

NRC Publications Archive Archives des publications du CNRC

WheatCRISPR: a web-based guide RNA design tool for CRISPR/Cas9-mediated genome editing in wheat

Cram, Dustin; Kulkarni, Manoj; Buchwaldt, Miles; Rajagopalan, Nandhakishore; Bhowmik, Pankaj; Rozwadowski, Kevin; Parkin, Isobel A. P.; Sharpe, Andrew G.; Kagale, Sateesh

This publication could be one of several versions: author's original, accepted manuscript or the publisher's version. / La version de cette publication peut être l'une des suivantes : la version prépublication de l'auteur, la version acceptée du manuscrit ou la version de l'éditeur.

For the publisher's version, please access the DOI link below. / Pour consulter la version de l'éditeur, utilisez le lien DOI ci-dessous.

Publisher's version / Version de l'éditeur:

<https://doi.org/10.1186/s12870-019-2097-z>

BMC Plant Biology, 19, 1, 2019-11-06

NRC Publications Archive Record / Notice des Archives des publications du CNRC :

<https://nrc-publications.canada.ca/eng/view/object/?id=b2ba8291-8753-4b1c-aaa7-b266a09d174e>

<https://publications-cnrc.canada.ca/fra/voir/objet/?id=b2ba8291-8753-4b1c-aaa7-b266a09d174a>

Access and use of this website and the material on it are subject to the Terms and Conditions set forth at

<https://nrc-publications.canada.ca/eng/copyright>

READ THESE TERMS AND CONDITIONS CAREFULLY BEFORE USING THIS WEBSITE.

L'accès à ce site Web et l'utilisation de son contenu sont assujettis aux conditions présentées dans le site

<https://publications-cnrc.canada.ca/fra/droits>

LISEZ CES CONDITIONS ATTENTIVEMENT AVANT D'UTILISER CE SITE WEB.

Questions? Contact the NRC Publications Archive team at

PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca. If you wish to email the authors directly, please see the first page of the publication for their contact information.

Vous avez des questions? Nous pouvons vous aider. Pour communiquer directement avec un auteur, consultez la première page de la revue dans laquelle son article a été publié afin de trouver ses coordonnées. Si vous n'arrivez pas à les repérer, communiquez avec nous à PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca.

DATABASE

Open Access



WheatCRISPR: a web-based guide RNA design tool for CRISPR/Cas9-mediated genome editing in wheat

Dustin Cram¹, Manoj Kulkarni¹, Miles Buchwaldt², Nandhakishore Rajagopalan¹, Pankaj Bhowmik¹, Kevin Rozwadowski², Isobel A. P. Parkin², Andrew G. Sharpe^{1,3} and Sateesh Kagale^{1*} 

Abstract

Background: CRISPR/Cas9 gene editing has become a revolutionary technique for crop improvement as it can facilitate fast and efficient genetic changes without the retention of transgene components in the final plant line. Lack of robust bioinformatics tools to facilitate the design of highly specific functional guide RNAs (gRNAs) and prediction of off-target sites in wheat is currently an obstacle to effective application of CRISPR technology to wheat improvement.

Description: We have developed a web-based bioinformatics tool to design specific gRNAs for genome editing and transcriptional regulation of gene expression in wheat. A collaborative study between the Broad Institute and Microsoft Research used large-scale empirical evidence to devise algorithms (Doech et al., 2016, *Nature Biotechnology* 34, 184–191) for predicting the on-target activity and off-target potential of CRISPR/*SpCas9* (*Streptococcus pyogenes* Cas9). We applied these prediction models to determine on-target specificity and potential off-target activity for individual gRNAs targeting specific loci in the wheat genome. The genome-wide gRNA mappings and the corresponding Doench scores predictive of the on-target and off-target activities were used to create a gRNA database which was used as a data source for the web application termed WheatCRISPR.

Conclusion: The WheatCRISPR tool allows researchers to browse all possible gRNAs targeting a gene or sequence of interest and select effective gRNAs based on their predicted high on-target and low off-target activity scores, as well as other characteristics such as position within the targeted gene. It is publicly available at <https://crispr.bioinfo.nrc.ca/WheatCrispr/>.

Keywords: CRISPR, Cas9, gRNA design tool, Genome editing, Transcriptional regulation, Wheat

Background

Genome editing technology based on a bacterial adaptive immune system, termed CRISPR (Clustered, Regularly Interspersed, Palindromic Repeats) / Cas9 (CRISPR-associated endonuclease 9 [1–4]); has sparked a new revolution in biological and agricultural research [5, 6]. CRISPR/Cas9 technology originating from *Streptococcus pyogenes* relies on two important components, a Cas9 endonuclease and a single guide RNA (sgRNA) formed by fusing two small RNA molecules, namely CRISPR

RNA (crRNA) and an auxiliary trans-activating crRNA (tracrRNA) that together guide Cas9 nuclease to a specific DNA site [7, 8]. Each crRNA unit contains a 20-nt guide sequence complementary to a target site, designated as guide RNA (gRNA). Another critical feature of the Cas9 system is the Protospacer-Adjacent Motif (PAM) flanking the 3'-end of the DNA target site that dictates the target search mechanism of Cas9 [9]. The PAM comprises a triplet of base pairs with a canonical sequence 5'-NGG-3' where "N" is any nucleotide [9]. Other non-canonical PAM triplets have also been described, including NAG, NCG and NGA that support less efficient CRISPR/Cas9 functions [10, 11], and thus may contribute to off-target activity.

* Correspondence: Sateesh.Kagale@nrc-cnrc.gc.ca

¹National Research Council Canada, 110 Gymnasium Place, Saskatoon, SK S7N 0W9, Canada

Full list of author information is available at the end of the article



Although CRISPR/Cas9 applications promise to accelerate the pace and course of crop improvement [5, 6], a number of hurdles exist that limit full exploitation of this innovative technology, especially in crops with large polyploid genomes. Wheat is an economically important cereal crop providing 20% of the calorie and protein intake for the global population. It harbours a complex allohexaploid genome of 16 Gb with approximately 85% repetitive elements and estimated 107,921 high confidence and a further 161,537 low confidence annotated genes [12]. Due to the presence of up to six homoeoalleles per gene and large gene families, off-target gRNA binding and cleavage is one of the most critical issues that affect implementation of CRISPR/Cas9 technology in wheat. The gRNA is an important component of the CRISPR/Cas9 system as it determines the efficacy and specificity of Cas9 nuclease. An effective gRNA should have high on-target activity and low off-target potential. Thus, rational design and optimization of functional gRNA sequences is essential to achieving maximal effectiveness and highest targeting specificity for intended genomic location(s).

Multiple bioinformatics tools have been developed to facilitate the design of gRNAs and prediction of off-target sites [13–24]; however, only two of these programs, including E-CRISP [17] and CRISPRdirect [19] support design of gRNAs for wheat. CRISPRdirect predicts specific gRNAs based on in silico prediction of specificity but the lack of implementation of evidence-based metrics to predict off-target sites is a notable caveat. E-CRISP identifies off-target sites by aligning gRNAs to the genome with Bowtie2. However, Bowtie2 does not guarantee that all possible hits will be found, especially when the number of mismatches is high [11]. This results in an underestimation of potential off-target sites. A collaborative effort between scientists at the Broad Institute and machine learning experts at Microsoft Research used large-scale empirical evidence based on cleavage potential of thousands of gRNAs targeting a panel of 15 genes to uncover position-specific sequence features that are predictive of gRNA efficacy and specificity, including the position and frequency of single and di-nucleotides, the GC content of the gRNA, the location of the gRNA within the protein coding region and melting temperatures of the first 5, middle 8 and last 5 base pairs of the gRNA [11, 25]. The findings from these large-scale empirical data were utilized to devise new rules for gRNA on-target activity [rule set (rs) 2] and cutting frequency determination (CFD) scores to predict gRNA off-target effects [11], that can be broadly applied. In this study, we applied these prediction models to determine on-target specificity and potential off-target activity of individual gRNAs targeting any locus in the wheat genome, and designed a web-based bioinformatics

Table 1 Survey of PAM sites in the IWGSC v1.0 Chinese Spring wheat reference genome sequence

Region	Canonical (NGG)	Non-canonical (NAG, NCG and NGA)
Coding	33,543,850	93,876,010
Promoter	46,236,876	133,448,536
UTR/Intron	32,345,815	101,719,313
Intergenic	1,520,739,628	4,374,331,934

portal (WheatCRISPR) for design of highly specific gRNAs for CRISPR/Cas9-mediated genome editing and CRISPR-based transcriptional regulation of gene expression in Chinese Spring wheat.

gRNA database construction and content

Based on the current annotation [12], the bread wheat genome has approximately 35 million canonical PAM sites in coding regions and over 6 billion potential off-target sites across the entire genome including intergenic regions and non-canonical PAM sites (Table 1). Although applying the Doench algorithms is conceptually simple, the huge number of PAM sites in the large wheat genome makes the task of predicting off-target activity computationally challenging. Running the prediction models on all possible pairs of on-target (canonical coding and promoter) and off-target (all) sites is a daunting computational endeavour.

To reduce the number of potential off-target sites that must be considered to apply the Doench CFD algorithm, we limited our search to only those sites that have at most k mismatches to on-target sites. While searching for sites with k mismatches is much faster than applying the Doench algorithm to every possible site, it is still computationally intensive. To make this solution more tractable, we varied the maximum value of k in different regions of the genome such that the likelihood of detrimental off-target effects is minimized. For canonical NGG PAM sites in exons and promoter regions (2 kb upstream of the start codon), we searched up to $k = 6$. For other genic regions [i.e. introns and untranslated region (UTR)], we searched up to $k = 4$, and for intergenic regions we searched up to $k = 3$ (Table 2). For non-canonical PAM sites (NAG, NCG, NGA) we searched at

Table 2 Tiered k (mismatches) levels applied for the search of off-targets in the wheat genome

PAM	Off-target region	k (mismatches)
Canonical (NGG)	Exons and promoters	6
	Introns and UTRs	4
	Intergenic	3
Non-canonical (NAG, NCG, NGA)	Exons and promoters	4
	Introns and UTRs	3
	Intergenic	2

$k=4$, $k=3$, and $k=2$ in these regions, respectively (Table 2). In addition to these absolute k limits in each region, if at least 20 off-target matches were found up to a given k mismatches, we did not proceed to search for $k+1$ mismatches. Testing a subset of 1 M on-target:off-target gRNA pairings suggested that the wheat genome has very few active off-target sites with $k \geq 3$ (Fig. 1), making the tiered k approach a far more efficient search mechanism for the large wheat genome.

To implement this strategy, we extracted the PAM and gRNA sequences from all possible PAM sites in the

IWGSC v1.0 wheat (Chinese Spring) genome [12], separating the output into canonical and non-canonical sites. These sites were further categorized by their genomic location: coding, promoter, other genic (introns and UTR), and intergenic. The canonical coding and promoter sets were designated as on-target datasets and the rs2 algorithm was applied to each on-target gRNA sequence. Each of the on-target gRNA sequences was then searched against all eight datasets for off-by k mismatches, setting the maximum k as described in Table 2. The Doench CFD algorithm was applied to the

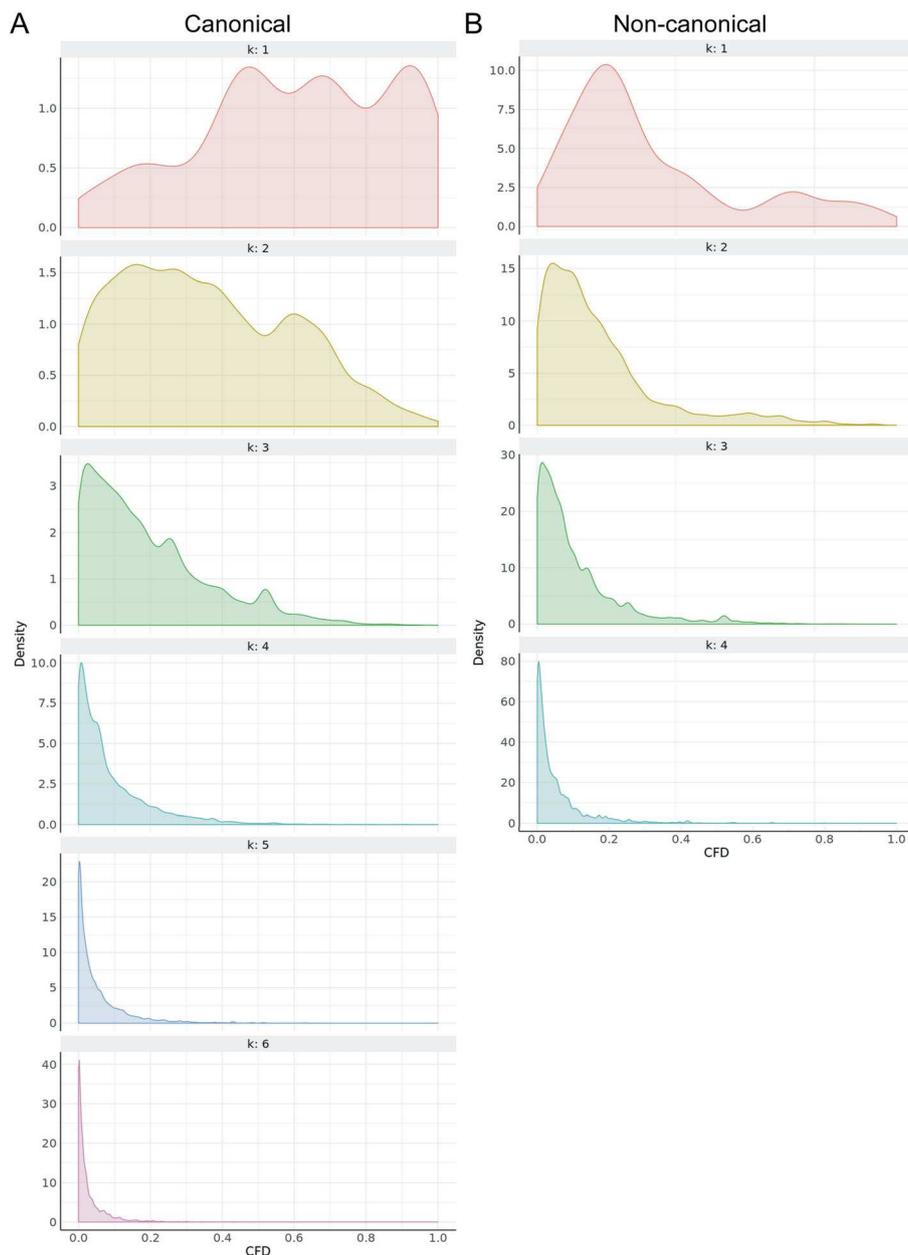


Fig. 1 Assessment of off-target effects of gRNAs with a tiered k (mismatches) approach. CFD densities are plotted by k (mismatches) using a subset of approximately one million canonical (a) and non-canonical (b) on-target:off-target gRNA pairings

resulting set of on-target::off-target mappings. The steps involved in creating the gRNA database are illustrated in Fig. 2. The gRNA mappings, the Doench scores, and the positions of the sequences in the genome were used as the data source for the WheatCRISPR web application.

Utility and discussion

WheatCRISPR (<https://crispr.bioinfo.nrc.ca/WheatCrispr/>) provides a convenient interface to browse the gRNA database, and allows researchers to view a set of predicted gRNAs targeting a gene or sequence of interest, and select them based on their predicted on-target and off-target

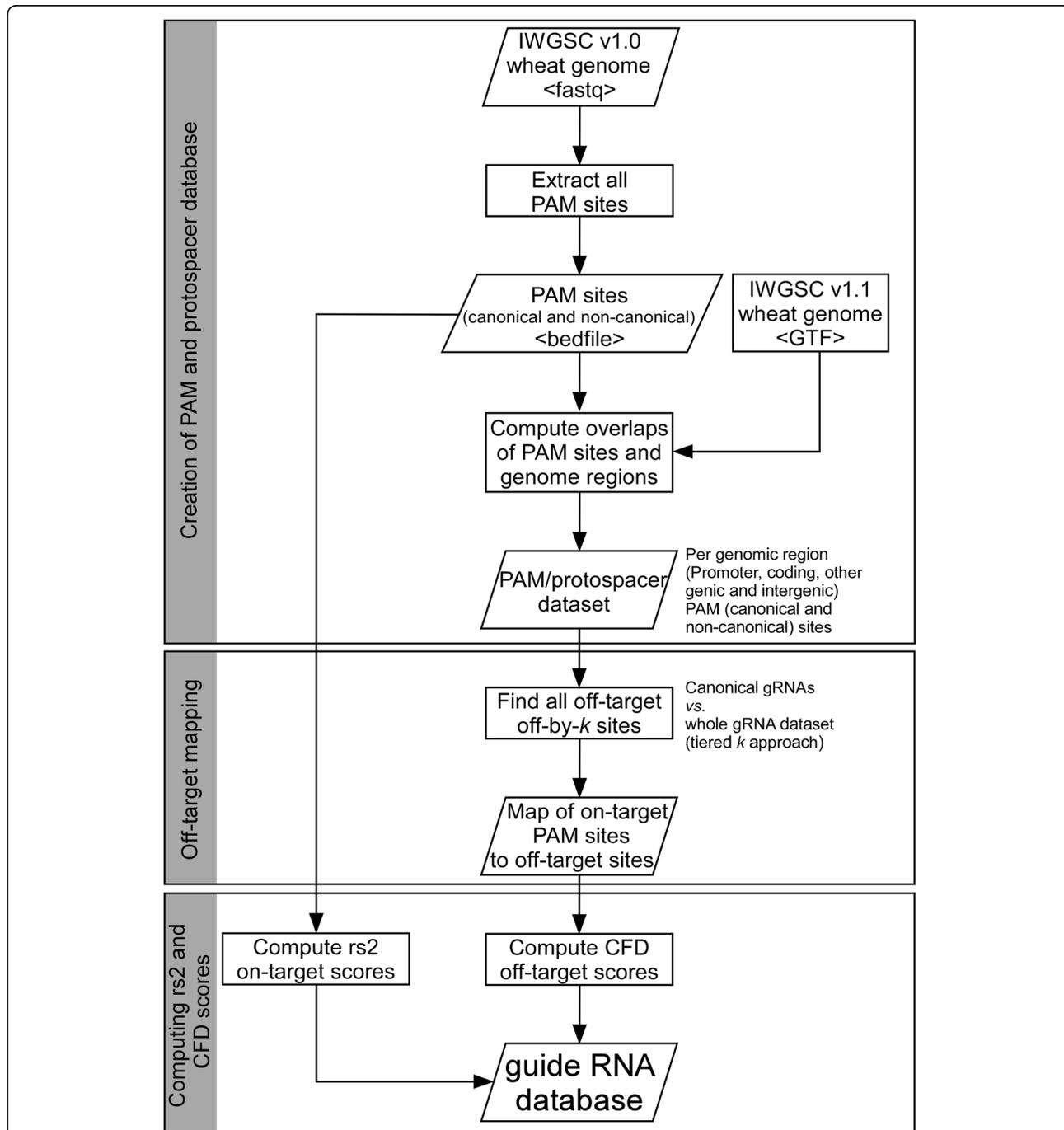


Fig. 2 Flowchart of the steps involved in creating the gRNA database. The workflow comprises three major steps, including (1) genome-wide scanning and extraction of PAM sites and adjoining gRNA sequences, (2) mapping off-targets and (3) computing on-target activity (rs2) and cutting frequency determination (CFD) scores based on the Doench prediction models (Doench et al., 2016)

activity, and the position of the gRNA within the targeted gene. The application presents summary statistics in graphs and tables that expedites the quick finding of the most effective candidate gRNAs for the gene of interest (Fig. 3 and Additional file 1: Table S1). By default, detailed information is displayed for the ten highest scoring gRNAs to facilitate rapid identification of the most likely efficacious candidate gRNA sequences. An interactive interface allows the user to browse all other gRNAs if desired.

A key summary statistic for evaluating the off-target activity of a gRNA is the maximum CFD score for the gRNA, i.e. the single worst off-target hit. The gRNA plot (exemplified in Fig. 3a) and table (Additional file 1: Table S1) for a given gene presents the rs2 score and the

maximum CFD score for each of the four genomic regions: coding, promoter, other genic, and intergenic. This facilitates selection of specific gRNAs by characterizing the potential severity of off-target effects based on the likelihood of unintended activity resulting in functional change to a coding region.

WheatCRISPR assists the user to find a trade-off between high on-target activity and low off-target activity by calculating an overall score for each gRNA that rewards high rs2 scores and penalizes high CFD scores (Additional file 1: Table S1). The overall score is a weighted average of the rs2 and maximum CFD scores. An optional variation of this score can be toggled on or off if the user wishes to target all homoeologous copies

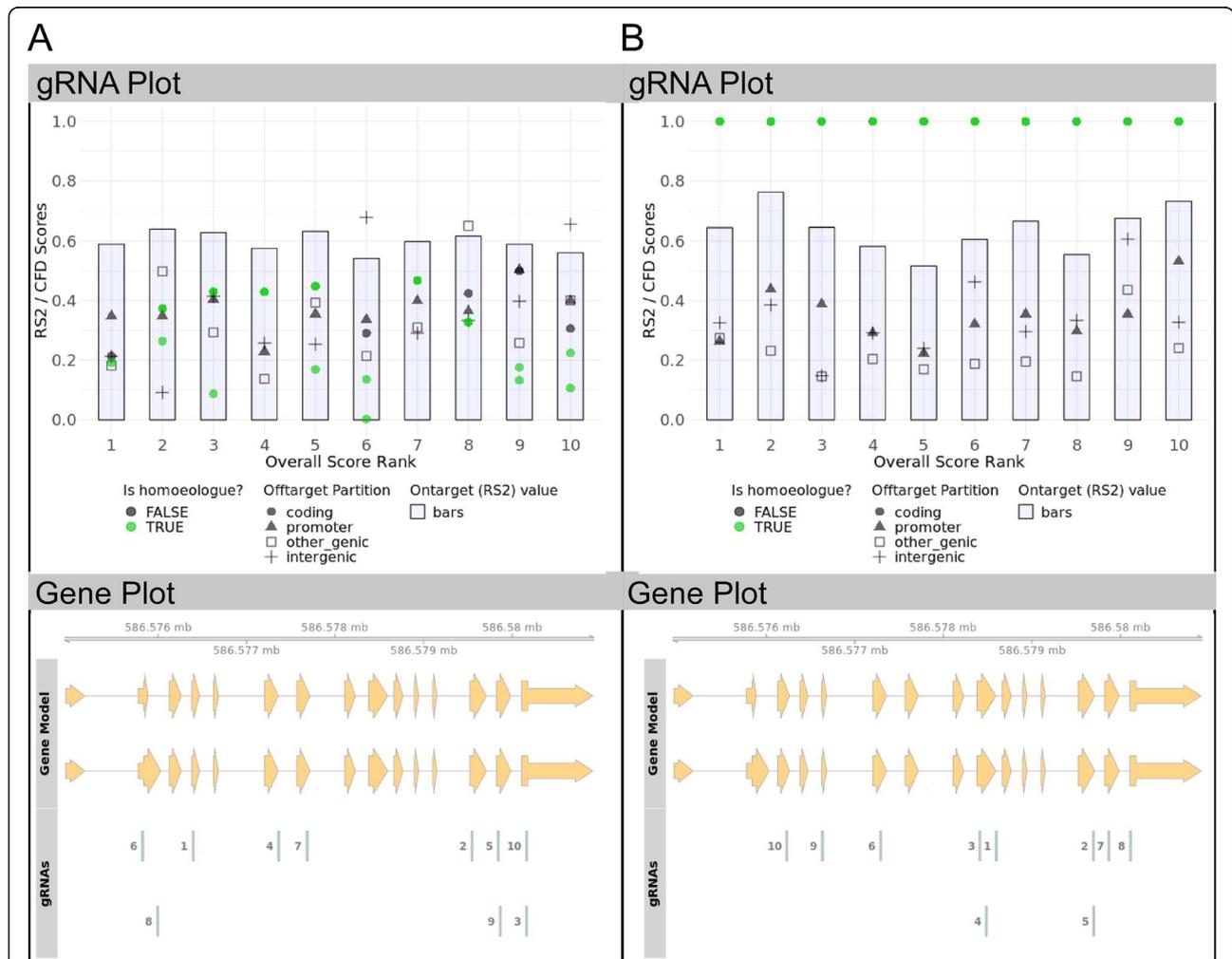


Fig. 3 Outputs from the WheatCRISPR web application. Exemplary gRNA and gene plots for the phytoene desaturase (PDS) produced using the WheatCRISPR web application are shown: **a** gRNAs specifically targeting TraesCS4B02G300100, and **b** gRNAs targeting all three homoeologues of the PDS gene (TraesCS4B02G300100, TraesCS4A02G004900, TraesCS4D02G299000). The gRNA plot displays a visualization of the rs2 and CFD scores and the score for any off-target hits to homoeologues. The blue-gray bars show the rs2 (on-target activity) score, the black points indicate the worst CFD (off-target activity) scores for each region (coding, promoter, other genic, intergenic), and the green points, if any, show the CFD scores for homoeologues. The gene plot displays the physical location of the gRNAs against the gene models. Each row in the gene model represents an isoform of the gene. The gray lines indicate introns and orange bars indicate exons. The thinner orange bars indicate UTRs

of the gene. In this variation, high CFD scores in homoeologues are rewarded while the maximum CFD in non-homoeologues remains penalized (Fig. 3b). Homoeologs were identified by the annotation available at ensembl-genomes.org. An overall score is used to rank all gRNAs for a gene so that the user can quickly identify the most likely candidate gRNAs. The overall scoring function is not based on any empirical evidence, so it is simply an intuitive estimate designed to help accelerate the process of finding effective gRNAs. Users are strongly encouraged to consider the individual rs2 and CFD scores, and other factors such as the location of gRNA within the protein coding region of the gene, before selecting a gRNA. The exact function used when targeting a specific gene (the default mode) is:

$$0.5(rs2) + 1 - (0.5(0.7(\max(\text{cfd_coding}, \text{cfd_promoter})) + 0.2(\max(\text{cfd_other_genic})) + 0.1(\max(\text{cfd_intergenic})))) + 0.5(rs2) + 1 - (0.5(0.7(\max(\text{cfd_coding}, \text{cfd_promoter})) + 0.2(\max(\text{cfd_other_genic})) + 0.1(\max(\text{cfd_intergenic}))))$$

and when targeting homoeologs is enabled:

$$0.33(rs2) + (1 - (0.33(0.7(\max(\text{cfd_coding}, \text{cfd_promoter})) + 0.2(\max(\text{cfd_other_genic})) + 0.1(\max(\text{cfd_intergenic})))))) + 0.34(\text{mean}(\text{cfd_hmlgs}))$$

Besides the predicted on-target and off-target activity metrics, the location of the gRNAs within a gene can also be important. It is often desirable to select gRNAs from exons that occur in all splice isoforms of a gene to ensure that all alternative transcripts are targeted. To identify the location of gRNAs within a gene, WheatCRISPR presents a genome browser-style Gene Plot with tracks for the gene models and the selected gRNA (Fig. 3).

The precomputed on-target to off-target mappings improve performance but limit the target sites to annotated genes. To search for targets outside annotated genes, WheatCRISPR also allows the user to paste in an arbitrary sequence of interest. In such cases, gRNAs are extracted, and off-target sequences and scores are computed on the fly. In this mode, functionality is limited for performance reasons. The maximum number of mismatches is limited as described in Table 2, and targeting a set of homoeologs is not possible.

To validate the accuracy of prediction of gRNA efficacy by WheatCRISPR, we compared the overall ranking of a subset of gRNAs and their targeting efficiency reported in the literature (Additional file 2: Table S2) [26–29]. The wheat gRNAs reported to be successful in targeting Q (the spelt factor gene), TaGW2 (wheat grain width and weight 2), TaLpx1 (wheat lipoxigenase 1), TaPDS1 (wheat phytoene desaturase 1) and INOX (inositol oxygenase) were predicted to have high on-target specificity and low off-target activity (i.e., ranked higher based on WheatCRISPR overall score). On the contrary, gRNAs that

either failed or had low success in editing TaGW2, TaDEP1 (wheat dense and erect panicle 1) and TaPIN1 (wheat PIN-FORMED 1) ranked lower (Additional file 2: Table S2). We have also confirmed the functionality of a number of gRNAs targeting wheat genes, such as TaPDS (Fig. 4) and puroindoline A [30], using an in vitro nuclease assay. These examples validate the prediction accuracy of WheatCRISPR and

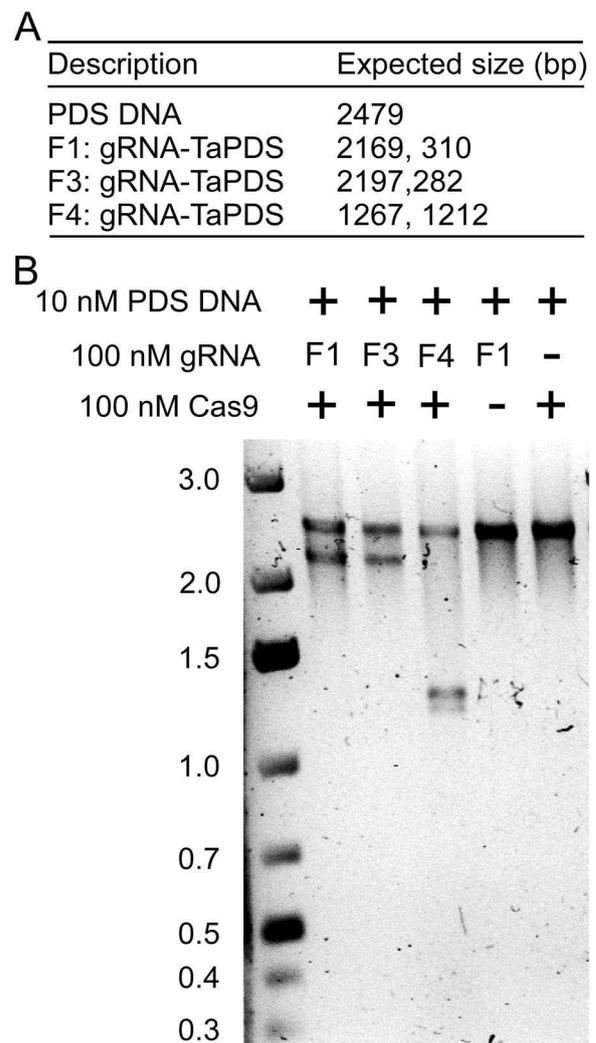


Fig. 4 Functional validation of TaPDS gRNAs by in vitro nuclease assay. **a** 2.479 kb fragment of wheat TaPDS (TraesCS4B02G300100) gene was amplified using wheat (Chinese Spring) genomic DNA and the primers TaPDS_F3 (5'-cgcagaggtgtttcacaagt-3') and TaPDS_R4 (5'-gagccatgctctctctacac-3'). **a** The expected band sizes of cleaved products using different guide RNAs against the 2.479 Kb input DNA. **b** In vitro nuclease assay of the input DNA using Cas9 and different guide RNAs. The Cas9 endonuclease (100 nM), guide RNAs (100 nM) and PCR amplified TaPDS target DNA fragment (10 nM) were mixed together at a molar ratio of 10:10:1 in a total assay reaction volume of 30 μ l. The reactions were incubated at 37 $^{\circ}$ C for 30 min. The assay was stopped by the addition of Proteinase K and products were analyzed using an agarose gel electrophoresis

demonstrate its utility in genome editing applications in wheat.

Conclusion

As an elegant alternative to reliance on natural or induced mutagenesis, CRISPR/Cas9-based gene editing technology has the potential to change the pace and course of crop breeding. To facilitate the application of this innovative technology in wheat, we have developed a robust web-based bioinformatics tool (WheatCRISPR) to enable selection of specific gRNAs for user-specified target gene or sequence and prediction of potential off-target sites. The current implementation of WheatCRISPR supports the selection of gRNAs to guide *S. pyogenes* Cas9 to genomic locations in the wheat genome. Identification of guide sequences with different PAMs reported for Cas9 variants, such as StCas9 (*Streptococcus thermophilus* Cas9), NmCas9 (*Neisseria meningitidis* Cas9), SaCas9 (*Staphylococcus aureus* Cas9) and FnCpf1 (*Francisella novicida* RNA-guided endonuclease) would be highly desirable. However, the reliance of Doench algorithms on empirical data (gRNA efficacy and specificity) specific to PAM sites of SpCas9 limits extension of WheatCRISPR to PAM sites of other Cas9 variants. Additionally, in wheat there will be a few genes for which finding unique gRNAs would be difficult due to polyploidy, high content of repetitive DNA content and genes typically existing as members of multi-gene families with high levels of sequence identity. In such cases, the users may have to consider other strategies (for example, dual gRNAs) to improve targeting specificity.

Availability and requirements

The WheatCRISPR web application is publicly available at <https://crispr.bioinfo.nrc.ca/WheatCrispr/>.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12870-019-2097-z>.

Additional file 1: Table S1. The gRNA table for a phytoene desaturase (TraesCS4B02G300100) gene.

Additional file 2: Table S2. Experimental corroboration and validation of gRNA efficiency predictions made by WheatCRISPR.

Abbreviations

Cas9: CRISPR-associated endonuclease; CFD: Cutting frequency determination; CRISPR: Clustered, regularly interspersed, palindromic repeats; crRNA: CRISPR RNA; gRNA: Guide RNA; PAM: Protospacer-Adjacent Motif; rs2: Rule set 2; tracrRNA: trans-activating crRNA; UTR: Untranslated region

Acknowledgements

We would like to thank Dr. Dan Tulpan for critical reading of the manuscript.

Authors' contributions

SK conceived the study. SK and AGS coordinated the study. DC designed and implemented the WheatCRISPR application. MK and PB provided technical and data analysis support. KR provided vectors and reporter

constructs. MB and IAPP contributed to the development of on the fly blast-based gRNA design tool. NR performed in vitro nuclease assays for gRNA validation. SK, MK and DC wrote the manuscript. All authors edited the manuscript and approved the final version.

Funding

This research was supported by the Canadian wheat improvement flagship program, National Research Council Canada. The funding agency had no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Availability of data and materials

All data generated or analysed in the study are included in this published article [and its supplementary information files].

Ethics approval and consent to participate

This research did not need ethics approval and consent as it did not involve human subjects, material or data.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹National Research Council Canada, 110 Gymnasium Place, Saskatoon, SK S7N 0W9, Canada. ²Agriculture and Agri-Food Canada, 107 Science Place, Saskatoon, SK S7N 0X2, Canada. ³Global Institute for Food Security, University of Saskatchewan, 110 Gymnasium Place, Saskatoon, SK S7N 4J8, Canada.

Received: 10 July 2019 Accepted: 23 October 2019

Published online: 06 November 2019

References

- Bhaya D, Davison M, Barrangou R. CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. *Annu Rev Genet.* 2011;45:273–97.
- Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science.* 2012;337(6096):816–21.
- Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science.* 2013;339(6121):819–23.
- Terns MP, Terns RM. CRISPR-based adaptive immune systems. *Curr Opin Microbiol.* 2011;14(3):321–7.
- Voytas DF, Gao C. Precision genome engineering and agriculture: opportunities and regulatory challenges. *PLoS Biol.* 2014;12(6):e1001877.
- Doudna JA, Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science.* 2014;346(6213):1258096.
- Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell.* 2014;157(6):1262–78.
- Sander JD, Joung JK. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat Biotechnol.* 2014;32(4):347–55.
- Anders C, Niewoehner O, Duerst A, Jinek M. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature.* 2014;513(7519):569–73.
- Arbuthnot P. Chapter 3 - engineering sequence-specific DNA binding proteins for antiviral gene editing. In: *Gene therapy for viral infections.* Amsterdam: Academic Press; 2015. p. 63–94.
- Doench JG, Fusi N, Sullender M, Hegde M, Vaimberg EW, Donovan KF, Smith I, Tothova Z, Wilen C, Orchard R, et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol.* 2016;34(2):184–91.
- International Wheat Genome Sequencing Consortium. Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science.* 2018;361(6403):eaar7191.
- Ding Y, Li H, Chen L-L, Xie K. Recent Advances in Genome Editing Using CRISPR/Cas9. *Front Plant Sci.* 2016;7:703.
- Wong N, Liu W, Wang X. WU-CRISPR: characteristics of functional guide RNAs for the CRISPR/Cas9 system. *Genome Biol.* 2015;16:218.

15. Bae S, Park J, Kim JS. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics*. 2014;30(10):1473–5.
16. Labun K, Montague TG, Gagnon JA, Thyme SB, Valen E. CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. *Nucleic Acids Res*. 2016;44(W1):W272–W276.
17. Heigwer F, Kerr G, Boutros M. E-CRISP: fast CRISPR target site identification. *Nat Methods*. 2014;11(2):122–3.
18. Lei Y, Lu L, Liu HY, Li S, Xing F, Chen LL. CRISPR-P: a web tool for synthetic single-guide RNA design of CRISPR-system in plants. *Mol Plant*. 2014;7(9):1494–6.
19. Naito Y, Hino K, Bono H, Ui-Tei K. CRISPRdirect: software for designing CRISPR/Cas guide RNA with reduced off-target sites. *Bioinformatics*. 2015; 31(7):1120–3.
20. Hodgkins A, Farne A, Perera S, Grego T, Parry-Smith DJ, Skarnes WC, Iyer V. WGE: a CRISPR database for genome engineering. *Bioinformatics*. 2015; 31(18):3078–80.
21. Liu H, Ding Y, Zhou Y, Jin W, Xie K, Chen LL. CRISPR-P 2.0: an improved CRISPR-Cas9 tool for genome editing in plants. *Mol Plant*. 2017;10(3):530–2.
22. Perez AR, Pritykin Y, Vidigal JA, Chhangawala S, Zamparo L, Leslie CS, Ventura A. GuideScan software for improved single and paired CRISPR guide RNA design. *Nat Biotechnol*. 2017;35(4):347–49.
23. Haeussler M, Schönig K, Eckert H, Eschstruth A, Mianné J, Renaud J-B, Schneider-Maunoury S, Shkumatava A, Teboul L, Kent J, et al. Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biol*. 2016;17:148.
24. Chuai GH, Wang QL, Liu Q. In Silico meets in vivo: towards computational CRISPR-based sgRNA design. *Trends Biotechnol*. 2017;35(1):12–21.
25. Doench JG, Hartenian E, Graham DB, Tothova Z, Hegde M, Smith I, Sullender M, Ebert BL, Xavier RJ, Root DE. Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat Biotechnol*. 2014; 32(12):1262–7.
26. Wang W, Pan Q, He F, Akhunova A, Chao S, Trick H, Akhunov E. Transgenerational CRISPR-Cas9 activity facilitates multiplex gene editing in allopolyploid wheat. *CRISPR J*. 2018;1(1):65–74.
27. Wang W, Simmonds J, Pan Q, Davidson D, He F, Battal A, Akhunova A, Trick HN, Uauy C, Akhunov E. Gene editing and mutagenesis reveal inter-cultivar differences and additivity in the contribution of TaGW2 homoeologues to grain size and weight in wheat. *Theor Appl Genet*. 2018;131(11):2463–75.
28. Zhang Y, Liang Z, Zong Y, Wang Y, Liu J, Chen K, Qiu JL, Gao C. Efficient and transgene-free genome editing in wheat through transient expression of CRISPR/Cas9 DNA or RNA. *Nat Commun*. 2016;7:12617.
29. Upadhyay SK, Kumar J, Alok A, Tuli R. RNA-guided genome editing for target gene mutations in wheat. *G3 (Bethesda)*. 2013;3(12):2233–8.
30. Rajagopalan N, Kagale S, Bhowmik P, Song H. A two-step method for obtaining highly pure Cas9 nuclease for genome editing, biophysical, and structural studies. *Methods Protoc*. 2018;1(2):17.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

