Selection signatures across seven decades of hard winter wheat breeding in the Great Plains of the United States

Article *in* The Plant Genome · September 2020

DOI: 10.1002/tpg2.20032

CITATIONS
0

READS
10

5 authors, including:

Habtamu Ayalew
Noble Research Institute
37 PUBLICATIONS **184** CITATIONS

SEE PROFILE

Mark E Sorrells
Cornell University
442 PUBLICATIONS **29,991** CITATIONS

SEE PROFILE

P. Stephen Baenziger
University of Nebraska at Lincoln
529 PUBLICATIONS **8,799** CITATIONS

SEE PROFILE

Received: 10 December 2019 | Accepted: 7 May 2020

DOI: 10.1002/tpg2.20032

The Plant Genome

ORIGINAL RESEARCH

Selection signatures across seven decades of hard winter wheat breeding in the Great Plains of the United States

Habtamu Ayalew[1] | **Mark E. Sorrells**[2] | **Brett F. Carver**[3] | **P. Stephen Baenziger**[4] | **Xue-Feng Ma**[1]

[1] Noble Research Institute, Ardmore, OK 73401, USA

[2] Plant Breeding and Genetics, School of Integrative Plant Science, Cornell University, Ithaca, NY 14853, USA

[3] Department of Plant and Soil Sciences, Oklahoma State University, Stillwater, OK 74078, USA

[4] Department of Agronomy and Horticulture, University of Nebraska, Lincoln, NE 68583, USA

Correspondence
Xue-Feng Ma, Noble Research Institute, Ardmore, OK 73401, USA.
Email: xma@noble.org

Funding information
Foundation for Food and Agriculture Research, Grant/Award Number: 0001; Noble Research Institute, LLC

Abstract

Classical plant breeding has been instrumental in changing the genetic makeup of crop plants for better ecological adaptation and improved quality. This paper provides insights of the genomic changes effected in hard winter wheat (*Triticum aestivum* L.) through decades of breeding and selection in the Great Plains of the United States. Population structure and differentiation analyses were conducted on 185 wheat cultivars released from 1943 to 2013. Cultivars were grouped into four distinct clusters using discriminant analysis of principal components (DAPC). One of the clusters was unique in that 15 out of the 18 individuals were recent releases (2000–2010), while 12 of the 18 shared the cultivar 'Jagger' in their genetic background. Jagger carries a 2NS/2AS translocation segment from *Aegilops ventricosa*, an important segment for resistance to several foliar diseases. Using the outlier approach, Wright's population fixation index (Fst) identified 450 loci that were directionally selected. The largest signature of selection was found on chromosome 2A. Genetic diversity was high while the inbreeding coefficient was low, indicating extensive hybridization and germplasm exchange among breeding programs within the region. Foliar disease pressure and selection for resistance helped shape the microevolution of wheat in the southern Great Plains. The results showed that high genetic diversity remains in hard winter wheat cultivars adapted to the Great Plains of the USA, and modern plant breeding did not cause any sizable reduction in genetic diversity of the crop in this region.

1 | INTRODUCTION

Common wheat (*Triticum aestivum*) is an allohexaploid (AABBDD) species with a basic chromosome number of seven (2n = 6x = 42). Because of its cumbersome genome size (~16 Gb) and high genome redundancy, genomic analysis of this species has remained one of the most challenging targets in plant genomics

Abbreviations: DAPC, discriminant analysis of principal components; MAF, minor allele frequency; PCA, principal component analysis; SNP, single nucleotide polymorphism; TCAP, Triticeae Coordinated Agricultural Project.

(Chapman et al., 2015; IWGSC, 2018; Zimin et al., 2017). On the other hand, traditional wheat breeding has made breakthroughs with remarkable increases in yield and quality (Fu, 2015; Pingali, 2012; Sharma et al., 2012). Breeding and selection endeavors of the past have developed hundreds of improved wheat cultivars that have revolutionized food and feed market worldwide. The occurrence of the Green Revolution between 1960 and 1980 represents one of the most significant contributions of plant breeding in human history (Borlaug, 1983; Gollin, Hansen, & Wingender, 2018).

Plant breeding tools help create genetic variation, whereas selection reduces the frequency of undesirable or rare alleles in the population and thus may cause genetic bottlenecks (Hyten et al., 2006; Voss-Fels et al., 2015). Genetic diversity has been under constant change or reshuffling because of artificial selection and environmental adaptation. While reduction of genetic diversity from wild forms to landraces and from landraces to modern cultivars is apparent (Gao, Zhao, Huang, & Jia, 2017; Reif et al., 2005), previous reports do not agree on the impact of modern plant breeding on genetic diversity of crop plants. Among others, Fu et al. (2005) and Fu and Somers (2009) reported reductions in genetic diversity in Canadian bread wheat, while Reif et al. (2005), Warburton et al. (2006), and White et al. (2008) reported constant or increasing diversity in wheat in other parts of the world. Notably, much of the success in CIMMYT during the Green Revolution was achieved at the cost of an overall reduction in wheat genetic diversity; however, this trend has been reversed through the more recent introgression of novel materials (Reif et al., 2005; Warburton et al., 2006). Generally, phenotypic improvement in modern crop cultivars is caused by a relatively small number of loci while a large proportion of the genome remain unchanged (Hyten et al., 2006). Those areas of the genome that underwent changes due to selection are referred to as signatures of selection (Cutter & Payseur, 2013).

Identifying those signatures of selection enables better understanding of the genetic makeup of the improved plant, leading to efficient crop breeding and genetic resource-based conservation programs. Advances in genomic technologies have made it possible to identify candidate genomic regions underlying local adaptation (Pool, Hellmann, Jensen, & Nielsen, 2010). Among such approaches are genome scan or outlier detection methods which have become particularly popular in identifying population structuring and adaptive differentiation (De Mita et al., 2013). Genomic differentiation or population fixation indices were helpful to identify and map functional mutations that facilitate marker-assisted selection (Leinonen, McCairns, O'hara, & Merilä, 2013; Maccaferri et al., 2019; N'Diaye et al., 2018).

Core Ideas

- Genomic changes may be tracked among historical wheat cultivars in the Great Plains of the United States.
- Population structure analysis revealed four distinct clusters of cultivars in the region.
- Genetic differentiation analysis among subpopulations revealed variation in selection signatures.
- Wright's population fixation index (Fst) identified directionally selected genetic loci.
- Selection signatures resulted mainly from selection for diseases resistance in the region.

Bread wheat in the United States was first introduced to the eastern part of the country in the early 17th century by the earliest colonialists and it expanded to the Midwest (often referred to as the "Wheat Belt") by the late 1800s with the introduction of 'Turkey' hard red winter wheat (Ball, 1930; Carleton, 1900).

Breeding in these areas has produced hundreds of improved cultivars. In this study, we analyzed population genetic parameters of 185 commercially released wheat cultivars across the Great Plains of the United States (Supplemental Table S1). These cultivars were a subset of the hard winter wheat association mapping panel (HWWAMP) previously genotyped using the 90K single nucleotide polymorphism (SNP) infimum assay (Wang et al., 2014) under the Triticeae Coordinated Agricultural Project (TCAP). Linkage disequilibrium and association mapping studies on some agronomic traits were published based on this mapping panel (Ayana et al., 2018; Guttieri et al., 2015a, 2017; Maulana et al., 2018). The SNP data could be further analyzed to identify loci that were directionally selected and to identify selection signatures in relation to the crop traits targeted for improvement. Therefore, the objective of this study was to retrospectively identify genome regions that were directionally selected, and link these regions with characterized genes known to affect wheat adaptation.

2 | MATERIALS AND METHODS

2.1 | Plant material and SNP genotyping

The HWWAMP consisted of 299 wheat lines that included released cultivars and elite experimental lines. We excluded experimental lines and only used 185 commercial

cultivars that have available records of year of release and state/program of origin (Guttieri et al., 2015a, 2015b). These cultivars were released between 1943 and 2013; their pedigrees and origins are presented in Supplemental Table S1.

The SNP data were accessed from the TCAP database (TCAP90K_HWWAMP, https://tcap.pw.usda.gov/wheat/genotyping/) (Blake et al., 2016). After filtering out SNPs with minor allele frequency (MAF) of less than 5% and missing data of more than 10%, 16,054 SNPs (out of 21,555 total SNPs) were retained and used in this study. Chromosomal positions were based on the 90K SNP consensus map by Wang et al. (2014) because many SNPs were not able to be mapped to the newly released wheat reference sequence (IWGSC, 2018).

2.2 | Population structure

All computational analyses were performed using R statistical software (R Core Team, 2018). Population structure of the 185 commercially released wheat cultivars was investigated using discriminant analysis of principal components (DAPC) with the *adegenet*/R package (Jombart & Ahmed, 2011; Jombart, Devillard, & Balloux, 2010). The optimum number of principal components (PCs) to retain for DAPC was interactively determined by using the *find.clusters* function in the *adegent*/R package with 10,000 iterations (Jombart & Ahmed, 2011). K-means clustering of PCs was used to identify groups of lines. The optimal number of k-means was determined by using the Bayesian information criterion (BIC) as a statistical measure of goodness of fit. Relationships among the 185 genotypes were further visualized using DAPC.

2.3 | Identification of selection signatures and genetic differentiation

Population genomic analyses were carried out using the *hierFstat* package in R (Goudet & Jombart, 2015). Population genomics parameters, including observed heterozygosity within subpopulations (Ho), expected heterozygosity within subpopulations (Hs), heterozygosity within the total population (Ht), Wright's fixation index (Fst), and inbreeding coefficient (Fis), were analyzed in terms of the total population (T), subpopulations (S), and individuals (I). Wright's Fis is the inbreeding coefficient of an individual with respect to the subpopulation it belongs to and Fst is the average inbreeding coefficient of subpopulations relative to the total population. Therefore, Fst was used as a measure of genetic differentiation *among* subpopulations (clusters) while Fis was used to estimate genetic differentiation *within* subpopulations (Hudson, Slatkin, & Maddison, 1992). All of the population parameters were assessed following Nei (1978).

Wright's fixation index (Fst) was calculated as: Fst = (Ht − Hs)/Ht. Similarly, inbreeding coefficient (Fis) was calculated as: Fis = (Hs − Ho)/Hs, where Hs is the mean expected heterozygosity within subpopulations, Ht is the expected heterozygosity within the total population, and Ho is the observed heterozygosity within subpopulations.

To identify genetic loci subjected to selection, we used the high-Fst outlier method corresponding to the distribution of Fst. From the Fst values, we determined the first and third quartiles (Q1 and Q3). Values higher than [Q3 + 1.5*(Q3 − Q1)] are considered positive outliers. Significance of these outliers were tested using Rosner's test (Rosner, 1983) with the EnvStats package in R (Millard, 2013).

Pair-wise genetic distances based on Fst values were calculated between subpopulations using the modified Roger's distance (Wright, 1978). Fst values were plotted against their chromosomal positions to show selection signatures along the genome. Allelic frequencies of highly selected loci were plotted across subpopulations to show their relative differences for those loci.

3 | RESULTS

3.1 | Genetic diversity and population structure

Preliminary variable transformation using principal component analysis (PCA), prior to DAPC analysis, showed that the first 100 principal components explained close to 90% of the total genetic variation (Figure 1a inset, PCA eigenvalues). The Bayesian information criterion (BIC) was lowest at $k = 4$ (Figure 1b), indicating a high probability of four distinct clusters (Figure 1a). The first three principal components of discriminant analysis (DA) captured most of the genetic structure of the total population (Figure 1a inset, DA eigenvalues).

Cultivars from the northern Great Plains tended to cluster together and the same was true for cultivars from the southern Great Plains. For example, cluster 1 was dominated by cultivars from Nebraska, Kansas, Colorado, and South Dakota, whereas Kansas, Texas and Oklahoma cultivars dominated cluster 3 and Colorado and Texas dominated cluster 4 (Table 1). Cultivars from Kansas and Oklahoma appeared in all four clusters.

We also analyzed temporal patterns by classifying cultivars in 10-year intervals. Cluster membership based on year of cultivar release did not clearly follow the 10-year intervals. However, most cultivars in cluster 2 (15/18) were

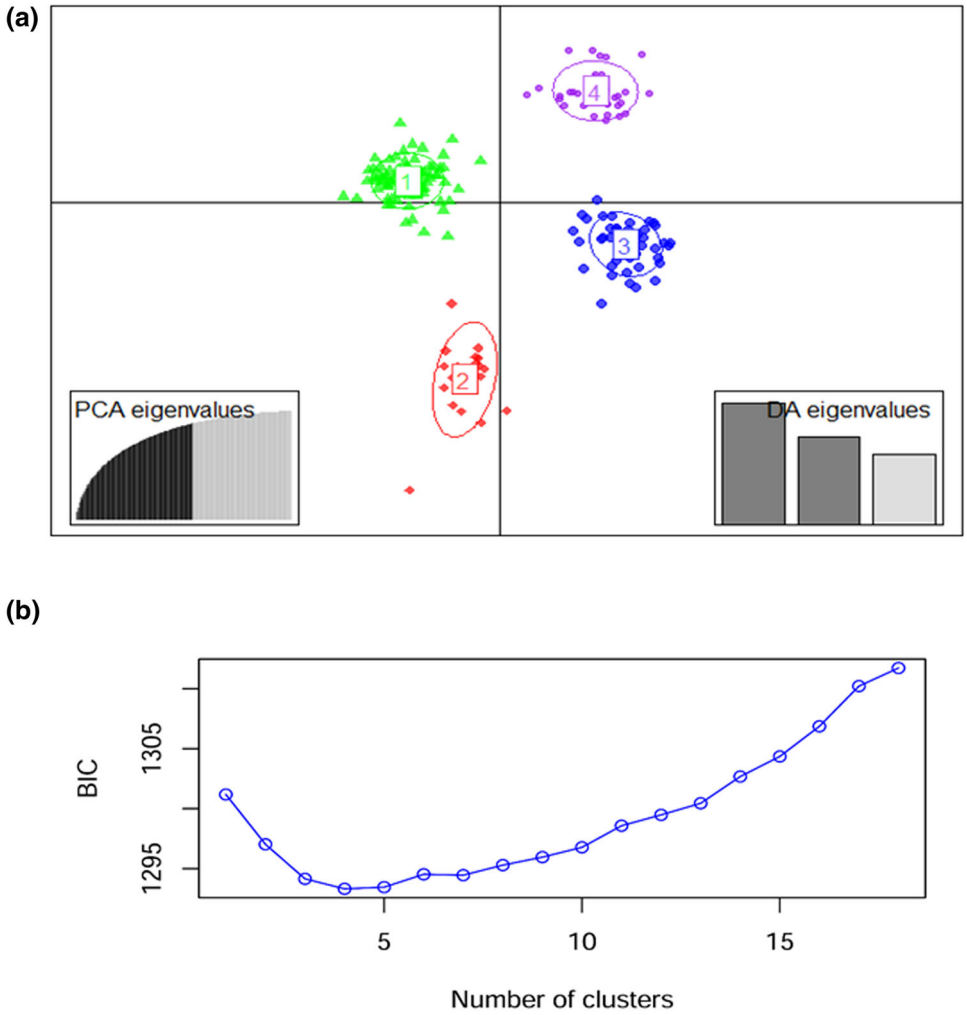

FIGURE 1 Population structure analysis of the 185 wheat cultivars. (a) Scatter plot of discriminant analysis of principal components (DAPC). The main figure shows the relative scatter within the four clusters, in which each dot represents a unique cultivar. The insets indicate the proportion of principal components used for DAPC (bottom left) and number of principal components for the discriminant analysis (DA) (bottom right). (b) The optimum number of clusters was determined using the *find.clusters* function in R. The graph shows the values for the Bayesian information criterion (BIC) relative to the numbers of clusters (k) tested

TABLE 1 Distribution of wheat cultivars across the four clusters

State	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Total
Colorado	15		1	9	25
Kansas	17	14	18	5	54
Montana	8		1	1	10
Nebraska	35		1	3	39
Oklahoma	2	2	14	1	19
South Dakota	12		4		16
Texas		2	9	11	22
Total	89	18	48	30	185

released between 2000 and 2010. Cultivars in this cluster were mostly from Kansas (14/18), while the rest (4/18) were from Texas and Oklahoma (Table 1; Supplemental Table S1). Twelve of the 18 cultivars in this cluster had cultivar Jagger in their pedigrees.

3.2 | Directional selection and population differentiation

The four clusters were used as subpopulations to analyze genetic diversity and subpopulation differentiation. There was high within-subpopulation heterozygosity (0.77), while the inbreeding coefficient (−0.67) and Wright's fixation index (0.02) were low (Figure 2; Supplemental Table S2). Similarity in expected heterozygosity in the total population (0.46) and heterozygosity within

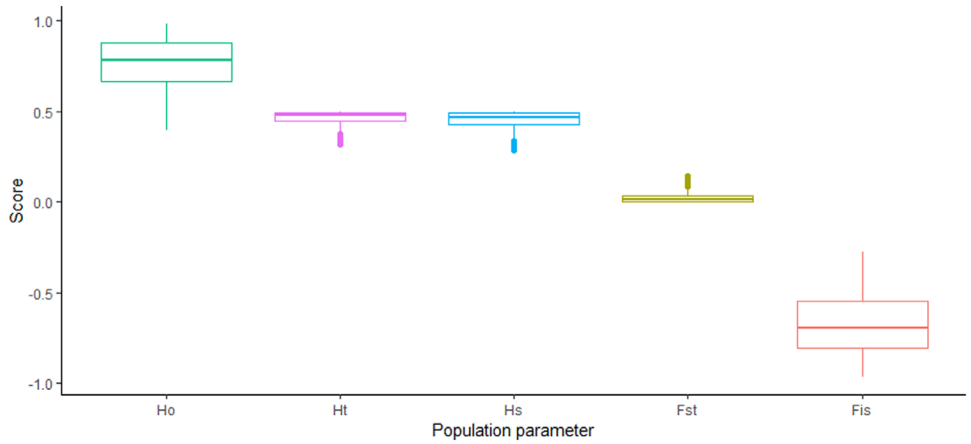

FIGURE 2 Summary statistics (boxplots) of genetic variation in US Great Plains wheat cultivars released between 1943 and 2013 based on the four clusters. Ho, observed heterozygosity within subpopulations; Hs, expected genetic diversity (heterozygosity) within subpopulations; Ht, expected heterozygosity in the random-mating total population; Fst, fixation index; Fis, inbreeding coefficient

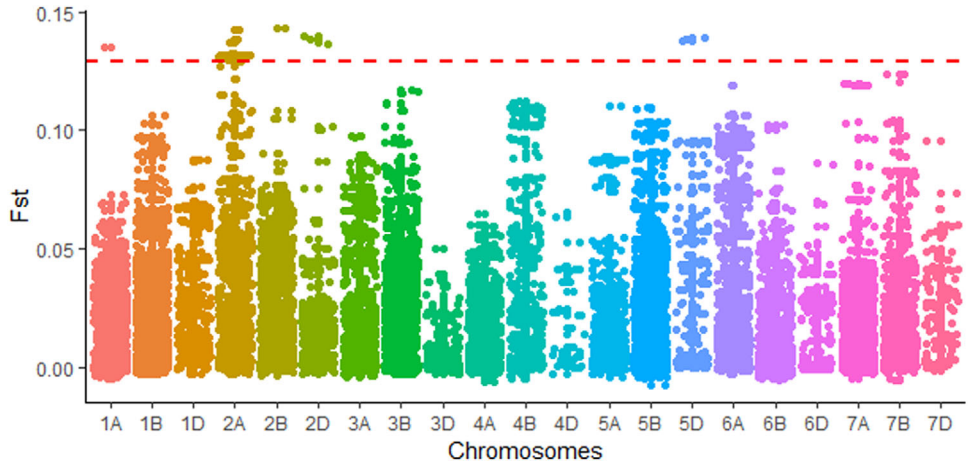

FIGURE 3 Genomic distribution of fixation index (Fst) values as a function of chromosome position in the wheat genome. The red dashed line corresponds to Fst = 0.129, the significant threshold value based on Rosner's outlier test

subpopulation (0.45) reflected low population differentiation. Fst values, using the four clusters as population substructures, ranged from −0.007 to 0.143 (negative Fst values are treated as zero, Supplemental Table S2). Close to 10% (1,438) of the total SNPs were not affected by selection (Fst = 0), whereas 450 loci were found to be outliers (Fst > 0.085). Fifty-one out of the 450 outliers were found to be significant outliers (Fst > 0.129) using Rosner's outlier test. Major selection signatures were found on chromosomes 1A, 2A, 2B, 2D, and 5D (Figure 3). Thus, group 2 showed significant signals of directional selection across all three homoeologous chromosomes 2A, 2B, and 2D.

Average pairwise Fst among clusters showed moderate population differentiation that ranged from 0.014 between clusters 1 and 3 to 0.043 between clusters 2 and 4 (Table 2). The seven states of cultivar origin were also used as population substructures and the Fst test was conducted (Supplemental Table S3). Average Fst was relatively lower in this case (0.01) than using the four clusters (0.02) identified. Based on pairwise Fst values between the states of germplasm source, Montana and Texas (0.021) had the largest differentiation, followed by Montana and Colorado (0.019), and Montana and Oklahoma (0.017), compared to the rest of the states because Montana is most distantly located relative to other states (Table 3). It is not surprising that the smallest differentiation values were observed

TABLE 2 Average pairwise Fst-based genetic distance among the four clusters

	Cluster 1	Cluster 2	Cluster 3
Cluster 2	0.040		
Cluster 3	0.014	0.036	
Cluster 4	0.021	0.043	0.020

TABLE 3 Average pairwise Fst-based genetic distance among the seven states of cultivar origin

	Kansas	Oklahoma	Colorado	Nebraska	S. Dakota	Montana
Oklahoma	0.002					
Colorado	0.009	0.012				
Nebraska	0.008	0.012	0.007			
South Dakota	0.006	0.009	0.009	0.001		
Montana	0.014	0.017	0.019	0.012	0.010	
Texas	0.006	0.002	0.010	0.015	0.014	0.021

between neighboring states, such as Nebraska and South Dakota (0.001), and Texas and Oklahoma (0.002), because of more similar agro-ecologies and frequent germplasm exchange between neighboring states. Similar analysis was conducted using the 10-year intervals during which cultivars were released as subpopulation structures (Supplemental Table S4). Despite the fact that these cultivars were released over a span of seven decades, they exhibited very low temporal genetic differentiation with an average Fst value of 0.001.

3.3 | Allele frequency differences among clusters

Directional selection leads to an increase in the frequency of favored alleles over time. BobWhite_c25359_132, GENE-1273_59, Excalibur_c15379_1305, Kukri_c28182_129, BS00 093111_51, RAC875_c90426_151, BobWhite_c7604_181, Kukri_c31776_1621, and Excalibur_c25599_358 were the top representative loci that showed significant directional selection (Fst > 0.137) (Supplemental Table S2). As shown in Figure 4, the alternative alleles of the indicated loci were highly concentrated in cluster 2 relative to the other three clusters. There was a complete shift (allele fixation) towards alternative alleles (alleles "a" to "g" shift) of loci BobWhite_c25359_132, BobWhite_c7604_181, and Kukri_c28182_129 in cluster 2 (Figure 4). Allele frequencies of those directionally selected loci indicated that cluster 2, which mainly consisted of cultivars released in Kansas between 2000 and 2010, represents a very recent selection signature in wheat breeding. Traits associated with these selection signatures deserve further investigation.

4 | DISCUSSION

Wheat is a recent introduction to the United States, which, as a result, might have experienced genetic bottlenecks early in the breeding process due to a founder effect associated with a restricted ancestral base – predominately Turkey and numerous selections from it. The 185 bread wheat cultivars in this report were grouped into four distinct clusters using discriminant analysis of principal components (DAPC). There was high genetic diversity within and among subpopulations; however, population differentiation as a whole was not high (Figure 2). In a similar study using Diversity Array Technology (DArT) markers, White et al. (2008) reported higher genetic diversity in wheat cultivars from the United States compared with that of Australia and the United Kingdom. Reif et al. (2005) and Warburton et al. (2006) reported that breeders were able to avert constriction of genetic diversity and subsequently enriched genetic diversity through the introgression of novel materials. The seemingly small genetic erosion and very low inbreeding coefficient in this study may be partly explained by extensive intercrossing of parental lines coupled with constant exchange of breeding materials between breeding programs within the Great Plains and introduction of novel germplasm beyond the Great Plains. Artificial selection, unlike introduction bottlenecks, showed very low impact on the genetic diversity of soybeans in the United States (Hyten et al., 2006).

Evidence for directional selection was investigated using Wright's fixation indices estimated based on subpopulations defined in different ways. Wheat cultivars across the Great Plains of the United States showed low to moderate population differentiation despite decades of selective breeding. Genetic differentiation based on ecological adaptation and generation classes (decade of cultivar release) was not significant. The extensive intercrossing between parental lines and germplasm exchange between breeding programs might have eliminated both temporal and spatial genetic differentiation in the cultivars studied. Differentiation among clusters was more pronounced. There was significant allelic shift of some loci (significant outliers) between cultivars in cluster 2 and the other three clusters (Figure 4). Cluster 2 was dominated by cultivars that share inheritance from the cultivar Jagger, a broadly adapted and highly disease resistant cultivar when it was released. Population differentiation in this study tended to identify the unique pedigree background that was resistant to multiple

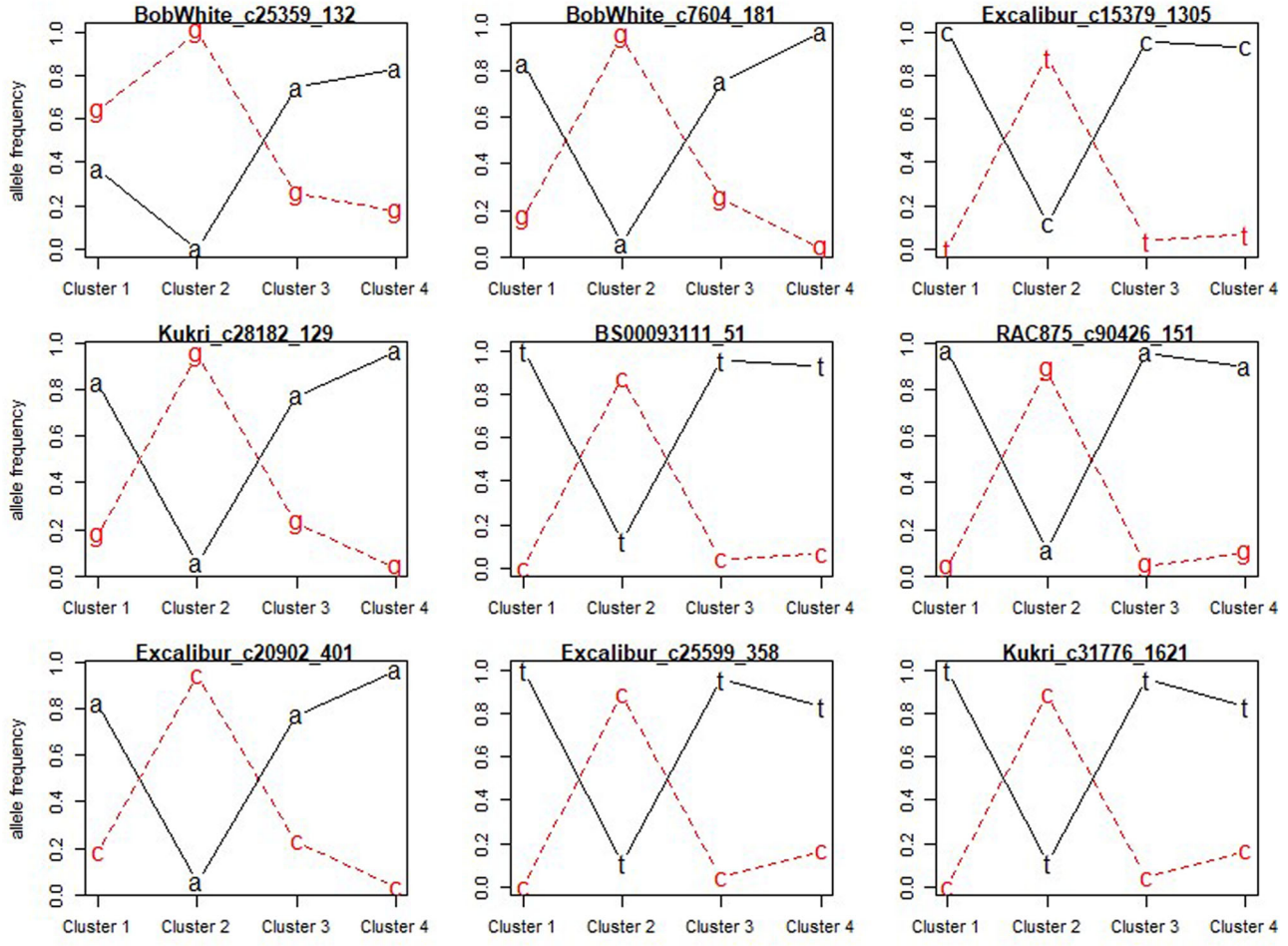

FIGURE 4 Allele frequency changes of nine representative significant single nucleotide polymorphism (SNP) loci across the four clusters. Cluster 2 shows the inverse frequency of alleles relative to the other clusters. There was a complete shift (allele fixation) towards alternative alleles of loci BobWhite_c25359_132, BobWhite_c7604_181, and Kukri_c28182_129 in cluster 2

foliar diseases due to the 2NS/2AS translocation segment discussed below.

Disease pressure and selection for resistance appeared to shape the microevolution of wheat in the southern Great Plains. Cultivars from Kansas, Oklahoma and Texas showed the least genetic distance, as might be expected because public breeding programs from those states share their germplasm freely and frequently. Commonality of disease pressures tends to lead to commonality of gene sources for resistance, intended or unintended. Nevertheless, wheat breeding in this region should continue to incorporate germplasm from diverse sources to enrich the genetic pool for future breeding progress. It was interesting to observe that cluster 2 was composed of newer cultivars, mostly from Kansas (14/18), while the remaining four were from Texas and Oklahoma. The uniqueness of cluster 2 in this study is attributable to cultivar Jagger and its relatives. Jagger was released in 1994 to the Great Plains of the United States for its high grain yield potential, strong general disease resistance, and excellent bread-baking quality (Sears et al., 1997). This cultivar and its relatives provided reliable sources of resistance to stripe rust (caused by *Puccinia striiformis f. sp. Tritici*) that remained effective in the southern Great Plains until 2010 (Fang et al., 2011; Xue, Kolmer, Wang, & Yan, 2018).

Jagger carries a 2NS/2AS translocation segment from *Aegilops ventricosa* on the short arm of chromosome 2A, spanning a distance of 27.8 Mb from the telomere (Xue et al., 2018). This translocation segment carries disease resistance genes for leaf rust (*Lr37*) caused by *Puccinia triticina* Erikss (Kolmer, 2017; Xue et al., 2018), stem rust (*Sr7a* and *Sr38*) caused by *Puccinia graminis* (Turner, Jin, Rouse, & Anderson, 2016), wheat blast caused by *Magnaporthe oryzae* (Cruz et al., 2016), and stripe rust (*Yr17*) (Fang et al., 2011) although some of the resistance was lost recently due to occurrence of new virulence. The top three directionally selected loci identified in the present study (Excalibur_c15379_1305, Kukri_c31776_1621,

and Excalibur_c25599_358) were mapped on the short arm of 2A between 6.04–25.9 Mb, further validating that the 2NS alien fragment or disease resistance was the main driver of differentiation between cultivars in cluster 2 and the other clusters. A recent study by Juliana et al. (2019) identified several QTL on chromosomes 2A and 2D which are associated with resistance to four stem rust races. Chromosome arms of 1AS and 5DL also harbor several powdery mildew resistance genes (Ma, Sorrells, & Tanksley, 1994). Disease pressure remains a significant production constraint, ever since wheat was introduced to the United States (Ball, 1930). To no surprise, the most important selection signatures in the present study were associated with disease resistance, the result of intensive and long-term research on improving resistance to a pest often considered a moving target.

The five major selection signatures (on chromosomes 1A, 2A, 2B, 2D, and 5D) identified in this study consisted of the genomic fingerprints of other highly selected traits besides disease resistance. Five dwarfing genes were previously identified and mapped to chromosomes 2AS (Chaudhry, 1973; Yang, Zhang, Liu, & Wang, 1993), 2BL (Ellis, Rebetzke, Azanza, Richards, & Spielmeyer, 2005), 2DS (Worland, Korzun, Röder, Ganal, & Law, 1998), and 5DL (Chen et al., 2015). Chromosome 2DS also harbors several grain yield QTL (Nielsen, Backes, Stougaard, Andersen, & Jahoor, 2014), which are linked to the *Rht8*, and the daylength-insensitive *Ppd-D1* gene. *Ppd-D1* is closely linked with *Rht8* (Würschum, Langer, Longin, Tucker, & Leiser, 2017) but *Ppd-D1* has also been reported to have pleiotropic effects on plant height (Börner, Worland, Plaschke, Schumann, & Law, 1993). Here again, the top four significant outlier loci we identified (GENE-1273_59, BS00093111_51, Excalibur_c33173_557, and RAC875_c90426_151) co-located with the *Rht8* and *Ppd-D1* genes on the short arm of chromosome 2D. Short stature wheat was one of the main factors for the success of the Green Revolution in the 1960s (Borlaug, 1983).

In conclusion, wheat breeding in the Great Plains of the United States did not show any major reduction in genetic diversity. Many of the selection signatures were associated with the introgression of a chromosome segment from *Ae. ventricosa* for disease resistance. Wheat cultivars from the southern Great Plains are more related to each other than wheat from other states in this study. As a result, in addition to the disease resistance genes conferred by Jagger and its offspring, the genetic pool of wheat in the southern Great Plains stands to be further enriched through nonconvergent introduction of genetic diversity.

ACKNOWLEDGMENTS

The authors sincerely thank Meg Drake for critical reading of the manuscript and the Triticeae Coordinated Agricultural Project (TCAP), funded by USDA Agriculture and Food Research Initiative Competitive Grant 2011-68002-30029, for making genotypic data used in this study publicly available. This project was supported by Noble Research Institute, LLC and Foundation for Food and Agriculture Research (FFAR Award ID: 0001).

ORCID

Habtamu Ayalew https://orcid.org/0000-0002-4778-9008
Xue-Feng Ma https://orcid.org/0000-0002-0942-9116

REFERENCES

Ayana, G. T., Ali, S., Sidhu, J. S., Gonzalez-Hernandez, J. L., Turnipseed, B., & Sehgal, S. K. (2018). Genome-wide association study for spot blotch resistance in hard winter wheat. *Frontiers in Plant Science*, *9*, 926. https://doi.org/10.3389/fpls.2018.00926.

Ball, C. R. (1930). The history of American wheat improvement. *Agricultural History*, *4*, 48–71.

Blake, V. C., Birkett, C., Matthews, D. E., Hane, D. L., Bradbury, P., & Jannink, J.-L. (2016). The Triticeae Toolbox: Combining phenotype and genotype data to advance small-grains breeding. *The Plant Genome*, *9*, 1–10. https://doi.org/10.3835/plantgenome2014.12.0099.

Borlaug, N. E. (1983). Contributions of conventional plant breeding to food production. *Science*, *219*, 689–693. https://doi.org/10.1126/science.219.4585.689.

Börner, A., Worland, A. J., Plaschke, J., Schumann, E., & Law, C. N. (1993). Pleiotropic effects of genes for reduced height (Rht) and day-length insensitivity (Ppd) on yield and its components for wheat grown in middle Europe. *Plant Breeding*, *111*, 204–216. https://doi.org/10.1111/j.1439-0523.1993.tb00631.x.

Carleton, M. A. (1900). *Russian cereals adapted for cultivation in the United States*. USDA Division of Botany Bulletin 23. Washington, DC: USDA.

Chapman, J. A., Mascher, M., Buluç, A., Barry, K., Georganas, E., Session, A., … Oliker, L. (2015). A whole-genome shotgun approach for assembling and anchoring the hexaploid bread wheat genome. *Genome Biology*, *16*, 26.

Chaudhry, A. S. (1973). *A genetic and cytogenetic study of height in wheat*. University of Cambridge.

Chen, S., Gao, R., Wang, H., Wen, M., Xiao, J., Bian, N., … Bie, T. (2015). Characterization of a novel reduced height gene (Rht23) regulating panicle morphology and plant architecture in bread wheat. *Euphytica*, *203*, 583–594. https://doi.org/10.1007/s10681-014-1275-1.

Cruz, C. D., Peterson, G. L., Bockus, W. W., Kankanala, P., Dubcovsky, J., Jordan, K. W., … Valent, B. (2016). The 2NS translocation from *Aegilops ventricosa* confers resistance to the *Triticum* pathotype of *Magnaporthe oryzae*. *Crop Science*, *56*, 990–1000. https://doi.org/10.2135/cropsci2015.07.0410.

Cutter, A. D., & Payseur, B. A. (2013). Genomic signatures of selection at linked sites: Unifying the disparity among species. *Nature Reviews Genetics*, *14*, 262. https://doi.org/10.1038/nrg3425.

De Mita, S., Thuillet, A. C., Gay, L., Ahmadi, N., Manel, S., Ronfort, J., & Vigouroux, Y. (2013). Detecting selection along environmental gradients: Analysis of eight methods and their effectiveness for

outbreeding and selfing populations. *Molecular Ecology*, *22*, 1383–1399.

Ellis, M. H., Rebetzke, G. J., Azanza, F., Richards, R. A., & Spielmeyer, W. (2005). Molecular mapping of gibberellin-responsive dwarfing genes in bread wheat. *Theoretical and Applied Genetics*, *111*, 423–430. https://doi.org/10.1007/s00122-005-2008-6.

Fang, T., Campbell, K. G., Liu, Z., Chen, X., Wan, A., Li, S., … Yan, L. (2011). Stripe rust resistance in the wheat cultivar Jagger is due to Yr17 and a novel resistance gene. *Crop Science*, *51*, 2455–2465. https://doi.org/10.2135/cropsci2011.03.0161.

Fu, Y.-B. (2015). Understanding crop genetic diversity under modern plant breeding. *Theoretical and Applied Genetics*, *128*, 2131–2142. https://doi.org/10.1007/s00122-015-2585-y.

Fu, Y.-B., Peterson, G. W., Richards, K. W., Somers, D., DePauw, R. M., & Clarke, J. M. (2005). Allelic reduction and genetic shift in the Canadian hard red spring wheat germplasm released from 1845 to 2004. *Theoretical and Applied Genetics*, *110*, 1505–1516. https://doi.org/10.1007/s00122-005-1988-6.

Fu, Y.-B., & Somers, D. J. (2009). Genome-wide reduction of genetic diversity in wheat breeding. *Crop Science*, *49*, 161–168. https://doi.org/10.2135/cropsci2008.03.0125.

Gao, L., Zhao, G., Huang, D., & Jia, J. (2017). Candidate loci involved in domestication and improvement detected by a published 90K wheat SNP array. *Scientific Reports*, *7*, 44530. https://doi.org/10.1038/srep44530.

Gollin, D., Hansen, C. W., & Wingender, A. (2018). Two blades of grass: The impact of the green revolution. *National Bureau of Economic Research*, *2018*.

Goudet, J., & Jombart, T. (2015). Hierfstat: Estimation and tests of hierarchical F-statistics. R package version 0.04-22. Retrieved from http://www.r-project.org.

Guttieri, M. J., Baenziger, P. S., Frels, K., Carver, B., Arnall, B., Wang, S., … Waters, B. M. (2015a). Prospects for selecting wheat with increased zinc and decreased cadmium concentration in grain. *Crop Science*, *55*, 1712–1728. https://doi.org/10.2135/cropsci2014.08.0559.

Guttieri, M. J., Baenziger, P. S., Frels, K., Carver, B., Arnall, B., & Waters, B. M. (2015b). Variation for grain mineral concentration in a diversity panel of current and historical Great Plains hard winter wheat germplasm. *Crop Science*, *55*, 1035–1052. https://doi.org/10.2135/cropsci2014.07.0506.

Guttieri, M. J., Frels, K., Regassa, T., Waters, B. M., & Baenziger, P. S. (2017). Variation for nitrogen use efficiency traits in current and historical Great Plains hard winter wheat. *Euphytica*, *213*, 87. https://doi.org/10.1007/s10681-017-1869-5.

Hudson, R. R., Slatkin, M., & Maddison, W. P. (1992). Estimation of levels of gene flow from DNA sequence data. *Genetics*, *132*, 583–589.

Hyten, D. L., Song, Q., Zhu, Y., Choi, I.-Y., Nelson, R. L., Costa, J. M., … Cregan, P. B. (2006). Impacts of genetic bottlenecks on soybean genome diversity. *Proceedings of the National Academy of Sciences*, *103*, 16666–16671. https://doi.org/10.1073/pnas.0604379103.

IWGSC (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science*, *361*, eaar7191. https://doi.org/10.1126/science.aar7191.

Jombart, T., & Ahmed, I. (2011). adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics*, *27*, 3070–3071. https://doi.org/10.1093/bioinformatics/btr521.

Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC genetics*, *11*, 94. https://doi.org/10.1186/1471-2156-11-94.

Juliana, P., Poland, J., Huerta-Espino, J., Shrestha, S., Crossa, J., Crespo-Herrera, L., … Kumar, U. (2019). Improving grain yield, stress resilience and quality of bread wheat using large-scale genomics. *Nature Genetics*, *51*, 1530–1539.

Kolmer, J. A. (2017). Genetics of leaf rust resistance in the hard red winter wheat cultivars Santa Fe and Duster. *Crop Science*, *57*, 2500–2505. https://doi.org/10.2135/cropsci2016.12.1029.

Leinonen, T., McCairns, R. J. S., O'hara, R. B., & Merilä, J. (2013). Q ST–F ST comparisons: Evolutionary and ecological insights from genomic heterogeneity. *Nature Reviews Genetics*, *14*, 179. https://doi.org/10.1038/nrg3395.

Ma, Z., Sorrells, M., & Tanksley, S. (1994). RFLP markers linked to powdery mildew resistance genes Pm1, Pm2, Pm3, and Pm4 in wheat. *Genome*, *37*, 871–875.

Maccaferri, M., Harris, N. S., Twardziok, S. O., Pasam, R. K., Gundlach, H., Spannagl, M., … Milner, S. G. (2019). Durum wheat genome highlights past domestication signatures and future improvement targets. *Nature Genetics*, *1*. https://doi.org/10.1038/s41588-019-0381-3.

Maulana, F., Ayalew, H., Anderson, J. D., Kumssa, T. T., Huang, W., & Ma, X.-F. (2018). Genome-wide association mapping of seedling heat tolerance in winter wheat. *Frontiers in Plant Science*, *9*, 1272. https://doi.org/10.3389/fpls.2018.01272.

Millard, S. P. (2013). *EnvStats: An R package for environmental statistics*. New York: Springer.

N'Diaye, A., Haile, J. K., Nilsen, K., Walkowiak, S., Ruan, Y., Singh, A. K., … Pozniak, C. J. (2018). Haplotype loci under selection in Canadian durum wheat germplasm over 60 years of breeding: Association with grain yield, quality traits, protein loss and plant height. *Frontiers in Plant Science*, *9*, 1589. https://doi.org/10.3389/fpls.2018.01589.

Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, *89*, 583–590.

Nielsen, N. H., Backes, G., Stougaard, J., Andersen, S. U., & Jahoor, A. (2014). Genetic diversity and population structure analysis of European hexaploid bread wheat (*Triticum aestivum* L.) varieties. *PLoS One*, *9*, e94000. https://doi.org/10.1371/journal.pone.0094000.

Pingali, P. L. (2012). Green revolution: Impacts, limits, and the path ahead. *Proceedings of the National Academy of Sciences*, *109*, 12302–12308. https://doi.org/10.1073/pnas.0912953109.

Pool, J. E., Hellmann, I., Jensen, J. D., & Nielsen, R. (2010). Population genetic inference from genomic sequence variation. *Genome Research*, *20*, 291–300. https://doi.org/10.1101/gr.079509.108.

R Core Team (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Reif, J. C., Zhang, P., Dreisigacker, S., Warburton, M. L., van Ginkel, M., Hoisington, D., … Melchinger, A. E. (2005). Wheat genetic diversity trends during domestication and breeding. *Theoretical and Applied Genetics*, *110*, 859–864. https://doi.org/10.1007/s00122-004-1881-8.

Rosner, B. (1983). Percentage points for a generalized ESD many-outlier procedure. *Technometrics*, *25*, 165–172. https://doi.org/10.2307/1268549.

Sears, R. G., Moffatt, J. M., Martin, T. J., Cox, T. S., Bequette, R. K., Curran, S. P., … Witt, M. D. (1997). Registration of 'Jagger' wheat. *Crop Science*, *37*, 1010–1010. https://doi.org/10.2135/cropsci1997.0011183X003700030062x.

Sharma, R. C., Crossa, J., Velu, G., Huerta-Espino, J., Vargas, M., Payne, T. S., & Singh, R. P. (2012). Genetic gains for grain yield in CIMMYT spring bread wheat across international environments. *Crop Science*, *52*, 1522–1533. https://doi.org/10.2135/cropsci2011.12.0634.

Turner, M. K., Jin, Y., Rouse, M. N., & Anderson, J. A. (2016). Stem rust resistance in 'Jagger' winter wheat. *Crop Science*, *56*, 1719–1725. https://doi.org/10.2135/cropsci2015.11.0683.

Voss-Fels, K., Frisch, M., Qian, L., Kontowski, S., Friedt, W., Gottwald, S., & Snowdon, R. J. (2015). Subgenomic diversity patterns caused by directional selection in bread wheat gene pools. *The Plant Genome*, *8*, 1–13. https://doi.org/10.3835/plantgenome2015.03.0013.

Wang, S., Wong, D., Forrest, K., Allen, A., Chao, S., Huang, B. E., … Cattivelli, L. (2014). Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. *Plant Biotechnology Journal*, *12*, 787–796. https://doi.org/10.1111/pbi.12183.

Warburton, M. L., Crossa, J., Franco, J., Kazi, M., Trethowan, R., Rajaram, S., … Ginkel, M.v (2006). Bringing wild relatives back into the family: Recovering genetic diversity in CIMMYT improved wheat germplasm. *Euphytica*, *149*, 289–301. https://doi.org/10.1007/s10681-005-9077-0.

White, J., Law, J. R., MacKay, I., Chalmers, K. J., Smith, J. S. C., Kilian, A., & Powell, W. (2008). The genetic diversity of UK, US and Australian cultivars of *Triticum aestivum* measured by DArT markers and considered by genome. *Theoretical and Applied Genetics*, *116*, 439–453. https://doi.org/10.1007/s00122-007-0681-3.

Worland, A. J., Korzun, V., Röder, M. S., Ganal, M. W., & Law, C. N. (1998). Genetic analysis of the dwarfing gene Rht8 in wheat. Part II. The distribution and adaptive significance of allelic variants at the Rht8 locus of wheat as revealed by microsatellite screening. *Theoretical and Applied Genetics*, *96*, 1110–1120. https://doi.org/10.1007/s001220050846.

Wright, S. (1978). *Evolution and the genetics of populations, Volume 4: Variability within and among natural populations*. Chicago: University of Chicago Press.

Würschum, T., Langer, S. M., Longin, C. F. H., Tucker, M. R., & Leiser, W. L. (2017). A modern green revolution gene for reduced height in wheat. *The Plant Journal*, *92*, 892–903. https://doi.org/10.1111/tpj.13726.

Xue, S., Kolmer, J. A., Wang, S., & Yan, L. (2018). Mapping of leaf rust resistance genes and molecular characterization of the 2NS/2AS translocation in the wheat cultivar Jagger. *G3: Genes, Genomes, Genetics*, *8*, 2059–2065. https://doi.org/10.1534/g3.118.200058.

Yang, T., Zhang, X., Liu, H., & Wang, Z. (1993). Chromosomal arm location of a dominant dwarfing gene Rht21 in common wheat variety - XN0004. *Journal of Northwest Sci-Tech University of Agriculture and Forestry, 4*.

Zimin, A. V., Puiu, D., Hall, R., Kingan, S., Clavijo, B. J., & Salzberg, S. L. (2017). The first near-complete assembly of the hexaploid bread wheat genome, *Triticum aestivum*. *Gigascience*, *6*, gix097. https://doi.org/10.1093/gigascience/gix097.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Ayalew H, Sorrells ME, Carver BF, Baenziger PS, Ma X-F. Selection signatures across seven decades of hard winter wheat breeding in the Great Plains of the United States. *The Plant Genome*. 2020;e20032. https://doi.org/10.1002/tpg2.20032