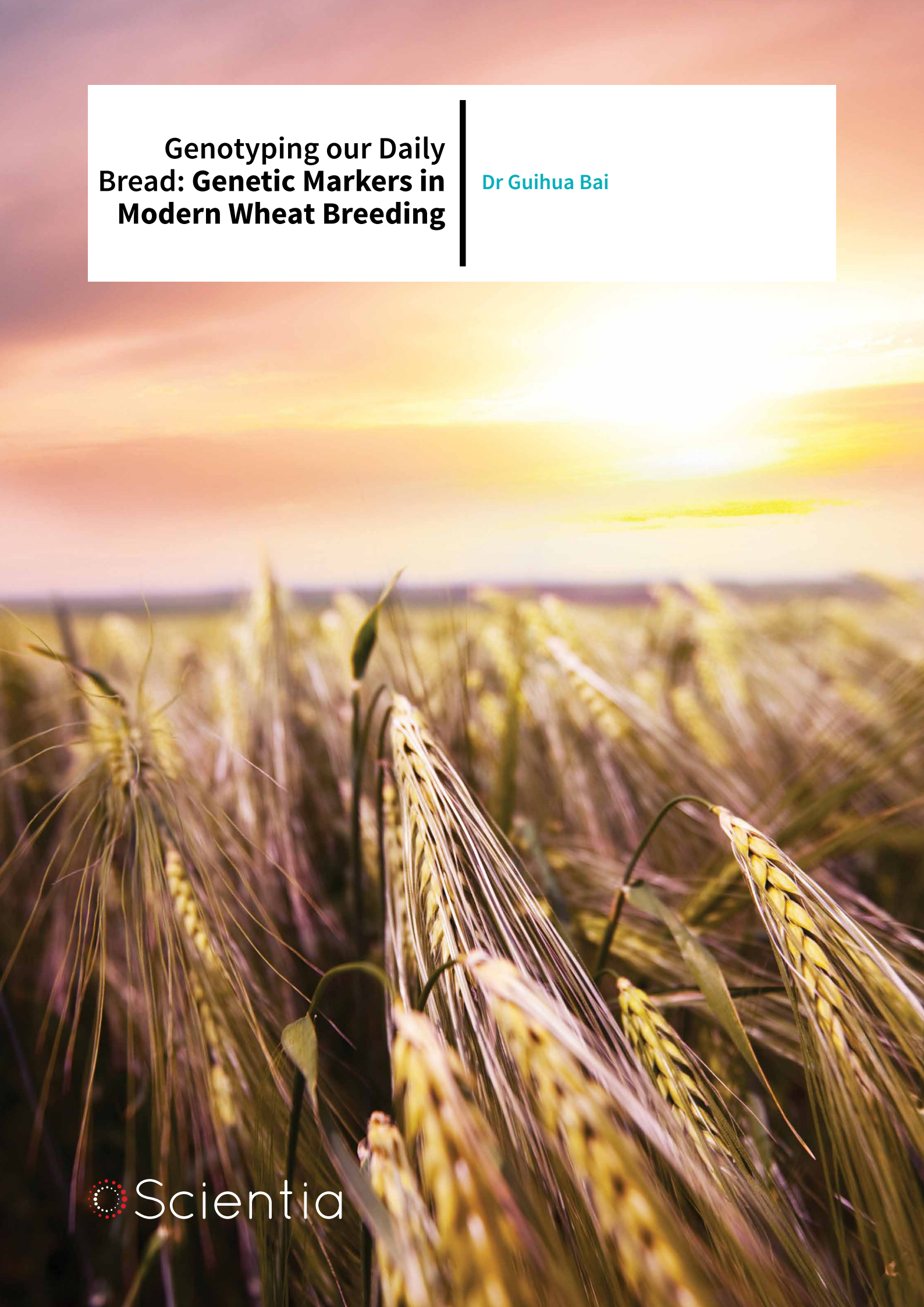
Genotyping our Daily Bread: **Genetic Markers in Modern Wheat Breeding**

Dr Guihua Bai



Scientia

GENOTYPING OUR DAILY BREAD: GENETIC MARKERS IN MODERN WHEAT BREEDING

Wheat is a staple crop, and is of utmost importance for global food supply. **Dr Guihua Bai** and his group at USDA-ARS conduct wheat genomics research to analyse wheat DNA markers, and assist in the breeding of new cultivars, to ensure high grain yields and quality, as well as resistance to multiple biotic and abiotic stresses.

Norman Borlaug (1914–2009), American agronomist and father of the Green Revolution, saw food security as a global priority, stating, 'Without food, people perish, social and political organisations disintegrate, and civilisations collapse.' Raised in rural Iowa by a Norwegian-American farming family during the Great Depression, Borlaug was all-too familiar with the socio-economic implications of food supply, and the inefficiencies of the standard agricultural practices of his day. Borlaug went on to meticulously cross-breed thousands of wheat varieties to produce advantageous traits such as high yields, disease-resistance, early maturation, dwarfism and tolerance to environmental conditions. Borlaug's pioneering programme kick-started the Green Revolution that saw global crop yields skyrocket. The productivity gains from Borlaug's agricultural innovations are thought to have saved over a billion people from starvation, as well as preserved large tracts of pristine wilderness from being converted to farmland. It was not for nothing then that Borlaugh was awarded the Presidential Medal of Freedom, the Congressional Gold Medal, and the Nobel Peace Prize, being one of only six people in history to receive these three prestigious awards.

Borlaugh's focus on wheat production to ensure food supply and security was well justified. Wheat is a staple food source for billions, with 729 million tonnes produced worldwide in 2014. While Borlaugh's efforts were revolutionary, further innovations in wheat breeding and production are necessary to feed a growing human population. But first, it is essential to understand the evolution and genetics of modern wheat varieties in order to evaluate the complex interplay between genotype, phenotype and environment.

The Complicated Phylogeny of Wheat

Cultivation of wheat (*Triticum* spp.) is thought to have started in the Fertile Crescent around 9600 BC and spread worldwide. Due to domestication and selective breeding of wheat, modern wheat varieties bear little resemblance to their wild ancestors, having ears that hold more grains, as well as greater tolerance to environmental conditions and features amenable to modern farming practices. Dramatic increases in genome size are a characteristic feature of the grass family (*Poacea*), of which wheat is a part of. Therefore, cereal crop plants are typically polyploid, having more than two pairs of homologous (pairing) chromosomes.

Most cultivated varieties of wheat belong to two species, the durum wheat (*Triticum durum or Triticum turgidum* subsp. *durum*) used in pasta and semolina, and bread wheat (*Triticum aestivum*). The phylogeny of both common domesticated wheat species is somewhat complicated by a series of genome duplications, hybridisations between species and trait selection. Around 7 million years ago, an ancestral *Triticeae* gave rise to two divergent diploid precursor genomes, denoted AA and BB. Then, about 5 million years ago, these formed the diploid DD genome (2n) of *Aegilops* genus (goatgrass). About 800,000 years ago, tetraploid wild emmer wheat appeared (*Triticum turgidum*), formed from a hybridisation between ancestral wheats with AA and BB genomes. The domestication of emmer wheat reflects the dawn of agriculture, as hunter-gatherers began deliberately cultivating wild emmer wheat. The selection of these first farmers resulted in plants with larger grains, giving rise to domesticated emmer wheat (*Triticum turgidum* subsp. *dicoccum*). Further selection produced durum wheat (*Triticum durum or Triticum turgidum* subsp. *durum*) around 9,000 years ago, the only tetraploid species of wheat of commercial importance that is widely cultivated today. Expansion of emmer wheat cultivation into wild areas led to chance of hybridisation between domesticated emmer (AABB) and wild goatgrass (DD), producing today's hexaploid bread wheat (*Triticum aestivum*) (AABBDD), with kernels that can easily be released from the spike, allowing free-threshing.

The Quest for Wheat Genetic Markers

Dr Guihua Bai from USDA-ARS is passionate about wheat genetics, and considers this important for the future of wheat cultivation. He is the principal investigator in the USDA Central Small Grain Genotyping Center in Manhattan, Kansas and provides DNA marker service to breeders in the hard winter wheat regions in the US Great Plains. His group applies state-of-the-art modern genotyping techniques to discover and validate wheat DNA markers that are subsequently used to select important traits. The genotyping data can then provide selection guides to wheat breeders seeking to create new cultivars in response to specific challenges, such as newly emergent plant pathogens or climate change. This is no easy feat, given the size and complexity of three sub-genomes of bread wheat.

The typical genome of *Triticum aestivum* is composed of 17 Gbp (giga-base pairs) of DNA, of which 80–90% is made of repetitive sequences, presenting an enormous challenge for DNA sequencing and analysis. The partial draft sequence of the *Triticum aestivum* genome in 2010 by a group in the UK represented a milestone in wheat genomic research. However, a recent coordinated international endeavour (http://www.wheatinitiative.org/) on subsequent re-sequencing, sequence re-assembly, and gene identification and annotation, gave birth to a relatively complete wheat genome reference sequence. The newly assembled reference sequence provides critical resources for gene cloning, DNA marker identification, and elucidating the genotype-phenotype relationships of agriculturally relevant varieties that Dr Bai focuses on.

The analysis of DNA markers lies at the heart of Dr Bai's research. Of particular interest to his team is the analysis of quantitative trait loci (QTLs) or genes that are important to wheat production. A QTL is a DNA section (a locus), usually containing one or more genes, responsible for a particular quantitative phenotypic trait. Genes are DNA sections encoding different amino acid sequences that make up proteins, which, at the molecular level, are ultimately responsible for determining physiological phenotypes. In most of the cases, the identity of gene(s) underlying a QTL is difficult to discover, but the DNA fragments nearby the gene can be easily identified and serve as landmarks for selecting the gene(s) in breeding. These polymorphic DNA markers physically close to genes of interest on a chromosome are more likely to be inherited together with the gene (DNA sections are 'shuffled' during meiosis) – this is known as genetic linkage. While most QTLs are postulated to have a minor effect on plant physiology, Dr Bai's group seeks to identify the ones that have major and stable effects using DNA markers.

In traditional breeding, breeders directly select the trait of interest, and these selections need to be done under certain environments. For example, selection for disease resistance has to be done where the plants are severely infected by the disease, whereas high yield traits have to be selected under disease free conditions. New cultivars must show high yields and good quality, and also resistance to multiple diseases, each of which needs to be selected in a different environment in traditional breeding. Using DNA marker-assisted selection (MAS), these traits can be selected by testing a small piece of leaf at the early seedling stage to predict presence of target genes or QTLs, significantly reducing phenotyping costs, speeding up the breeding process, and increasing selection accuracy. Dr Bai's group provides marker analysis services to breeders in the US Great Plains to help them deploy MAS in US hard winter wheat breeding.

As new generation sequencing (NGS) technologies, DNA marker technologies have been significantly improved. SNPs (single nucleotide polymorphisms) including insertions/deletions ('indels') are common polymorphic markers for genotyping. Polymorphic SNP markers can be easily found by comparing DNA sequences between genotypes. Recently, Dr Bai and his team found that a single SNP tightly linked to a major QTL on chromosome 7A of *T. aestivum* had a major impact on kernel

Wheat with a scab resistance gene, Fhb1, transferred by marker-assisted selection (right) and without Fhb1 (Left)

Wheat with TaPHS1 (left) and without TaPHS1 (right)

length and weight. They did this by employing an SNP Chip – a DNA chip technology that can simultaneously identify a large number of DNA polymorphisms for QTL identification. They then converted the chip-based SNP to a breeder-friendly marker type called KASP assay, a PCR-based technique involving the use of allele-specific primers and fluorescent probes, so breeders can use this KASP-SNP marker to directly predict the presence of the allele of interest. The group has recently demonstrated the use of NGS for low-cost multiplex analysis of several polymorphic markers associated with various important wheat traits, including disease and insect resistance, grain quality and plant height, which significantly reduced genotyping cost and increased efficiency.

Genomic Approaches to Tackling Pre-Harvest Sprouting

Long seed dormancy (SD) is advantageous to wild grasses in a harsh environment, allowing germination only under optimal conditions. This is a feature that has been 'bred out' of modern wheat varieties by artificial selection, as short and predictable germination is conducive to high productivities and straightforward breeding. However, a major disadvantage of short SD is propensity for pre-harvest sprouting (PHS), which involves premature germination of kernels in a spike before harvest. PHS is problematic for farmers as it causes significant losses in yields and grain quality. In regions with high rainfall during harvest season, the cultivation of wheat cultivars with appropriate SD and resistance to PHS is advocated. Identification of alleles and QTLs associated with PHS resistance or susceptibility, to inform breeding strategies, is a high priority research theme within the USDA Central Small Grain Genotyping Center.

TaPHS1 is a gene known to regulate PHS resistance on chromosome 3A in wheat. Dr Bai's group cloned the gene and identified two mutations in the positions +646 and +666 of the *TaPHS1* coding region that result in wheat PHS susceptibility in a white wheat Rio Blanco. Dr Bai's group screened 327 wheat accessions of wild and domesticated wheat progenitors with A genome using three KASP markers based on the two *TaPHS1* mutations and one mutation in the promoter region. The wheat species were also assessed for PHS tolerance under greenhouse conditions, to correlate the genotype to the phenotypic trait. It was found that most accessions of wild wheat progenitors were highly PHS-resistant, and all lacked the +646 mutation. Hexaploid wheat and *T. durum* were found to have the highest sprouting rates, demonstrating the diversification of SD during domestication. Interestingly, it was found that +646 mutation occurred independently in *T. monococcum* and *T. aestivum*, driven by the same selection pressure for reduced seed dormancy.

Plant breeders and farmers alike have sought to determine quantitative indicators of PHS susceptibility. What better indicator than grain colour (GC)? Colorimetric analysis is intuitive and amenable to straightforward observation without specialist equipment. Previous anecdotal and empirical observations have suggested that GC is associated with PHS, with red-grained wheats more tolerant to PHS than white-grained wheat. Dr Bai's group have confirmed this relationship between GC and PHS tolerance with a genome-wide association study (GWAS) in 185 elite lines and cultivars. For these cultivars, both sprouting studies in greenhouses and field experiments, and genotyping studies were carried out with wheat 9K and 90K SNP arrays. The GWAS study found that a number of genes, mainly on group 3 chromosomes, that control GC also regulate PHS resistance, but this relationship was observed in the field but not the greenhouse, highlighting the importance of environmental conditions on triggering gene-encoded phenotypes. In addition, several genes that do not relate to GC were also identified, which are important for breeding PHS tolerant white wheat.

Marker-Assisted Selection in Wheat Breeding – The Way Forward

In a sense, Dr Bai is following in the footsteps of Norman Borlaug, who began a revolutionary wheat breeding programme 50 years ago. However, plant breeding techniques have moved on, with new genomic tools and approaches that would have been unknown to Borlaugh in the 1960s. Dr Bai and his group deploy the latest genomic technologies to analyse QTLs and associated DNA markers, to probe the complex relationships between genotype and phenotype in wheat. They then utilise these insights to inform breeding strategies, with the aim of producing elite cultivars enriched with advantageous QTLs, while minimising deleterious QTLs. This is no easy task, given the complexities of gene segregation, and the reduced gene pool from millennia of selection. It is hoped that marker-assisted selection, coupled with innovations in agricultural practices, will fulfil the original vision of Borlaug's Green Revolution and further boost productivities to feed an ever-growing human population.

Meet the researcher

Dr Guihua Bai

Hard Winter Wheat Genetics Research Unit,
USDA-ARS
Manhattan, Kansas, USA

Dr Guihua Bai is an eminent researcher in wheat molecular genetics and breeding. His research interests include wheat genomics, analysis of wheat genes, QTLs and DNA markers for marker-assisted selection, wheat transcriptomics, and characterising and cloning wheat resistance genes to Fusarium head blight, leaf rusts, soil-borne mosaic virus and pre-harvest sprouting. From an early age, he has had a passion for agronomy, having completed his B.S. in Agronomy in 1982 and his M.S. in Plant Genetics and Breeding in 1985, both at Nanjing Agricultural University, Nanjing, China. After completing his M.S. degree, he worked as a wheat breeder at the Food and Crop Institute, Jiangsu Academy of Agricultural Science, Nanjing for four and a half years, before going on to gain a PhD at the Department of Botany and Plant Pathology, Purdue University, in 1995. He has since undertaken a number of research roles in Texas Tech University, University of Illinois, and USDA/ARS/NCAUR, before taking on an Assistant Professorship in Wheat Genetics and Breeding at Oklahoma State University between 1999 and 2002. Since then, he has settled in his present home in Manhattan, Kansas, having become the Director of the USDA Central Small Grain Genotyping Center and Research Plant Molecular Geneticist at USDA/ARS/HWWGRU, and an Adjunct Professor in the Department of Agronomy, Kansas State University.

CONTACT

E: guihua.bai@ars.usda.gov
T: (+1) 785 532 1124
W: http://hwwgenotyping.ksu.edu

KEY COLLABORATORS

Allan K Fritz, Kansas State University
Brett F Carver and Liuling Yan, Oklahoma State University
Scott Haley, Colorado State University
Stephen Beanziger, University of Nebraska
Sunish Sehgal, South Dakota State University
Jackie Rudd and Amir Ibrahim, Texas A&M University
Robert Graybosch, Xiangyang Xu, and James Kolmer, USDA-ARS

FUNDING

USDA, US Wheat and Barley Scab Initiative
National Research Initiative, USDA National Institute of Food and Agriculture.

REFERENCES

M Lin, D Zhang, S Liu, G Zhang, J Yu, AK Fritz and G Bai, Genome-wide association analysis on pre-harvest sprouting resistance and grain color in U.S. winter wheat, BMC Genomics, 2016, 17, 794.

Z Su, S Jin, Y Lu, G Zhang, S Chao and G Bai, Single nucleotide polymorphism tightly linked to a major QTL on chromosome 7A for both kernel length and kernel weight in wheat, Molecular Breeding, 2016, 36, 15.

A Bernardo, S Wang, P St. Amand and G Bai, Using Next Generation Sequencing for Multiplexed Trait-Linked Markers in Wheat, PloS One, 2015, 10, e0143890.

S Liu, SK Sehgal, M Lin, J Li, HN Trick, BS Gill and G Bai, Independent mis-splicing mutations in TaPHS1 causing loss of pre-harvest sprouting (PHS) resistance during wheat domestication, New Phytologist, 2015, 208, 928.

S Liu, SK Sehgal, J Li, M Lin, HN Trick, JM Yu, BS Gill and G Bai. Cloning and characterization of a critical regulator for pre-harvest sprouting in wheat, Genetics, 2013, 195, 263.

USDA Agricultural Research Service