Tutorial: How to use the Wheat TILLING database

1. Visit <u>http://dubcovskylab.ucdavis.edu/wheat_blast</u> to go to the BLAST page or click on the 'Wheat BLAST' button on the homepage.

Dubc	COV	sky	Lab	INCOMENTS OF L	on location	N.	Stales -
Resources for W	/heat Ge	netics & Br	eeding	NP ANA	a Waxal	NACE N	
Home	Wheat	t BLAST	Breeding Program	Research 🔻	People 🔻	Tools 🔻	Links
Navigation	n	Home	2				
• Sign In		Welcome to Jorge Dubc	o the Dubcovsky Lab's webs ovsky Bio & Publications	ite.			
		Resear	ch projects				

2. A) Paste in your nucleotide sequence, B) select the 'TILLING Genomic Reference (4x and 6x)' database, and then C) hit the "Basic search" button.

	Basic Search - using default BLAST parameter settings
	Enter query sequences here in Fasta format or paste an example
A) ——>	>gi 33333044 gb AF543316.1 Triticum aestivum MADS box protein (MADS) mRNA, complete cds ATGGGGAGAGGGGAGGTGGAGCTCAAGCGGATCGAGAACAAGATCAACCGCC AGGTCACCTTCGCCAAGC GCCGCAACGGCCTGCTCAAGAAGGCCTACGAGGCTCTCCGTCCTCTGCGATGC CGAGGTCGCCTCATCAT
C) ——	Or upload sequence fasta file: Choose File No file chosen Program blastn C Database(s) TILLING Genomic Ref (4x and 6x) IWGSC AB + UCW (4x TILLING) IWGSC ABD + UCW (4x TILLING) IWGSC ABD + U (6x TILLING) TILLING Capture Design Padded Exons IWGSCV2.2 Psuedomolecules T. turgidum (Kronos) Transcriptome UwgsCv3.2 Psuedomolecules T. turgidum (Kronos) Transcriptome
	And/or upload sequence fasta file: Choose File No file chosen Basic search Reset

3. Wait 5-20 seconds while the BLAST search runs. The time is dependent on number of query sequences and the length of the query.

4. When finished, you will be presented with a table of BLAST hit results. Here, A) you can view the BLAST HTML output, B) download tabular BLAST TSV output, C) check the box for contig sequence you want to download, D) click on Scores to view the BLAST HTML alignment for a specific hit, and E) click a link to view the annotation with possible mutant lines in the same region.

From this page, you can also **download full contig sequences** based on your results.



5. After clicking on a visualization link, allow JBrowse to load the sequence and annotation data. This may take anywhere from 5 seconds to 3 minutes depending on connection speed.

Once finished loading, you will have your hit region highlighted in yellow, with your BLAST HSPs depicted by green bars in the BlastHSP Results track.

Below the Blast HSP Results track, we can see the High Confidence Mutants track where Single Nucleotide Variants (SNVs) are color-coded depending on the severity of the variant effect. SNVs that are most severe are colored in RED (stop_gained, stop_lost, splice_site changes), medium severity SNVs are colored in PURPLE (missense/amino acid change), the least severe SNVs are colored in GREEN (synonymous/no amino acid change), and lastly SNVs are colored in BLUE for intron_variants or when the effect is unknown (I.E. in the case of all UCW chromosome assembled contigs or anything that is not annotated in IWGSC).

						Last U	Jpdated: 9/7/16
Available Tracks	IWGSC-5BS 🝷 Fil	e View Help					
🗙 filter by text	0 1,000	2,000 3,	000 4,000	5,000 6,0	00 7,000	8,000 9,	9
BlastHSP Results	(ର୍ 🕀 🗨 🛛 IWGSC_	CSS_5BS_scaff_2265844	WGSC_CSS_5BS_sca	ff_2265844:110703 (10.7	
Genome Reference Annotation 3	dust gene:Trae	es_585_60390FC55		3,000			*
MIPS v22 Annotation TGAC Annotation TGAC+MIPs Merged Annotation		286					
Mutants 3							
 High-Confidence Mutants predicted Mutation Effect Low-Confidence Mutants predicted Mutation Effect Medium-Confidence Mutants predicted Mutation Effect 	BlastHSP Results		MyBLASTHit_2	MyBLASTHit_11 MyBLASTHit	MyBLASTHI	<mark>_3 MyBLASTHIL_1</mark> .STHIL_4 MyBLASTHIL_	_
Reference sequence					MyBLASTHIL_9 N	yBLASTHIT_5 MyBLA	A
Reference sequence					MyBLASTHit_8 MyBLASTHit_12	MyBLASTHit_10	
	High-Confidence Mut	ants predicted Mutation Eff	ect Kronos351:C3352T	Kronos382:C4828T SNV C -> T	Kronos167:G6158A SNV G -> A	Kronos198:G7596A SNV G -> A	
1					Kronos776:C6176T SNV C -> T	Kronos829:G7702A SNV G -> A	
					Kronos264 SNV G -> /	6:G6874A Kronos7 SNV C -	<i>r</i> ∹
					Kronos33 SNV G ->	1:G6933A A	

6. When clicking on a particular SNV, an information window will be displayed with details.



7. In order to **download nucleotide sequence data** for only a particular region, you can use the Reference Sequence track to export the sequence in FASTA format. First, zoom into the region of interest so that it is displayed in the visualization window.

To zoom, hold the <SHIFT> key and use your mouse to click and drag over the exact region you want. You should see a RED line with the positional start information after pressing SHIFT.



8. Then, click on the reference track down arrow and choose 'Save track data'. This track may be at the bottom of the interface.



9. You will then be presented with a dialog window where you can choose from a couple of options: Highlighted region, visible region, or whole reference sequence. In this case, since we are trying to grab only a particular sub-region of the highlighted region, we will choose 'Visible region' and then click the 'Save' button



10. To download SNP data in TSV format (can be opened in excel), right-click on a SNP for different export options.

ect <u>5:C79</u> 98T > -> T		Krono SNV (s2127:0 G -> A	8088A		Kronos2 SNV G -	576:G817 > A	6A				Kronos: SNV C
Kronos4 SNV C ->	648:C8036	šТ	🗌 Vie	w details								1
kron	002465-C8	0457	🧵 Do	wnload A	LL Sampl	e Data fo	r Kronos3	347				
SNV	C -> T	0401	🚺 Do	wnload K	ronos334	7 Data on	Contig IV	VGSC_C	SS_5BL_	scaff_108	69581	
			🚺 Dor	wnload A	LL Data o	n Contig	WGSC_C	CSS_5BL	_scaff_10	869581		
			🖶 See	ed Order	Form							

Alternate download option:

• Downloading mutation data can also be searched based on IDs from the Sample Search page: <u>http://dubcovskylab.ucdavis.edu/wheat-tilling/sample-search</u>

Filter by Kronos or Cadenza and then click result to download data. C Kronos Cadenza Soth
Line Name: Kronos2105
Example Search Kronos2105
Gene Name:
Example Search: Traes_1AS_BF353B963.3, Traes_1
Scaffold Name:
Example Search: IWGSC, IWGSC_CSS_1AL_scaff_1091068
Search by any combination of line, gene, or scaffold
Example Search: Traes_1AS_BF353B963.3,Kronos2101,Kronos2106, IWGSC_CSS_1AL_scaff_1091068
○ Kronos ○ Cadenza Both Submit

- The first 3 options (Line Name, Gene Name, or Scaffold Name) will search automatically and present you with a list of options. Clicking on one in the dropdown box will automatically download the mutation data for the search term. Please note the search is CASE SENSITVE.
- Additionally, the bottom box can be used to paste in mixed ID list to automatically generate one file for all the searched mutation data.