow spectrum of affected species.

Acknowledgements

We are grateful to the 1KITE community, especially Karen Meusemann, for providing access to their datasets.

Funding

This work has been supported by the Australia India Strategic Research Fund, Grand Challenge Scheme, Grant Number GCF010009.

Conflict of Interest: none declared.

References

Arziman,Z. *et al*. (2005) E-RNAi: a web application to design optimized RNAi constructs. *Nucleic Acids Res*., **33**, W582–W588.

Bachman,P.M. *et al*. (2013) Characterization of the spectrum of insecticidal activity of a double-stranded RNA with targeted activity against Western Corn Rootworm (*Diabrotica virgifera virgifera* LeConte). *Transgenic Res*., **22**, 1207–1222.

Bolognesi,R. *et al*. (2012) Characterizing the mechanism of action of double-stranded RNA activity against western corn rootworm (*Diabrotica virgifera virgifera* LeConte). *PLoS ONE*, **7**, e47534.

Christiaens,O. and Smagghe,G. (2014) The challenge of RNAi-mediated control of hemipterans. *Curr. Opin. Insect Sci*., **6**, 15–21.

Crane,Y.M. and Gelvin,S.B. (2007) RNAi-mediated gene silencing reveals involvement of Arabidopsis chromatin-related genes in Agrobacterium-mediated root transformation. *Proc. Natl. Acad. Sci*., **104**, 15156–15161.

Dietzl,G. *et al*. (2007) A genome-wide transgenic RNAi library for conditional gene inactivation in Drosophila. *Nature*, **448**, 151–156.

Kamath,R. (2003) Genome-wide RNAi screening in *Caenorhabditis elegans*. *Methods*, **30**, 313–321.

Langmead. *et al*. (2008) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*., **10**, R25–R25.

Li,J. *et al*. (2013) Advances in the use of the RNA interference technique in Hemiptera. *Insect Sci*, **20**, 31–39.

Mathews,D.H. (2010) Using OligoWalk to identify efficient siRNA sequences. *Methods Mol. Biol*., **629**, 109–121.

McHale,M. *et al*. (2013) A 22-nt artificial microRNA mediates widespread RNA silencing in *Arabidopsis*. *Plant J*., **76**, 519–529.

Misof,B. *et al*. (2014) Phylogenomics resolves the timing and pattern of insect evolution. *Science*, **346**, 763–767.

Mutti,N.S. *et al*. (2008) A protein from the salivary glands of the pea aphid, *Acyrthosiphon pisum*, is essential in feeding on a host plant. *Proc. Natl. Acad. Sci. U.S.A*., **105**, 9965–9969.

Palli,S.R. (2014) RNA interference in Colorado potato beetle: steps toward development of dsRNA as a commercial insecticide. *Curr. Opin. Insect Sci*., **6**, 1–8. pg

Romeis,J. *et al*. (2013) Deriving criteria to select arthropod species for laboratory tests to assess the ecological risks from cultivating arthropod-resistant genetically engineered crops. *Chemosphere*, **90**, 901–909.

Slater,G. *et al*. (2005) Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics*, **6**, 31

Smagghe,G. and Swevers,L. (2014) Pests and resistance- RNAi research in insects. *Curr. Opin. Insect Sci*., **6**, iv–iv.

Tian,H. *et al*. (2009) Developmental control of a lepidopteran pest *Spodoptera exigua* by ingestion of bacteria expressing dsRNA of a non-midgut gene. *PLoS One*, **4**, e6225.

Whyard,S. (2015) Plant science. Insecticidal RNA, the long and short of it. *Science*, **347**, 950–951.