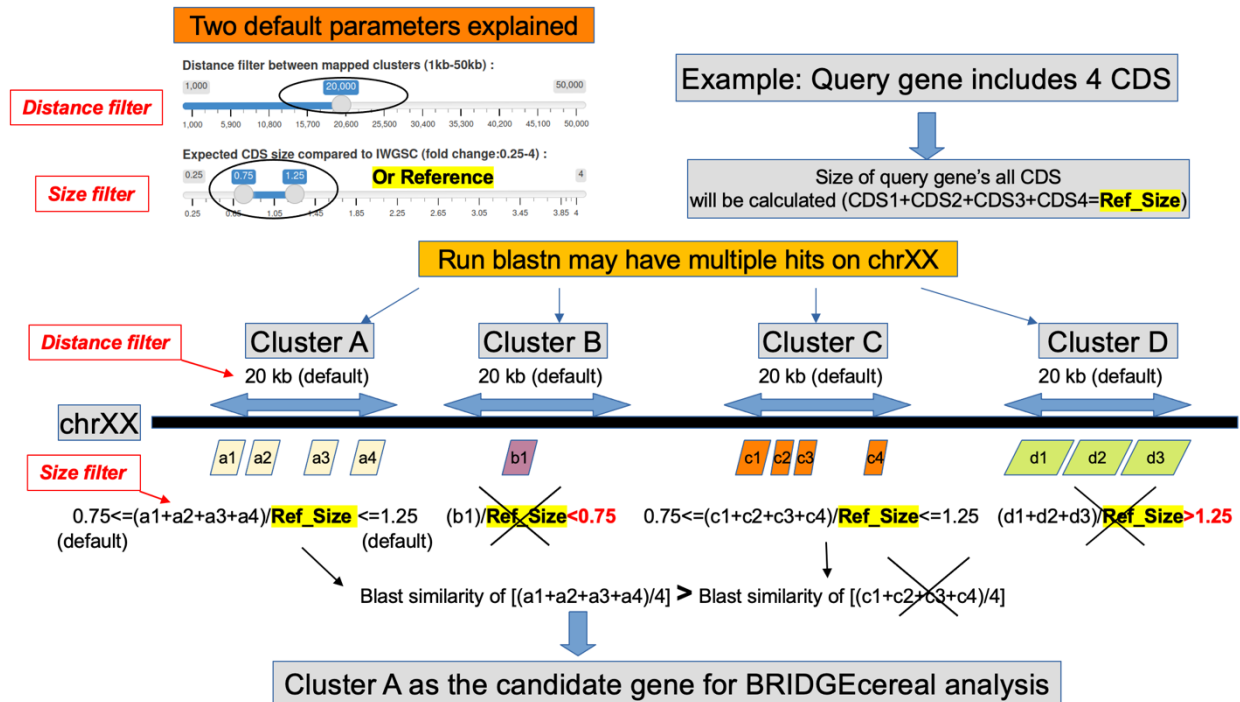


1, Two default parameters explained

- 1, If you expect large insertions in orthologs, please increase the distance filter.
- 2, If you expect that the ratio of (ortholog's coding region) / (query's coding region) is too large or too small, please adjust the size filter accordingly.



2, Paste sequence

Users want to blast a transcript sequence against pan-genomes

- 1). Input your sequence ID name, such as YourID, in the Gene Name box.

DO NOT click Button “(1) Check Gene ID”!

- 2). Manually select the Reference and Chromosome, select “fasta_seq” in the “CDS” box, then paste the sequence in the “Your fasta sequence” box. For your fasta sequence, the first line should be >YourID, the same name as the input shown in the Gene Name box.

- 3). Click “(2) Submit” button to start the process. User’s “YourID” query sequence together with other pan-genomes will be plotted in Panel 1.

Gene name (such as TraesCS4A02G058900 for IWGSC) or YourID for fasta sequence

YourID ← Job's name

(1) Check Gene ID **DO NOT click this button**

Pick Genome (Please select one!) :

IWGSC ← Reference genome (Wheat)

Chromosome (Please select one!)

chr4A ← Target chromosome (Wheat)

CDS (Coding sequence); OR your fasta sequence :

fasta_seq ← Select this option

Your fasta sequence (Please add first line: >YourID Before pasting your DNA sequence!)

>YourID ← Paste your sequence here with the name "YourID"

```

ATGGGTCGGGGGAAGGTGGAGATGAGGCGGATCGAGAACAAAGATAAGCCGGCAGGTGACGT
TCGCCAAGCGCCGGAATGG
GCTGCTCAAGAAGGCCTACGAGCTCTCGCTGCTCTGCGACGCCGAGGTCGCCCTCATCATCTT

```

3, Upload a contig or a chromosome

Users want to compare their sequenced chromosome (or large DNA fragment/contig) with pan-genomes on specific gene.

Prepare your *****.fa** and *****.fa.gz** files following this instruction: Rename query chromosome.fa file (Wheat example):

```

>chr4A
ATGCATGC...

```

Here, chr should be in lowercase!

For other crops (chromosome 4): Maize (>chr4); Sorghum (>chr4); Rice (>chr4); Barley (>chr4H)

For a contig or a large fragment containing your interested gene, just use chromosome as fasta file's first line. For example, your sequenced wheat contig/fragment contains a gene, which is known located at chromosome4A in the default reference genome IWGSC, your first fasta line should be written as >chr4A

Rename this file as:

Parent1_chr4A.fa (the second one as Parent2_chr4A.fa)

Then Run (to get Parent1_chr4A.fa.gz):

bgzip Parent1_chr4A.fa

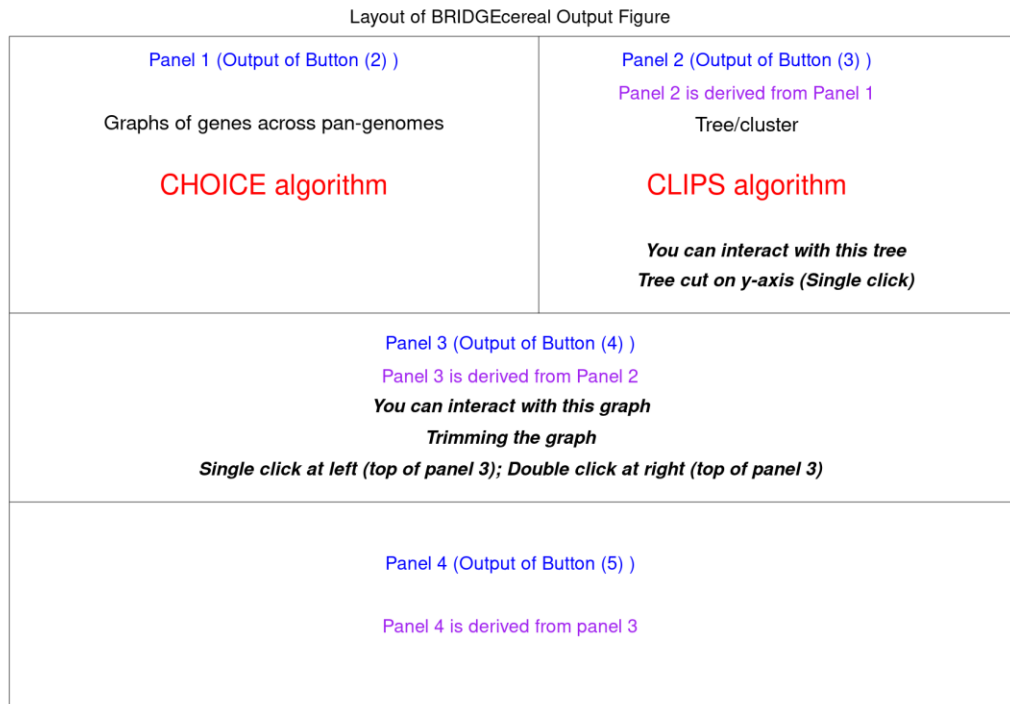
- 1). Input the gene model ID in the Gene Name box, then click “(1) Check Gene ID” button to fill the Boxes for Reference and Chromosome. A hyperlink, connected to the corresponding crop’s genomic database, will be available (in blue color) for the query gene ID.
- 2). To upload the compressed Parent1_chr4A.fa.gz file (max size=300 MB), click “Browse” button and upload your file, and BRIDGEcereal will let you know the file’s uploading progress ... (Time depends on your bandwidth speed).

The screenshot shows the BRIDGEcereal web interface. At the top, there's a box for 'Gene name (such as TransCS4A02G058900 for IWGSC) or Your ID for fasta sequence' with the value 'TraesCS4A02G058900' entered. Below it is a button labeled '(1) Check Gene ID'. To the right, there's a status bar showing 'There have been 1845 visitors!' and 'There're 6 genes or jobs have been submitted on this page! You can try yours.' Below the status bar, a blue link says 'The query gene is located at: chr4A:52,605,554-52,617,528, (+)strand.' Below the gene name box, there are dropdown menus for 'Pick Genome (Please select one):' (set to 'IWGSC'), 'Chromosome (Please select one):' (set to 'chr4A'), and 'CDS (Coding sequence): OR your fasta sequence:' (set to 'CDS'). Below these is a text area for 'Your fasta sequence (Please add first line: >YourID Before pasting your DNA sequence!)'. At the bottom, there's a file upload section with a 'Browse...' button and a file named 'Parent1_chr4A.fa.gz' selected. A button labeled 'Upload complete' is visible. To the right of the main form, there's a progress bar showing 'Processing your compressed file Parent1 ... almost there...' with an arrow pointing to it and text saying 'After uploading, a progress bar will give you an idea when you can submit this job'.

- 3). When you see the progress bar (blue) showing “Upload complete”, then click “(2) Submit (large file)” Button (yellow color) to run BRIDGEcereal. Your “Parent1” genome together with other pan-genomes will be plotted in panel 1.

The screenshot shows the filter section of the BRIDGEcereal web interface. It includes two input fields for 'Upstream (kb), max input should <=100 (kb)' and 'Downstream (kb), max input should <=100 (kb)', both set to '1.2'. Below these is a section for 'Genomes (Default: all genomes selected):' showing '12 items selected'. There are two sliders: 'Distance filter between mapped clusters (1kb-50kb):' with a value of '20,000' and 'Expected CDS size compared to Reference (fold change:0.25-4):' with a value of '1.25'. Both sliders have arrows pointing to them with text saying 'Filter used in CHOICE, most of time should be fine' and 'Filter used in CHOICE' respectively. At the bottom, there are two buttons: '(2) Submit' and '(2) Submit (large file)'. An arrow points to the '(2) Submit (large file)' button with text saying 'Click here to start'.

4, Figure layout



5, Arrangement of tables

Arrangement of BRIDGEcereal Output Tables

<p style="color: red; font-weight: bold;">Variety Table (From CLIPS algorithm and User's tree cut)</p> <p style="color: purple; font-weight: bold;">Varieties in each group based on User's tree cut</p>
<p style="color: blue;">Table1 (From CHOICE algorithm)</p> <p>Summary of all mapped positions (HSPs) on target chromosome and candidate gene selected for downstream analysis</p>
<p style="color: blue;">Table2 (From Table1's candidate cluster(gene))</p> <p>Candidate gene selected for plotting all figures</p>
<p style="color: blue;">Table3 (From Table1's Non-candidate clusters)</p> <p>Not used for BRIDGEcereal downstream analysis BUT, maybe helpful for troubleshooting And, provide all information about mapped positions (HSPs) for query gene</p>

6, Downloadable files

	Output files in the compressed .zip	Details
1	GeneID.png	Trimmed figure (Panel 4) in .png format
2	GeneID_Blast_Original	blast output (original), all HSPs on target chromosome
3	GeneID_Haplotype_syn	CHOICE processed blast output, filtered HSPs on target chromosome
4	GeneID_CDS.fasta	Gene's CDS fasta file
5	GeneID_Haplotype.fasta	genomic sequences (.fasta) extracted from pan-genomes
6	GeneID_Haplotype_N_Gaps	Gaps found in pan-genomes
7	GeneID_Haplotype-Self_out_m8	blast output of genomic DNA in pan-genomes (pairwise)
8	GeneID_ref_CDS-Haplotype_out_m8	blast output of CDS and pan-genomes
9	GeneID_repMask2	blast output of repeats
10	GeneID_User_Selected.fasta (optional)	output of User selected DNA (.fasta) based on trimmed coordinates