



University
of Idaho

**CANDIDATE GENE ANALYSIS OF THE
7DS QTL FOR DWARF BUNT
RESISTANCE USING TARGETED
CAPTURE SEQUENCING**

Guriqbal Dhillon, Pabitra Joshi, Yulin
Yan, Fangming Xiao, Will Krause,
Magaret Krause, and **Jianli Chen**

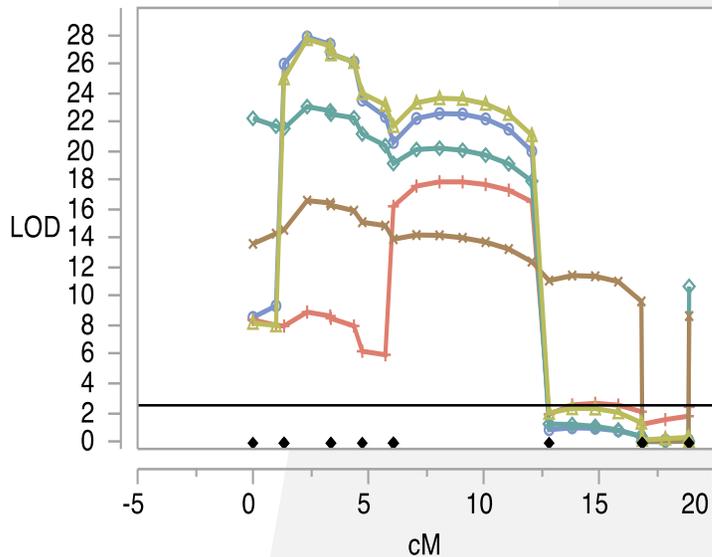
IMPORTANCE OF BUNT DISEASES



- Dwarf bunt caused by *T. controversa* and common bunt caused by *T. laevis* Kühn and *T. tritici* (Bjerk.) are two destructive diseases in wheat
- Cause significant yield losses and quality reduction
- Assessment of the disease is difficult and expensive
- Developing resistant wheat varieties crucial for sustainable crop management
- Requires a thorough understanding of the underlying genetic mechanisms
- Genomic analysis has emerged as a powerful tool to identify and characterize genes involved in resistance to bunt diseases



A MAJOR QTL *QDB.UI-7D* FOUND BY CHEN *ET AL.*, 2016



- Mapping population developed using IDO444 and Rio Blanco
- 159 RILs phenotyped across 4 environments in Logan, UT
- Genotyped using DArT markers and 9K SNPChip
- Three QTLs identified on chromosomes 1A, 2B, and 7DS
- QTL *QDB.ui-7D* detected with LOD 18.2 and R^2 43.4%
- Another QTL *QDB.ufa-7D* was also reported just downstream of the *QDB.ui-7D* 5 Mb apart (Muellner et al., 2020)

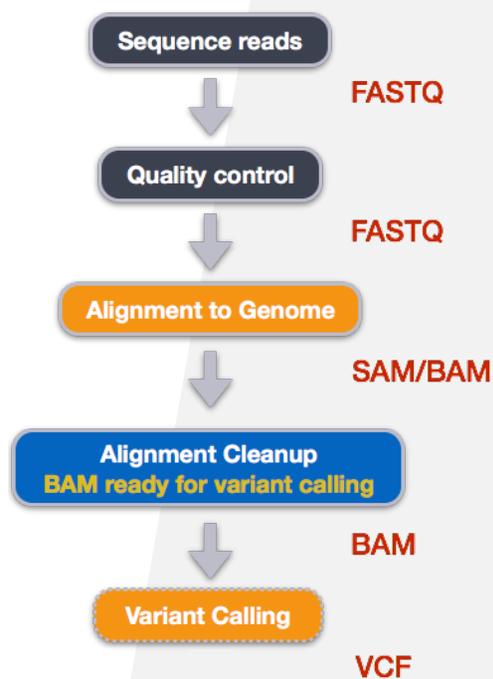


RESEARCH OBJECTIVES

- To use capture sequencing in a targeted region flanking the major QTL *QDB.ui-7D*
- To use bioinformatic tools to analyze and identify candidate genes underlining the *QDB.ui-7D*
- To validate the candidate genes using loss of function and gain of function experiments
- To develop the resistant cultivars using genomics assisted technologies



TARGETED CAPTURE SEQUENCING



- A 3-11 Mb region of the 7DS was targeted using Illumina PE151 paired-end sequencing.
- Resistant parent 'IDO444' and susceptible parent 'RioBlanco' plus 29 other lines were sequenced.
- More than 10 million reads per sample obtained.
- Low quality bases and adapter sequences removed.
- Filtered sequence mapped on refseqV1.0 using the BWA.
- Resulting raw BAM files were filtered to region of interest.
- Resulted in an average of 2.72 million reads per sample.



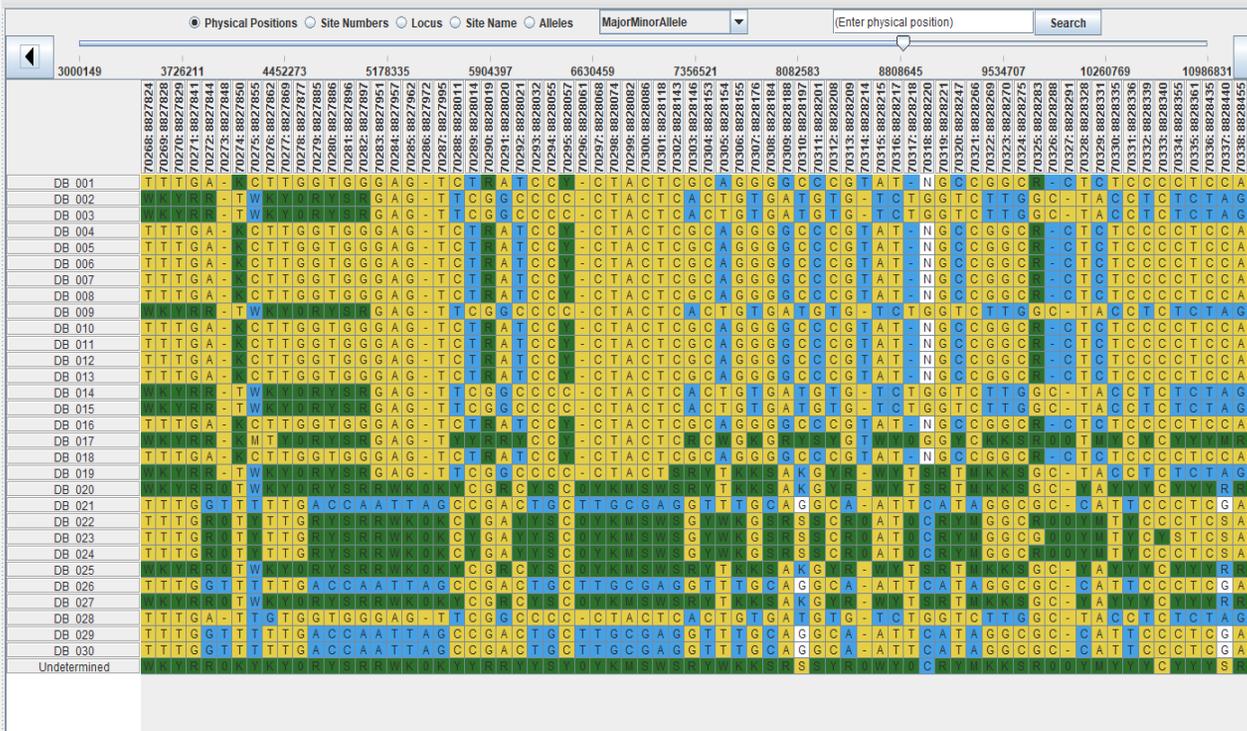
READ MAPPING STATISTICS

Sample	IDO444	IDO444 filtered	RioBlanco	RioBlanco filtered
reads mapped	18,140,184	2,736,437	16,352,209	2,284,793
reads mapped percent	99.820	99.784	99.787	99.744
raw total sequences	18,172,890	2,742,362	16,387,086	2,290,663
error rate	0.011	0.016	0.012	0.017
non primary alignments	0	0	0	0
reads MQ0 percent	7.824	7.391	8.275	7.118
reads properly paired percent	95.665	96.344	95.381	96.079

RAW VARIANTS: 94647 SITES FROM 3,000,149 - 10,986,839

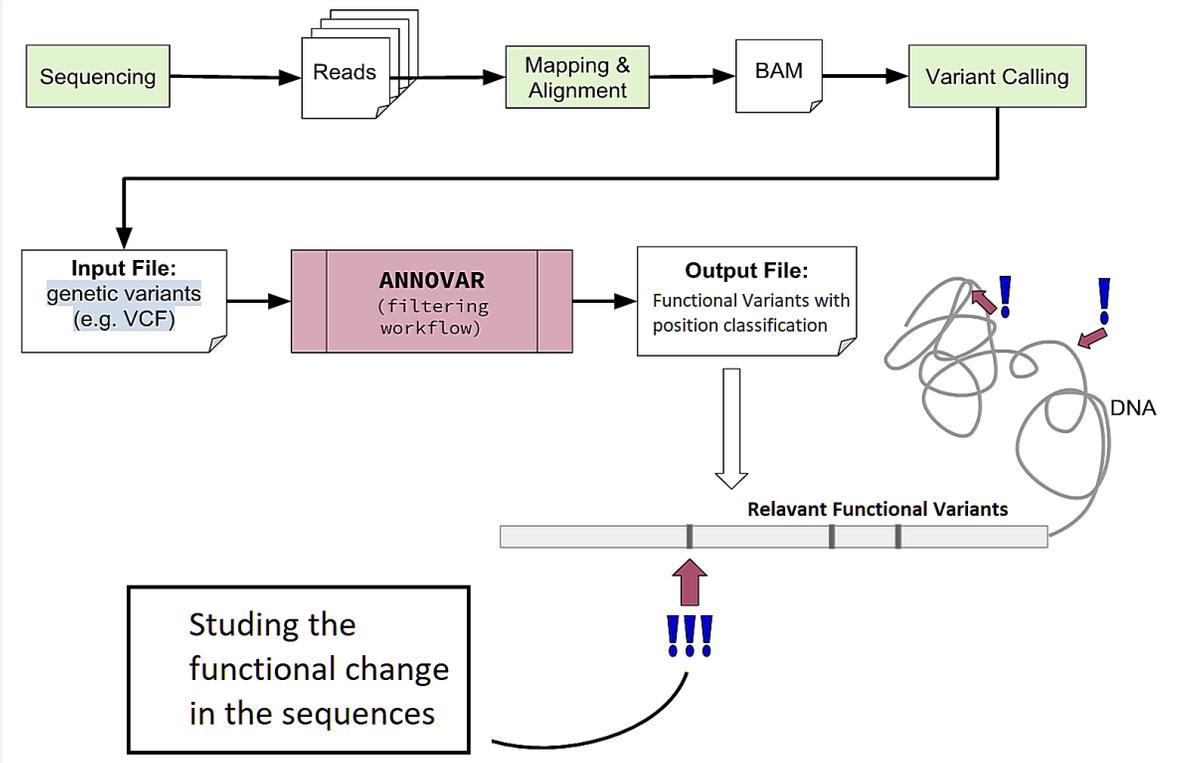


Code	Genotype
DB-001	IDO444
DB-002	RioBlanco
DB-003 to -019	RILs
DB-020 to -027	S Landraces
DB-028	Bt0
DB-029	Bt12
DB-030	Bt13





IDENTIFYING FUNCTIONAL VARIANTS (ANNOVAR)





POSITION AND FUNCTION-BASED CLASSIFICATION OF VARIANTS

Position Type	No of variants
intergenic	3111
intronic	115
exonic	444
UTR3	170
downstream	232
upstream	185
UTR5	8
ncRNA_exonic	2
All	4267

Mutation Type	Exonic variants
frameshift Indels	14
stopgain	4
synonymous SNV	179
nonsynonymous SNV	235
nonframeshift Indels	12
All	444



GENES WITH EXONIC VARIANTS

- 171 annotated genes spanning 145 genetic loci found in 3-11 Mb region on 7DS
- Only 29 genes showed variation in the exonic regions in IDO444
- These genes showed wide variation in the proteins encoded
 - Receptor Kinases
 - F-box proteins
 - NB-LRRS
 - Rp1-like proteins
 - Cytochrome P450
 - Ring finger like proteins



GENES (29) WITH EXONIC VARIANTS

Variant Type	Color Code
stopgain	green
synonymous SNV	red
frameshift Indel	Blue
Non Syn SNV	Black

GeneID	Function	Variants	NonSyn Variants
<i>TraesCS7D02G007400</i>	NBS-LRR disease resistance protein-like	1	1
<i>TraesCS7D02G008800</i>	Acid invertase 1	1	1
<i>TraesCS7D02G011600</i>	Acetyltransferase component of pyruvate dehydrogenase complex	1	0
<i>TraesCS7D02G011700</i>	Protein arginine N-methyltransferase 7	1	0
<i>TraesCS7D02G012300</i>	Zn-dependent exopeptidases superfamily protein	5	4
<i>TraesCS7D02G012400</i>	receptor kinase 1	19	13
<i>TraesCS7D02G012900</i>	Rp1-like protein	9	5
<i>TraesCS7D02G013200</i>	receptor kinase 1	39	17
<i>TraesCS7D02G013300</i>	Rp1-like protein	48	31
<i>TraesCS7D02G013400</i>	disease resistance family protein / LRR family protein	47	28
<i>TraesCS7D02G017600</i>	NBS-LRR-like resistance protein	25	11
<i>TraesCS7D02G017700</i>	NBS-LRR-like resistance protein	46	32
<i>TraesCS7D02G017900</i>	NBS-LRR disease resistance protein-like protein	32	16
<i>TraesCS7D02G019200</i>	Glycosyltransferase	2	0
<i>TraesCS7D02G020300</i>	F-box protein	18	9
<i>TraesCS7D02G020400</i>	Benzyl alcohol O-benzoyltransferase	3	1
<i>TraesCS7D02G021100</i>	BURP domain protein RD22	16	11
<i>TraesCS7D02G021700</i>	F-box protein	6	3
<i>TraesCS7D02G022000</i>	receptor-like protein kinase 1	17	6
<i>TraesCS7D02G022100</i>	Disease resistance protein (NBS-LRR class) family	27	16
<i>TraesCS7D02G022200</i>	Protein kinase family protein	6	5
<i>TraesCS7D02G022300</i>	Disease resistance protein (NBS-LRR class) family	4	1
<i>TraesCS7D02G022500</i>	Protein FRA10AC1	1	1
<i>TraesCS7D02G022600</i>	RING finger protein	1	1
<i>TraesCS7D02G022900</i>	Invertase inhibitor	2	1
<i>TraesCS7D02G023000</i>	RING finger protein	9	4
<i>TraesCS7D02G023100</i>	RING finger family protein	2	1
<i>TraesCS7D02G023300</i>	Cytochrome P450	21	8
<i>TraesCS7D02G023400</i>	F-box protein	35	26



POTENTIAL TARGETS IN NINE GENES

ID	IDO444	RioBlanco	Pos	Mutation	Gene	Function	Mutation NT	Mutation AA	Mutant domains	GeneStart	GeneStop	Size(kb)
S7D_4268568	AG	A	exon3	frameshift Indel	<i>TraesCS7D02G008800</i>	Acid invertase 1	c.1369delC	p.L457fs	C-terminal Transmembrane?	4267640	4270860	3.22
S7D_5331386	G	GTT	exon3	frameshift Indel	<i>TraesCS7D02G012400</i>	receptor kinase 1	c.211_212insTT	p.G71fs	Before binding site of protein kinase	5330809	5337313	6.504
S7D_5331411	ACG	A	exon3	frameshift Indel	<i>TraesCS7D02G012400</i>		c.237_238del	p.D79fs	Before binding site of protein kinase	5330809	5337313	6.504
S7D_5512644	AAATAG	A	exon8	frameshift Indel	<i>TraesCS7D02G013200</i>	receptor kinase 1	c.3939_3943del	p.F1313fs	C-terminal non domain region	5510346	5518574	8.228
S7D_5512713	AT	A	exon8	frameshift Indel	<i>TraesCS7D02G013200</i>		c.3874delA	p.I1292fs	C-terminal non domain region	5510346	5518574	8.228
S7D_5634198	G	A	exon1	stopgain	<i>TraesCS7D02G013300</i>	Rp1-like protein	c.C2602T	p.Q868X	Terminal end (C) of protein	5634184	5636799	2.615
S7D_5648558	TC	T	exon1	frameshift Indel	<i>TraesCS7D02G013400</i>	disease resistance family protein	c.49delG	p.E17fs	N-terminal before domain	5645812	5648607	2.795
S7D_7750174	T	C	exon2	stopgain	<i>TraesCS7D02G017700</i>	NBS-LRR-like resistance protein	c.C106T	p.R36X	in Rx N-terminal domain	7749593	7754061	4.468
S7D_7750708	C	T	exon2	stopgain	<i>TraesCS7D02G017700</i>		c.C640T	p.Q214X	in NB-ARC domain	7749593	7754061	4.468
S7D_7752652	A	AGCTTGGT AACGTTCC CGAT	exon2	frameshift Indel	<i>TraesCS7D02G017700</i>		c.2585_2603del	p.S862fs	after NB-ARC domain	7749593	7754061	4.468
S7D_7752984	CACAA	C	exon2	frameshift Indel	<i>TraesCS7D02G017700</i>		c.2916_2917insACAA	p.V972fs	after NB-ARC domain	7749593	7754061	4.468
S7D_10478651	TC	T	exon2	frameshift Indel	<i>TraesCS7D02G022100</i>	Disease resistance protein (NBS-LRR class) family	c.919delG	p.E307fs	after NB-ARC domain	10477394	10480078	2.684
S7D_10478655	T	TGGTA	exon2	frameshift Indel	<i>TraesCS7D02G022100</i>		c.915_916insTACC	p.I306fs	after NB-ARC domain	10477394	10480078	2.684
S7D_10478659	CG	C	exon2	frameshift Indel	<i>TraesCS7D02G022100</i>		c.911delC	p.S304fs	after NB-ARC domain	10477394	10480078	2.684
S7D_10478787	T	A	exon2	stopgain	<i>TraesCS7D02G022100</i>		c.A784T	p.R262X	after NB-ARC domain	10477394	10480078	2.684
S7D_10869822	TA	T	exon1	frameshift Indel	<i>TraesCS7D02G023300</i>	Cytochrome P450	c.512_513insT	p.A171fs	after cytochrome domain transmembrane?	10868164	10870414	2.25
S7D_10963456	T	TGC	exon1	frameshift Indel	<i>TraesCS7D02G023400</i>	F-box protein	c.203_204del	p.R68fs	in F box domain	10962062	10963681	1.619



ONGOING PROJECTS – USDA-NIFA AFRI

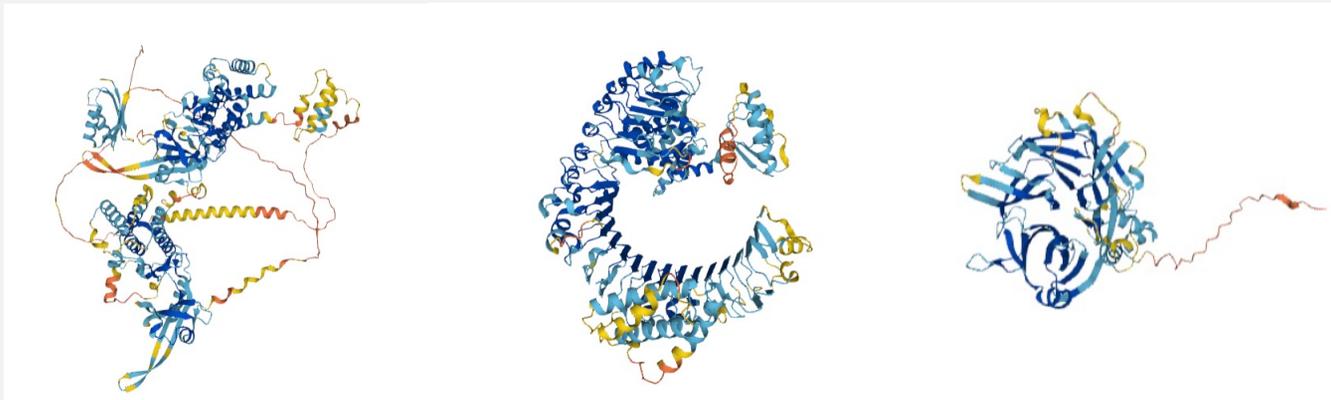
Validation of the candidate genes

- Three genes (primary targets) being cloned into vectors
- Transformation of RioBlanco with cloned genes
- Studying gain of function in transformed RioBlanco
- Studying loss of function using Crisper/Cas9 mediated transformation
- Studying loss of function using EMS mutation lines

ONGOING PROJECTS – USDA-NIFA AFRI

Validation of the candidate genes

- Sequencing the three candidate genes
- Developing in-situ gene models to compare the structural variations in the protein folding of the candidate genes which, in turn, would reveal the mode of action of the genes in imparting resistance to the bunt diseases.

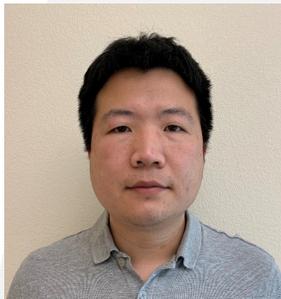




ONGOING PROJECTS – USDA-NIFA AFRI

- Fine mapping of the two major QTL, *QDB.ui-6DL* and *QDB.ui-7D*
- Developing genomic selection model to select major and minor genes
- Implementing marker assisted selection to pyramid multiple resistance genes
- Releasing bunt resistance cultivars

COLLABORATIVE TEAM



ACKNOWLEDGEMENT



University of Idaho

Idaho Agricultural Experiment Station