



AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
<http://ageconsearch.umn.edu>
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

Staff Paper Series

Scientific Selection: A Century of Increasing Crop Varietal Diversity in U.S. Wheat

by

Yuan Chai, Philip G. Pardey and Kevin A.T. Silverstein

Scientific Selection: A Century of Increasing Crop Varietal Diversity in U.S. Wheat

April 2022

Yuan Chai^{1*}, Philip G. Pardey^{1,2}, Kevin A.T. Silverstein^{1,3}

Affiliations and Acknowledgements

¹GEMS Informatics Center, University of Minnesota, St. Paul, MN 55108, US

²Department of Applied Economics, University of Minnesota, St. Paul, MN 55108, US

³Minnesota Supercomputing Institute, University of Minnesota, St. Paul, MN 55455, US

*Correspondence to: chai026@umn.edu

The authors are thankful for the outstanding research assistance they received from Connie Chan-Kang of the GEMS Informatics Center. This work was undertaken with primary support from the GEMS Informatics Center, and additional support from the International Science and Technology Practice and policy (InSTePP) center and the Minnesota Supercomputing Institute, all at the University of Minnesota. Partial support was also received from the Minnesota Agricultural Experiment Station (MIN-14-171).

Copyright © 2022 by Yuan Chai, Philip G. Pardey and Kevin A.T. Silverstein

Scientific Selection: A Century of Increasing Crop Varietal Diversity in U.S. Wheat

ABSTRACT

A prevalent and persistent biodiversity concern is that modern cropping systems lead to an erosion in crop genetic diversity. Although certain trait uniformity provides advantages in crop management and marketing, farmers are also incentivized to use diverse genetics to reduce risks from change in climate, pests and markets. These risk factors have spurred increased turnover in varietal use to address complex and spatially variable genetics, environment, and crop management (GxExM) interactions to optimize crop performance. Contrary to commonly held perceptions, phylogenetically blind and phylogenetically informed diversity metrics reveal that the intensive use of scientifically selected varieties has led to significant increases in both the spatial and temporal diversity of the U.S. wheat crop over the past century.

1. Introduction

Halting biodiversity loss is crucial to achieving many of the U.N.'s sustainable development goals, and is a leading development target in its own right (U.N. 2021; Delabre et al. 2021). Agriculture is seen as both a key cause of the global “biodiversity crisis” (UNEP 2016; U.N. 2021, pp. 19 and 52; Lu and Bullock 2021), and a principal means of addressing it (U.N.-SCBD 2020, p. 20). The nexus between biodiversity and agriculture is complex and multidimensional. With 36.7% of the world's land mass in cropping and animal agriculture (FAO 2021), promoting sustainable agricultural productivity growth is key to feeding a large and still growing world population at affordable prices, while at the same time stalling growth in or shrinking the footprint of agriculture to return areas to wild or at least more natural, biodiverse landscapes. The small number of crop species used to feed the planet is also a source of concern to many (e.g., National Research Council 1972; Fowler and Mooney 1990; Thrupp 2000; Smolders 2006). In 2018, more than half (51.6%) of the 7,967 quadrillion calories consumed by humans were sourced from just 5 crops—wheat (18.5% of total calories), rice (18%), sugar (6.8%), maize (5.3%), and soybeans (2.9%) (FAO 2021). This translates into large areas of the world growing the same crop type, for example, 216 million hectares (15.0% of total harvested area) of wheat, 197 million hectares (13.8%) of corn, and 162 million hectares (11.3%) of rice in 2019 (FAO 2021).

Another long-standing biodiversity concern is genetic erosion within cropping agriculture (e.g., Miller 1973; Harlan 1975), which Brush (2004, p. 154) succinctly defined as “...the loss of variability in crop populations.” This manifests itself as a narrowing of the genetic diversity across farmers' fields within a given crop species. The transition from growing landraces, essentially farmer-bred crop varieties, to scientifically bred varieties is seen as a pivotal point in

the narrowing of genetic diversity within agriculture (Frankel 1970; Harlan 1975). But the subsequent decades of using varieties developed by scientists rather than farmers is perceived as a further, if not the primary, cause of a narrowing genetic variability in crop populations (Fowler 1994, Appendix I; Brush 2004, ch. 7). Two commonly cited corollaries from these concerns are that a) a narrowing of the genetic variability in fields using scientifically-bred varieties makes cropping agriculture more vulnerable to losses associated with adverse climate, pest and disease shocks (NAS 1972; Miller 1973), and b) the encroachment of modern agriculture into cropping systems more reliant on the use of landraces erodes in situ access to the historical accumulation of genetic diversity that farmers through the ages have formally or informally selected into that material (Harlan 1975).

A persistent and still prevailing perception is that the use of modern, and in particular scientifically bred, varieties continues to be a major driver of genetic erosion (Frankel 1970; Hawkes 1983; UN-SCBD 2020). However, economic first principles suggest that genetic erosion within agricultural crops occasioned by the shift to, and, especially, the subsequent widespread uptake of, scientifically bred varieties is neither inevitable, or necessarily the most probable outcome. While economic incentives are likely to lead to crop uniformity in some traits, similar incentives are also likely to drive diversification in other crop attributes. As Duvick (1984) discussed, there are substantial economic gains by way of lower costs of food production, processing and consumption that come from uniformity in particular phenotypic traits in certain cropping systems (e.g., uniform crop emergence, flowering and harvesting times; plant, seed or fruit size, shape or composition; and so on). However, there are other economically valuable crop traits—e.g., resilience to changes in biotic and abiotic stresses over time, or responsiveness

to locational differences in the agroecological attributes that affect crop production—that incentivize the development and use of more diverse germplasm.

From this economic framing, it naturally follows that an informative assessment of crop genetic diversity considers both the temporal and the spatial dimensions of diversity. For example, while landraces (or farmer-bred varieties) may exhibit more spatial diversity at any point in time (Peng et al. 2011), they may be vulnerable to unprecedented (in nature and magnitude) biotic or abiotic shocks occasioned by relative rapid changes in climate and human- or naturally-mediated invasions of foreign pests and diseases. In contrast, the relatively rapid turnover of crop varieties in more intensive agriculture systems that use scientifically selected material (Brush 2004, ch. 7; Brennan and Byerlee 1991; Meng et al. 1998) opens up prospects for more rapid, and thereby more valuable, genetic responses to changing climate, pest and disease circumstances. Market forces are likely to favor varieties with superior performance attributes in these changed environmental realities. In addition, scientifically selected varieties that optimize their performance in locationally variable agroecological environments means that market forces are also likely to spur more spatially diverse seed development and deployments to achieve better performing G x E (genetics-by-environment) matchups.

The ecology, genetics, economics, information theory and other sciences have spawned a multitude of diversity measures. The measures used by prior studies of wheat diversity (see Table S1 for a summary) fall into two broad categories. One is phylogenetically blind entropy measures (e.g., Brennan and Byerlee 1991; Smale et al. 1998). These measures quantify diversity in terms of varietal “abundance” without consideration of relatedness among varieties. When applied to cropping agriculture, these measures effectively consider each variety of a given crop as being equally distinct from one another. The second category are phylogenetically informed

measures that attempt to account for the genetic, phenotypic, or functional relatedness among crop varieties. These include pedigree-based similarity measures, phenotypic (morphological or biochemical) traits-based measures, and molecular marker-based measures of genetic diversity.

A sizable number of studies have used a range of relatedness measures, including phenotypic traits, varietal pedigrees or molecular markers to generate measure of varietal relatedness among wheat populations (see, e.g., Table S1). Of note, these prior studies often do not take into consideration the spatio-temporal dynamics of wheat varietal diversity in the sense that the diversity is assessed among a collection of varieties irrespective of the locales and time where the varieties were grown (see, e.g., Poets et al. 2020).

Based on the economic notions broached above, it is by no means a foregone conclusion that the more intensive (in space and time) use of scientifically selected crop varieties leads to an erosion of crop diversity. Here we use a purpose-built data set to quantify the changing spatio-temporal pattern of varietal diversity in the U.S. wheat crop over the past century spanning the years 1919 to 2019; a period that encompasses most of the time that scientific crop breeding was informed by Mendelian methods of genetic selection (Biffen 1905; Ball 1930; Smýkal et al. 2016). To do so we draw on phylogenetically informed metrics of diversity developed in the ecological literature while also introducing the notion of a “meta-population” (where, in this instance, U.S. level crop varietal data are parsed into state-level spatial sub-units) to enable the joint assessment of changes in temporal and spatial crop diversity. We used the generalized phylogenetic diversity (PD) measure proposed by Chao et al. (2010), and applied the Hill numbers (or the effective number of varieties) framework to quantify the changing varietal diversity of the U.S. wheat crop during the past 100 years. The generalized PD and Hill numbers framework incorporates both varietal relatedness and varietal abundance to allow for the joint assessment of the spatio-

temporal dynamics of varietal diversity within a cropping system that is heavily reliant on scientifically selected varieties.

2. Methods

2.1 Data on Wheat Varieties in the U.S.

Data on planted area-by-variety were collected for the U.S. wheat crop for the period 1919-2019 by the authors and colleagues at the International Science and Technology Practice and Policy (InSTePP) center and the GEMS Informatics Center, University of Minnesota. Wheat crops were divided into three market classes—durum, spring and winter. From 1919 to 1984, acreage-by-variety data are reported quinquennially for a total of 42 states by USDA’s Statistical Bulletins on the Distribution of the Varieties and Classes of Wheat in the United States. Thereafter we sourced the required data from state-specific agricultural statistical services, which dropped to 16 states in 2000 and the years following. These 16 states accounted for 89% of total U.S. wheat area in 2019. For each market class, the states for which we have area-by-variety data account for around 97 percent of durum acreage, 100 percent of spring acreage, and 84 percent of winter acreage based on the 2019 NASS data. However, even within the 16 major wheat-growing states, not all states report area-by-variety data for all years. To construct a complete state-level panel on wheat variety areas, we filled in the missing years’ variety area information using linear interpolation based on years before and after any missing years. Furthermore, we used each variety’s year of release and last year of reported use nationally to determine the beginning and end year of each variety’s use when interpolating their state-specific areas. Information on variety’s year of release was collected from multiple sources, including the Genetic Resource Information System (GRIS) for Wheat and Triticale (CIMMYT 2017), the Germplasm Resources Information Network (GRIN) (USDA-NGRP 2021), the GrainGenes database

(USDA-ARS 2021), crop registration narratives from scientific journals such as *Crop Science*, *Journal of the American Society of Agronomy*, and *Journal of Plant Registration*, the Plant Variety Protection Office (USDA-PVPO 2021), and searches from elsewhere such as private company websites and university websites.

Additionally, information on each variety's name, market class, and crop pedigree were also collected. Varieties often have aliases or different spellings depending on the time and location they were marketed and adopted. To avoid double counting the same varieties with different names across states and over time, we reconciled the aliases of different varieties and standardized the names across all the reported wheat varieties grown commercially in the U.S. from 1919 to 2019. Multiple sources of wheat genetic and pedigree information were used to consolidate varietal names and their pedigrees, including the Genetic Resource Information System (GRIS) for Wheat and Triticale (CIMMYT 2017), the Germplasm Resources Information Network (GRIN) (USDA-NGRP 2021), the GrainGenes database (USDA-ARS 2021), and the Plant Variety Protection Office (USDA-PVPO 2021).

Where possible, the entire pedigree of each variety and their parents were traced back to either a landrace, a wild accession or a local variety. The pedigree information for all reported commercially grown varieties in the U.S. was collected and processed. Inconsistencies among reported pedigrees (and varietal name) were reconciled by PedTools, a python library developed by the GEMS Informatics Center at the University of Minnesota that maps input varieties for a crop to aliases, standardizes naming protocols, and recursively reconstructs pedigrees in principle back to landraces through repeated observations of parent-child relationships found in the literature. To process the 1,353 commercially grown varieties included in this study,

PedTools was applied to 2,597 varieties (including non-commercially grown crossing materials) for the final pedigree analysis.

2.2 Diversity Measures

A large number of biodiversity measures have been proposed in the disciplines of ecology, genetics, economics, information theory and other sciences (Fisher et al. 1943; Solow et al. 1993; Shannon 2001; Chao et al. 2014). Among them, the most commonly used are phylogenetically blind in the sense that they tally the distributions of each type of entity, regardless of each entity type's taxonomic, genetic or functional similarities. Such indexes include an entity richness index, the Shannon entropy index (Shannon 2001), and the Gini-Simpson index (Simpson 1949), all of which are special cases of the generalized Tsallis entropy measures (Keylock 2005; Jost 2006), also known as HCDT entropy indexes (Havrda and Charvát 1967; Daróczy 1970; Tsallis 1988). HCDT indexes can be converted into so-called “true diversity” measures. Such measures, also known as the effective number of species or “Hill numbers”, represent the hypothetical number of equally abundant entities that would give the same diversity index value as was actually observed (Hill 1973; Jost 2006). “Hill numbers” do not depend on the functional form of the index and satisfy the replication principle, whereby the index value doubles if each entity grouping was divided into two equal new groups (Jost 2006; Tuomisto 2012). Thus, Hill numbers allow for a unified and intuitive interpretation of diversity across locations.

For wheat, new varieties are typically developed by genetic crosses among existing varieties, and thus the contribution of each new variety to the overall crop diversity depends on their relatedness to existing varieties. Commonly used phylogenetically blind biodiversity measures such as species richness and Shannon entropy are unsuitable to differentiate areas growing many genetically similar, but nonetheless differentiated (by name), crop varieties from those areas with

many genetically distant crop varieties. Thus, characterizing the biodiversity of a crop species in modern agricultural landscapes requires a phylogenetically informed approach, which by construction incorporates the notion of varietal relatedness into the measure of diversity.

A growing number of phylogenetically informed biodiversity measures have been proposed to account for taxonomic, functional or phylogenetic similarities among species within a community, such as Rao's quadratic entropy (Rao 1982), taxonomic cladistics diversity (CD) (Vane-Wright et al. 1991), phylogenetic diversity (PD) (Faith 1992), pure diversity measure (Solow et al. 1993; Solow and Polasky 1994), functional diversity (FD) (Tilman 2001) and many others (e.g., Crozier 1992; Weitzman 1992; Warwick and Clarke 1995; Chao et al. 2010; Chiu and Chao 2014). Among these alternatives, one common approach to account for the genetic relatedness among biological individuals is the phylogenetic diversity (PD) measure proposed by Faith (1992). This diversity measure is defined as the sum of all the phylogenetic branches along the minimum spanning path to quantify the evolutionary history shared among individuals. Weitzman (1992) showed that a community's diversity value can be represented by the branch length of the hypothetical phylogenetic tree.

PD measures typically focus on the presence or absence of a species to measure the overall genetic variation within a community, without taking into account the relative abundance of each species. However, species abundance provides crucial additional information regarding the composition of the community, especially for agroecosystems where a few popular crop varieties may dominate the majority of the landscape, while numerous other varieties account for comparatively small portions of the overall cropped area. To incorporate *both* species abundance and species phylogenetic distances, Chao et al. (2010) generalized the traditional phylogenetic

measure and proposed a PD measure based on Hill numbers that quantifies “the mean effective number of species,” and in so doing unified many of the existing measures of biodiversity.

To calculate a generalized PD, a phylogenetic tree is first constructed based on distances between species of the community using the UPGMA method (unweighted pair group method with arithmetic mean). Both molecular markers and pedigree information have been used in major crop genetic diversity studies to derive genetic distances among crop varieties. For this study, we use the coefficient of parentage (COP) concept to infer genetic relatedness from pedigree information on all named U.S. wheat varieties planted during the period 1919 to 2019. Following Murphy et al. (1986), COP calculates the proportion of shared genetic material among varieties based on their respective pedigrees under the following assumptions: (1) a cultivar inherits half of its genes from each parent; (2) all parental lines are homozygous and homogeneous; and (3) all landraces are unrelated to each other. Defining a pair-wise dissimilarity index between variety i and variety j as d_{ij} , we can obtain a pair-wise dissimilarity matrix D for the collection of all wheat varieties over the entire study period in the U.S. Based on the pairwise dissimilarity matrix, a phylogenetic tree can be constructed for all U.S. wheat varieties. Using the phylogenetic tree, the generalized PD (denoted as \underline{H}) for a community of wheat varieties can then be calculated using Chao et al.'s (2010) method as:

$${}^q\bar{H}(\bar{T}) = \left[\sum_{i=1}^B T_i \times \left(\frac{a_i}{\bar{T}} \right)^q \right]^{\frac{1}{1-q}}$$

where B is the number of branch segments in the tree, T_i denotes the length of branch i ($i = 1, 2, \dots, B$), a_i denotes the branch abundance (sum of relative abundance of all species descended from branch i), q denotes the exponent value (i.e., order) given to the branch abundance normalized by the mean branch length, which is defined as $\bar{T} = \sum_{i=1}^B T_i a_i$. For special cases of

the order q spanning the entire age of the phylogenetic tree, it is shown that ${}^0\bar{H}(\bar{T})$ becomes the total branch length, which is the traditional Faith's PD; ${}^1\bar{H}(\bar{T})$ can be linked to a generalization of Shannon entropy to incorporate phylogenetic distances; and ${}^2\bar{H}(\bar{T})$ can be linked to Rao's quadratic entropy (Chao et al. 2010; Chiu and Chao 2014). The phylogenetic Hill number is then calculated as:

$${}^q\bar{D}(\bar{T}) = \frac{{}^q\bar{H}(\bar{T})}{\bar{T}} = \frac{1}{\bar{T}} \left[\sum_{i=1}^B T_i \times \left(\frac{a_i}{\bar{T}} \right)^q \right]^{\frac{1}{1-q}}$$

For a state with N equally common species that are completely distinct from each other along the phylogenetic tree, the diversity measure ${}^q\bar{D}(\bar{T})$ always gives exactly N . Thus, the phylogenetic Hill number ${}^q\bar{D}(\bar{T})$ can be interpreted as the effective number of maximally distinct lineages with equal relative abundance (Chao et al. 2010).

2.3 Spatial and Temporal Diversity Decomposition

The generalized PD indexes introduced above are static measures of biodiversity for a single crop (in this instance wheat) community (e.g., a U.S. state, or the U.S. as a whole). However, variation in both the spatial and temporal dimensions of an ever-changing mix of crop varieties are one of the most fundamental features of modern agricultural systems. Thus, decomposing biodiversity into its spatial and temporal components provides a means of characterizing the dynamic changes over space *and* time in agricultural biodiversity. A growing number of long-term datasets have been used to examine the spatial and temporal patterns of biodiversity change within ecological systems (e.g., Magurran et al. 2010). This study incorporates both spatial and temporal decompositions into an assessment of the diversity dynamics of a major crop species.

Whittaker (1960) first proposed the decomposition of the overall diversity (γ - diversity) into within-community (α - diversity) and between-community (β - diversity) components, using either an additive or multiplicative rule. For spatial decomposition within each time period, each U.S. state can be treated as a separate crop community, where the variation among different states reflects the spatial diversity across the landscape. For temporal decomposition with each U.S. state, each year can be treated as a separate crop community where the variation across years reflects the temporal diversity within each state.

Here we define a crop community as the collection of varieties planted within a given state for a given year for a given crop. Then a “spatial metacommunity” is defined as the collection of all communities within a single year (i.e., all states in the U.S. within a given year) and a “temporal metacommunity” is defined as the collection of communities over multiple years within the same state (i.e., all years within a given state). With these definitions of spatial and temporal metacommunities, following Marcon et al. (2014), the total species neutral HCDT entropy (γ - entropy) for a metacommunity ${}^qH_\gamma$ can be decomposed as:

$${}^qH_\gamma = {}^qH_\alpha + {}^qH_\beta = \sum_m w_m {}^qH_{\alpha m} + \sum_m w_m {}^qH_{\beta m}$$

where α - and β - entropies for the metacommunity (i.e., ${}^qH_\alpha$ and ${}^qH_\beta$) are the weighted sums of local entropies within each community (i.e., ${}^qH_{\alpha m}$ and ${}^qH_{\beta m}$). The weight w_m adjusts for sample size differences among communities, which is commonly defined as $w_m = n_m/N$ where n_m is the number of individuals (here, crop varieties) in a local community and N is the total number of individuals for a metacommunity. The α - and β - entropies for a community are calculated as:

$${}^qH_{\alpha m} = - \sum_i p_{im}^q \ln_q p_{im}$$

$${}^q_m H_\beta = \sum_i p_{im}^q \ln_q \frac{p_{im}}{p_i} \text{ where } p_i = \sum_m p_{im}$$

Similarly, as a linear transformation of generalized entropy, the generalized PD for the metacommunity ${}^q \bar{H}_\gamma(T)$ can be decomposed as:

$${}^q \bar{H}_\gamma(T) = {}^q \bar{H}_\alpha(T) + {}^q \bar{H}_\beta(T)$$

where

$${}^q \bar{H}_\gamma(T) = \sum_k \frac{T_k}{\bar{T}} {}^q H_{\gamma k}$$

$${}^q \bar{H}_\alpha(T) = \sum_m w_m \sum_k \frac{T_k}{\bar{T}} {}^q H_{\alpha k}$$

$${}^q \bar{H}_\beta(T) = \sum_m w_m \sum_k \frac{T_k}{\bar{T}} {}^q H_{\beta k}$$

The corresponding decomposition of the diversity index ${}^q \bar{D}_\gamma(T)$, also known as the phylogenetic Hill number, is then obtained as:

$${}^q \bar{D}_\gamma(T) = {}^q \bar{D}_\alpha(T) {}^q \bar{D}_\beta(T)$$

where

$${}^q \bar{D}_\gamma(T) = e_q^{{}^q \bar{H}_\gamma(T)}$$

$${}^q \bar{D}_\alpha(T) = e_q^{{}^q \bar{H}_\alpha(T)}$$

$${}^q \bar{D}_\beta(T) = e_q^{\frac{{}^q \bar{H}_\beta(T)}{1+(1-q){}^q \bar{H}_\alpha(T)}}$$

Here e_q^x is the transformed exponential defined as $e_q^x = [1 + (1 - q)x]^{1/(1-q)}$

Depending on the grouping of communities into a given “spatial metacommunity” or a “temporal metacommunity”, the above formula allows us to decompose diversity into its respective spatial or temporal dimensions. Essentially, for a metacommunity (i.e., a collection of local

communities), the overall γ -diversity is decomposed into an average local community diversity (α -diversity) and a measure of the effective number of communities (β -diversity). Intuitively, for a given year in the U.S. with an overall phylogenetic diversity (PD) index value γ , the α -diversity component of this “spatial metacommunity” indicates the average diversity of wheat varieties growing within a state (i.e., the average effectively distinct number of varieties from maximally distinct lineages with equal relative abundance in a state), while the β -diversity component of this “spatial metacommunity” indicates the effectively distinct number of states (i.e., the equivalent number of states that each has α effective number of varieties that are distinct from each other). Similarly, for a “temporal metacommunity” (i.e., multiple periods or epochs for a given state) with an overall phylogenetic diversity (PD) index value γ , the α -component indicates the average diversity of wheat varieties growing in each epoch in this state, while the β -component indicates the effectively distinct number of epochs (i.e., the equivalent number of epochs where each has α effectively distinct number of varieties). Using a unique long-run panel dataset for a major crop species, such spatial and temporal decompositions allow us to better understand the impact of crop varietal turnover on modern agricultural genetic diversity and address key questions concerning 1) the overall trends in the phylogenetic diversity of the U.S. wheat crop, and 2) changes in either the spatial or temporal dimensions of diversity in the U.S. wheat crop over the past century.

3. Results

3.1 Phylogenetically Blind Measures of Wheat Biodiversity

U.S. planted wheat area peaked in 1981 (88.3 million acres), thereafter declining steadily to just 45.5 million acres by 2019 (USDA-NASS 2021). In recent years, around 70% of the acreage was

planted to winter wheat, with spring wheat averaging around 25% and durum wheat less than 5% (USDA-NASS 2021). Spring and durum wheats are grown mostly in the Northern Plains, while winter wheat is mainly planted in the Central Plains and Pacific Northwest. Among all three classes of wheat that were commercially grown across 16 major wheat-producing states during the period 1919-2019, we identified a total of 1,353 unique named varieties, each accounting for at least 0.5% state-level area share by market class in the year they were grown.

As shown in Fig.1, the area dynamics of major varieties exhibits a strong regularity. First, the acreage shares attributed to the top wheat varieties decline over time (see decrease in overall colored areas in Fig. 1). In 1919 the top five varieties accounted for 88% of the U.S wheat planted acreage; by 1964 that share had shrunk to just 34%. The colored plus darker grey elements of Fig. 1 signify the share of area planted to the top 20 varieties (99% in 1919 down to 68% in 2019). Second, the temporal pattern of major variety uptake typically follows a logistic function (consistent with Griliches' 1957 classic study of the U.S. adoption of hybrid corn), followed by a period of dis-adoption as newer varieties gained popularity. Finally, Fig 1. graphically illustrates the waves of faster varietal turnover that have consistently swept through the U.S. wheat crop over the past century. Specifically, earlier varieties (such as *Turkey*) that were once dominant for several decades are no longer major varieties, while in recent years the top varieties tend to completely turn over within a period of 5 years.

Fig. 2 plots several phylogenetically blind diversity metrics that summarize important dimensions of the varietal dynamics shown in Fig. 1. Our data reveal that the number of major commercially grown wheat varieties increased steadily over time, from just 33 in 1919 to 186 in 2019. On a per-million-acre basis, the intensity of varietal use indicates an increasingly diverse spatial pattern. In 1919, varietal intensity averaged 0.8 varieties per million acres. A century

later, farmers were using a much more diverse portfolio of varieties; varietal intensity had increased more than 10-fold to average 9.1 varieties per million acres in 2019.

In 1919 just 1.3% of the planted area was sown to new varieties (i.e., <5 years old). By 2019, more than one third (36.0%) of the U.S. wheat area was planted to new varieties, such that the area-weighted age of commercially grown varieties declined dramatically from 36.4 years in 1919, to 16.0 years in 1960, and down to just 9.3 years in 2019. The decline in the area share of older varieties (i.e., >15 years since release) was particularly pronounced; 68.8% in 1919, and only 13.7% in 2019.

The average trends on varietal longevity mask a good deal of variation in the commercially useful life of individual varieties. For example, the winter wheat variety *Cheyenne*, bred at the Nebraska Agricultural Experiment Station and released for commercial use in 1933, was planted for a period of 80 years, disappearing from the varietal statistics in 2014. Its area peaked at around 2.5 million acres in 1959, falling steadily to less than 5,000 acres in 2013. In our collection, a total of 453 (33.5% of the 1,353 total) wheat varieties were long-lived (i.e., commercially grown for more than 15 years). In contrast, 93 varieties were especially short-lived, with recorded commercial use of just one year.

3.2 Spatio-temporal Phylogenetic Diversity Patterns

While the phylogenetically blind measures presented above unequivocally reveal increasing varietal diversity in the U.S. wheat crop, they can be potentially misleading indicators of the extent of genetic erosion associated with the use of scientifically selected crop varieties. It is the perception of a narrowing of the genetic variation in modern cropping systems that is most closely associated with the concerns over the resilience of these systems to current and prospective climate and pest shocks. Phylogenetically blind measures of diversity will overstate

the degree of genetic diversity within a given population or area extent (e.g., a field, state or country) when the varieties in that population or locale are genetically related through shared breeding materials.

To infer the genetic relatedness using pedigree data, in this study we utilize a comprehensive collection of U.S. wheat pedigree information (see Methods). As Fradgley et al. (2019) demonstrated in a recent large-scale wheat genetic diversity study across 454 wheat varieties, there is high correlation between pedigree- and marker-based kinship coefficients, confirming the value of using pedigree information to inform and manage wheat genetic diversity. The phylogenetic tree on the LHS of Fig. 3. graphically depicts the genetic relatedness of 1,353 commercial wheat varieties grown in the U.S. during the period 1919-2019. The genetic distance between each pair of wheat varieties is represented by the pedigree branch length between them, based on their respective coefficient of parentages (COPs) that we calculated using the method described by Murphy et al. (1986). The horizontal, colored lines in Fig. 3 indicate the presence or absence of each of these commercially grown wheat varieties for each of the years 1919-2019.

The phylogenetic variation among wheat varieties in the U.S. increased over time, as revealed by the expanding coverage across the phylogenetic tree over time. The three market classes for wheat differ in their clustered locations within the phylogenetic tree. Most of the durum wheats (red lines) are concentrated within a closely clustered region on the phylogenetic tree. Winter wheats exhibit a much more diverse phylogenetic background than durum or spring wheats, with varieties ranging across the entire phylogenetic tree. In addition, the phylogenetic background of winter wheats constantly changes as the use of particular varieties waxes and wanes over time.

Phylogenetic Diversity (PD) Indexes

We used the genetic relatedness among varieties based on their respective position within the phylogenetic tree, and the abundance of each variety based on its state-specific planted area shares, to calculate a PD index for each of the 16 major U.S. wheat producing states during the period 1919-2019 using the method described by Chao et al. (2010). The overall, and dominant, trend is for the phylogenetic diversity of wheat to increase over time in all the larger wheat growing states for all three market classes (Fig. 4). For winter wheat, many states—including Kansas, Kentucky, Montana, Nebraska, Oklahoma, Texas and Washington—had substantial increases in phylogenetic diversity, especially in the more recent decades, while some of the smaller wheat growing states (Oregon, Wyoming, Indiana) had rather stagnant phylogenetic diversity over the longer run. The PD indexes for spring wheat are generally similar to those for winter wheat within states growing both market classes (such as South Dakota, Montana and Washington). In California, Montana and South Dakota, the PD indexes for durum wheats are almost always less than those for winter wheat, whereas in North Dakota the phylogenetic diversity of durum wheat is comparable to that of spring and winter wheat.

Temporal Decomposition of PD

The phylogenetic diversity trends reported above track state-level changes over time in a PD index, where that index is calculated by treating each state at each point in time as a discrete community of wheat varieties. Adopting an approach introduced by Whitakker (1960) in his study of diversity changes in the vegetative cover of two forested regions, a more nuanced, and more insightful measurement approach to analyzing the variation in phylogenetic diversity among U.S. wheat varieties is to introduce the concept of a metacommunity. In this first instance, we treat the wheat varieties grown over the entire century in each state as a distinct “temporal metacommunity,” such that the overall phylogenetic diversity (aka γ_T -diversity) of

each state-specific metacommunity can be decomposed into two components: an α_T -diversity, which captures the average effective number of phylogenetically distinct varieties planted each year in a given state, and a temporal β_T -diversity, which measures the effective number of phylogenetically distinct epochs for a state. Thus, for example, states with higher numbers of phylogenetically distinct varieties in a year will have larger α_T -diversity values, while those states with higher turnover rates of phylogenetically distinct varieties will have higher β_T -diversity values. Fig. 5 plots the α_T -diversity (number of effectively distinct varieties per year) on the x-axis and temporal β_T -diversity (number of effectively distinct epochs) on the y-axis for each state, where the rectangle area (equal to $\alpha_T \times \beta_T$) represents each state's overall (spatio-temporal) γ_T -diversity.

Winter wheat states differ markedly in terms of their α_T -, β_T - and γ_T -diversities. Washington, Idaho and Texas are the top three ranked states in terms of their overall γ_T -diversities, with both high α_T -diversity (averaging 6.7, 6.7 and 6.8 effectively distinct varieties per year, respectively) and high temporal β_T -diversity (5.1, 5.0 and 4.7 distinctly different epochs during the 101-year period spanning 1919 to 2019). In contrast, Wyoming has the lowest overall γ_T -diversity, with both the lowest α_T -diversity (averaging just 3.2 effectively distinct varieties each year) and lowest temporal β_T -diversity (with just 2.7 phylogenetically distinct epochs). Kansas is the most β_T -diverse state, with 5.7 phylogenetically distinct epochs during the period 1919-2019, indicative of a rapid rate of varietal turnover.

For spring wheat, North Dakota has the highest overall γ_T -diversity with 4.4 effective varieties each year (α_T -diversity), and 6.0 phylogenetically distinct years (β_T -diversity). Idaho and Minnesota have the highest α_T -diversity and β_T -diversity respectively. In Washington, the α_T -diversity for spring wheat (4.7) is smaller than it is for winter wheat (6.7), with a difference of

2.0 effectively distinct varieties. Spring wheat in North Dakota and South Dakota have higher α_T -, β_T - and γ_T -diversities than wheats in the other two market classes, suggesting both a higher effectively distinct number of varieties and a faster rate of replacement for spring wheat in these states.

Among the four durum wheat states, Montana planted 3.6 effectively distinct varieties each year, higher than North Dakota, South Dakota, and California which ranged from 2.0 to 2.7 effectively distinct varieties annually. California has the highest temporal β_T -diversity (4.5 distinct epochs) among all four states.

Spatial Decomposition of PD

To characterize the spatial variation of wheat varieties across states, the collection of all major wheat growing states in the U.S. for each year is treated as a “spatial metacommunity”. In this instance, the γ_S -diversity represents the overall diversity in the U.S., which can be decomposed into α_S -diversity—representing the average number of effectively distinct varieties in a state—, and the spatial β_S -diversity—which represents the effective number of phylogenetically distinct states in the U.S. For the three market classes of wheat in the U.S., their overall γ_S -diversity and its α_S - and β_S - decompositions are plotted over time in Fig. 6. Generally, the diversity indexes are increasing over time for all three market classes of wheat in the U.S., where winter wheat has the highest overall γ_S -diversity followed by spring wheat. Winter wheat and spring wheat have similar α_S -diversity indexes over time, suggesting that the state-level wheat varietal diversities are similar among these two market classes. However, winter wheat has much higher spatial β_S -diversities than spring wheat, consistent with the fact that winter wheat is grown in more states over larger geographic areas than spring wheat. This is also consistent with the notion that market forces are likely to drive towards the development and deployment of varieties that better

align varietal genetics across the diverse, location-specific agroecologies in which they are grown.

With only four major durum wheat states, the spatial β_S -diversity and overall γ_S -diversity for durum are the lowest among the three market classes. The overall γ_S -diversities for the U.S. wheat crop are heavily influenced by both the α_S -diversity and β_S -diversity, suggesting that both the within-state varietal diversity and across-state spatial variations are important for the overall diversity of the wheat population across each of the three market classes.

4. Discussion

Our temporal, spatial and varietal coverage constitutes a superset of the coverage contained in previous studies of the changing crop diversity in U.S. wheat production (Table S1). And, going beyond these prior studies, we used the concept of a meta-community to examine both the spatial and temporal dimensions of crop diversity. Our conclusions derived from both phylogenetically blind diversity metrics, in which different crop varieties are counted as equally distinct varieties—which tend to inflate estimates of diversity—as well as phylogenetically informed metrics that take account of the relatedness of scientifically-selected varieties. Despite these analytical differences, our results provide further, more comprehensive, evidence that the increasingly intensive use of scientifically selected crop varieties generally leads to more, not less, biodiverse cropping practices, at least regarding diversity in the U.S. wheat crop over the past century.

Developing accurate estimates of varietal similarity is still challenging. De novo sequences of the complete genome, let alone hundreds of thousands of marker sites for every wheat variety, are still not available. Hence, we have used COPs—estimates of relatedness based on pedigrees—to

infer genetic relatedness. Despite the unrealistic assumption of the absence of selection that the COP calculation employs, it has been shown to align more closely with direct marker-based assessments of relatedness among wheat varieties as the number of assessed markers increases (Fradgley et al. 2019). This provides further confidence that the phylogenetically informed diversity statistics we have computed are indicative of the changing spatio-temporal pattern of wheat genetic diversity.

Our findings add to a growing body of empirical evidence of generally enhanced, not diminished, phylogenetic crop diversity in modern agricultural systems (see, e.g., the compilation in Table S1). Moreover, this finding can profoundly reframe the policy and practical implications of a sizable body of literature that ignores phylogenetic (within-crop species) diversity when assessing the implications of intensified cropping systems on agricultural biodiversity. For example, a recent meta-review of 89 studies across 1,475 locations on the relationships between biodiversity and cropping agriculture concluded that “Up to 50% of the negative effects of landscape simplification on ecosystem services was due to richness losses of service-providing organisms, with negative consequences for crop yields” Dasinense et al. (2019, p. 1). However, this study did not address the notion of within-crop (richness and abundance) diversity. Moreover, the spatial (and temporal) extents of the sampled locations included in the meta review are limited (i.e., generally data for a single or just a few years across individual fields, farms or nearby areas) relative to the U.S.-wide wheat acreage spanning a century of production included in the present study. Notably, and contrary to Dasinense’s conclusion, U.S. average wheat yields rose fourfold (from 12.9 bu/ac in 1919 to 51.7 bu/ac in 2019) in a monoculture cropping system that we show is getting genetically more complex from a phylogenetic perspective.

What if any lessons do these findings have in terms of the biodiversity concerns embodied in the U.N.'s sustainable development goals and similar concerns by many other institutions and individuals? Setting policies, and subsequently putting them into practice, to achieve desired outcomes involving the biodiversity-agricultural nexus is a complex, multi-objective problem. Moreover, conserving or enhancing (within species) crop biodiversity in and of itself is a multi-instrument, multi-objective problem, where the crop biodiversity outcomes are envisaged as a means (instrument) to other ends (e.g., yield growth; resilience to a multitude of climate or pest and disease shocks; and various food security, access and equity outcomes). Beyond the genetic diversity of the crop per se, there is above and below ground (e.g., plant, microbe and insect) biodiversity within cropping systems, the biodiversity inherent in dual animal-livestock production systems, and the much more encompassing notions of managed, natural or wild systems of biodiversity involving land outside of agriculture to be considered.

Getting some clarity on the distinction between policy instruments and targets is useful in thinking through the policies and practices concerning agricultural biodiversity. As Alston and Pardey (2021) point out, Tinbergen's Rule (Arrow 1958) calls on policymakers to use at least one policy instrument per policy target, while Corden's (1974) corollary argues for matching instruments closely to targets. Failing to keep policy instruments and targeted outcomes closely aligned is often ineffective (in that one instrument is unlikely to be equally optimal for achieving more than one outcome) and socially wasteful (in that misaligned instruments and targets are unlikely to be cost effective, incurring higher than otherwise direct and opportunity costs).

If enhancing crop (within species) diversity at consequential temporal and spatial scales is the target, our results show that the intelligent intensification of cropping systems using scientifically selected crop varieties can be an especially effective instrument. Moreover, in this particular

instance, this same instrument has multiple other desired outcomes. The higher yields associated with more-intensive cropping systems clearly have positive food security implications by increasing crop output, lowering the unit costs of production, or both. This in turn lowers the price of food, with especially equitable impacts on poorer people who spend a substantially larger share of their meager incomes on food (Muhammad et al. 2011).

If our results are generalizable such that intensive cropping systems generally increase both yields and crop genetic diversity, they are also likely to be impactful on biodiversity more generally by shrinking the footprint of cropping agriculture. In the spirit of Waggoner (1996) and Borlaug (2007), Stevenson et al. (2013), for example, carried out a counterfactual simulation to assess the land use consequences of global agriculture absent the Green Revolution yield gains associated with genetic improvement in wheat, rice, coarse grains and other crops (cassava, lentils, beans and potatoes) over the period 1961-2004. They concluded that "... the total crop area in 2004 would have been between 17.9 and 26.7 million hectares larger in a world that had not benefited from crop germplasm improvement since 1965. Of these hectares, 12.0–17.7 million would have been in developing countries, displacing pastures and resulting in an estimated 2 million hectares of additional deforestation (Stevenson et al. 2013, p. 8363)." In fact, U.S. wheat production increased by 3.3-fold over the past century (USDA-NASS 2021) while planted area declined by 41%. Similarly, the 2019 global area in wheat was roughly comparable to the acreage in the late 1960s, so that output grew by 3-fold in-line with the increase in average global wheat yields (FAO 2021).

If enhancing both within- and between-species biodiversity in modern cropping systems are the two prioritized policy targets, then the Tinbergen-Corden insights would argue in favor of using at least two instruments. For example, intensive use of scientifically selected varieties of a

particular crop could be coupled with other genetic and crop management instruments. However, that may incur sizable opportunity costs, if by seeking to simultaneously enhance both within- and between-species diversity the productivity (yield or farmer cost of production) performance of one or either crop species is compromised, thus leading to an increase in the overall footprint of cropping agriculture with unintended, albeit negative, consequences for biodiversity conservation more generally.

The choice of targets and instruments is naturally context sensitive. In poorer parts of the world, the priority policy target may be to improve the well-being of poor people, especially by way of increasing crop productivity (thus expanding the supply and lowering the price of food, with equitable implications for poorer consumers given larger shares of their incomes are typically spent on food). In this instance, our results indicate that an efficient instrument to cost-effectively achieve well-being targets for poorer people would be to intelligently intensify wheat production by way of expanding the use of scientifically selected wheat varieties; a strategy that calls for doubling down on the science and seed systems that at present often underserve farmers in poorer countries. The other positive finding from our study is that this particular configuration of policy instruments and targets is also likely to improve biodiversity outcomes, both directly by increasing the spatio-temporal diversity of wheat varieties in use, and indirectly by stalling the expansion or even shrinking the footprint of unnatural agricultural landscapes in favor of increasing areas in more natural environments.

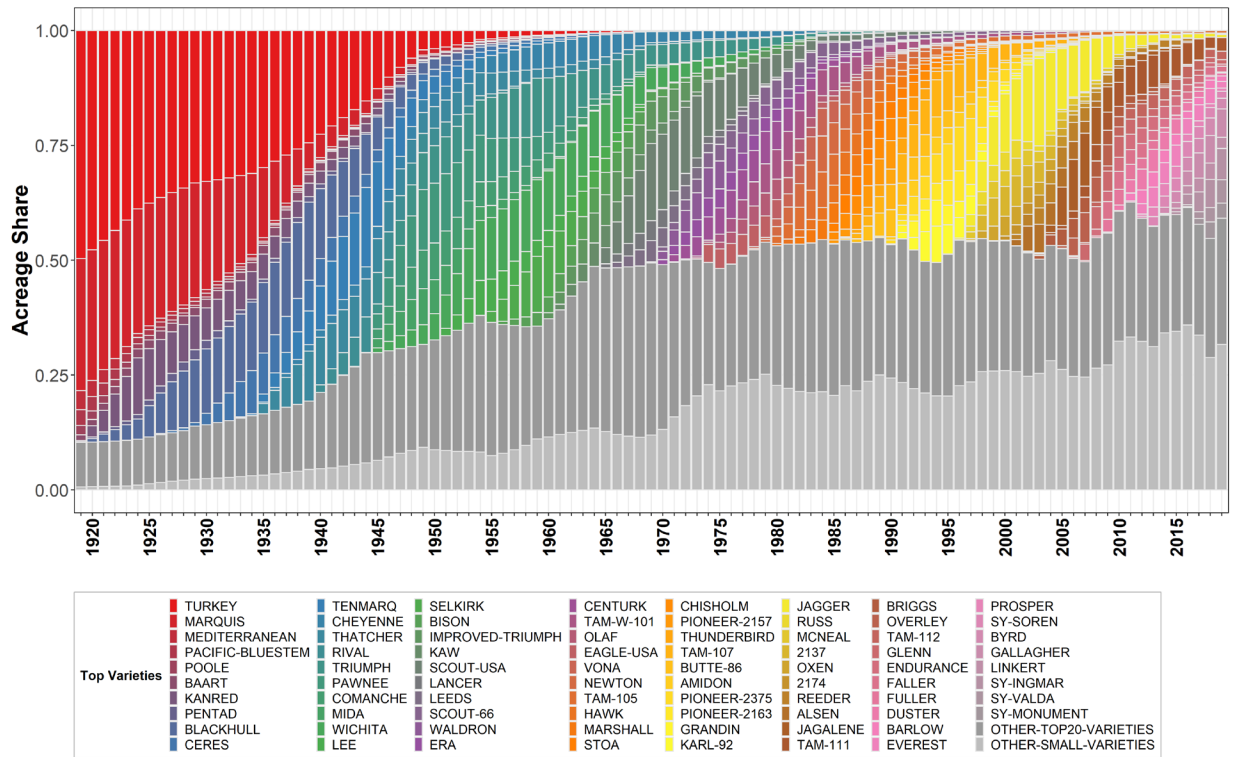


Fig. 1. Turnover of the top wheat varieties in the U.S., 1919-2019.

Notes: Colored varieties represent varieties that were among the top 5 in area share in at least one year; “OTHER-TOP20-VARIETIES” represent varieties that were among the top 20 in area share in at least one year (excluding the top 5 varieties colored already); “OTHER-SMALL-VARIETIES” represent the remaining named varieties (excluding unknown varieties).

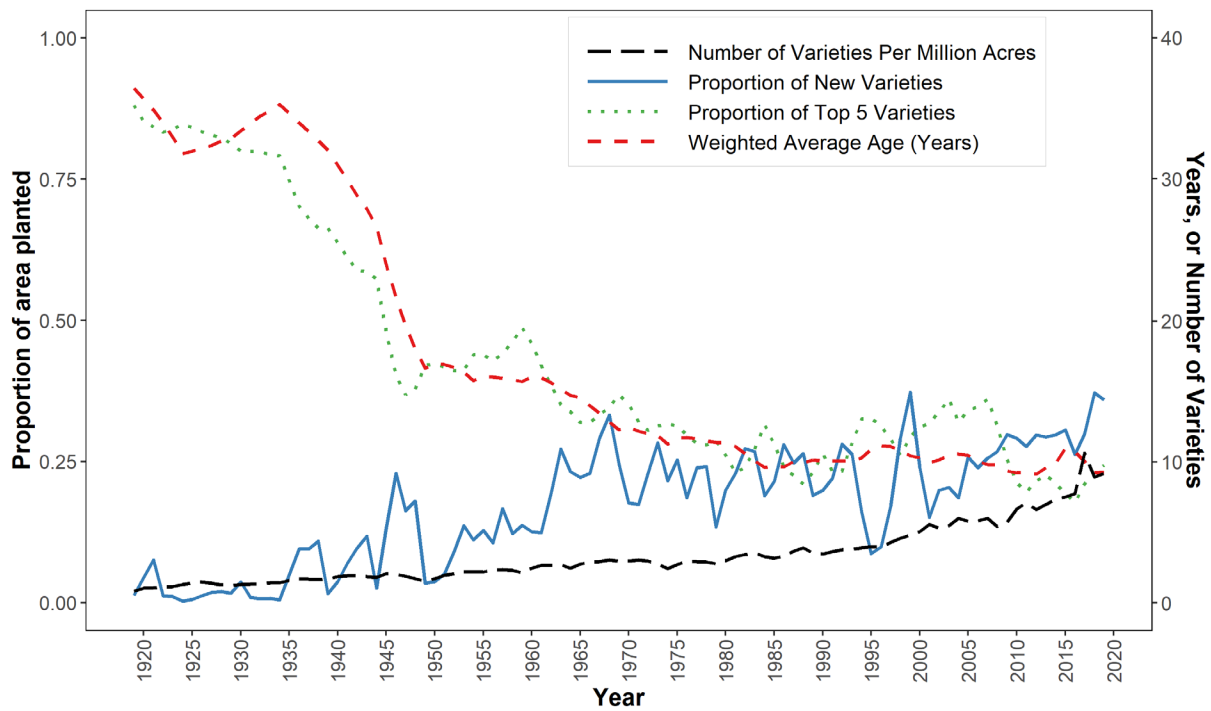


Fig. 2. Phylogenetically blind measures of U.S. wheat diversity, 1919-2019.

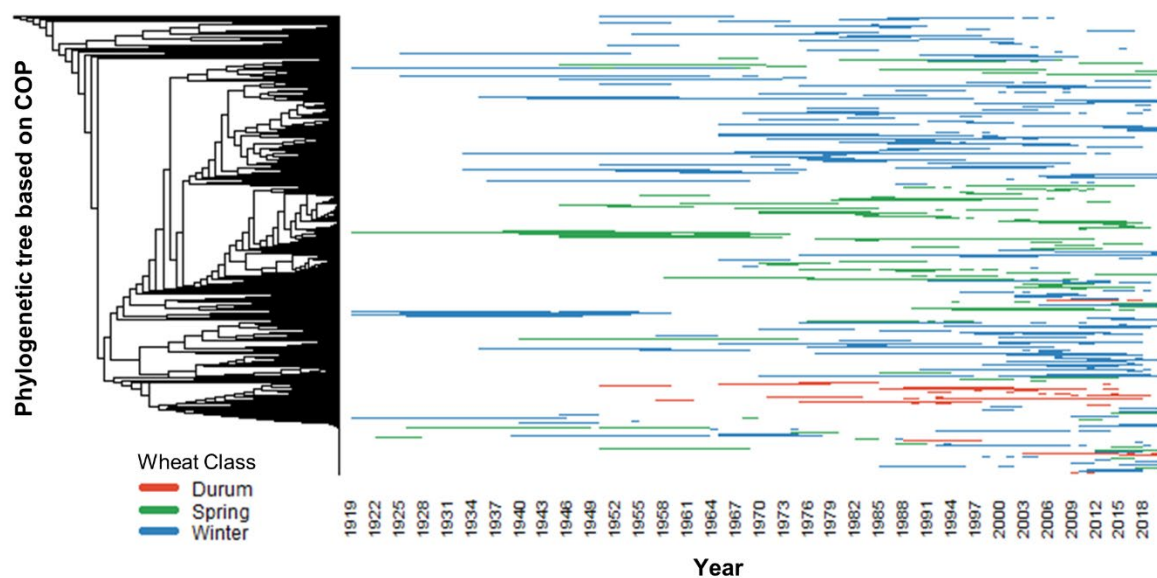


Fig. 3. Changes over time in the phylogenetic variation of commercially grown U.S. wheat varieties, 1919-2019.

Notes: The left-hand-side phylogenetic tree illustrates the distances among commercial wheat varieties grown in the U.S. during 1919-2019 based on their coefficient of parentage. Each colored dash line represent the use of a particular wheat variety for a given year in the U.S. with the color representing different wheat market classes.



Fig. 4. Phylogenetic diversity for major U.S. wheat producing states, 1919-2019.

Notes: Lines represent the Phylogenetic Diversity indexes of the order 1 for each state by different wheat market classes during 1919-2019. Similar trends are observed using PD indexes of order 2 (see Supplementary Materials).

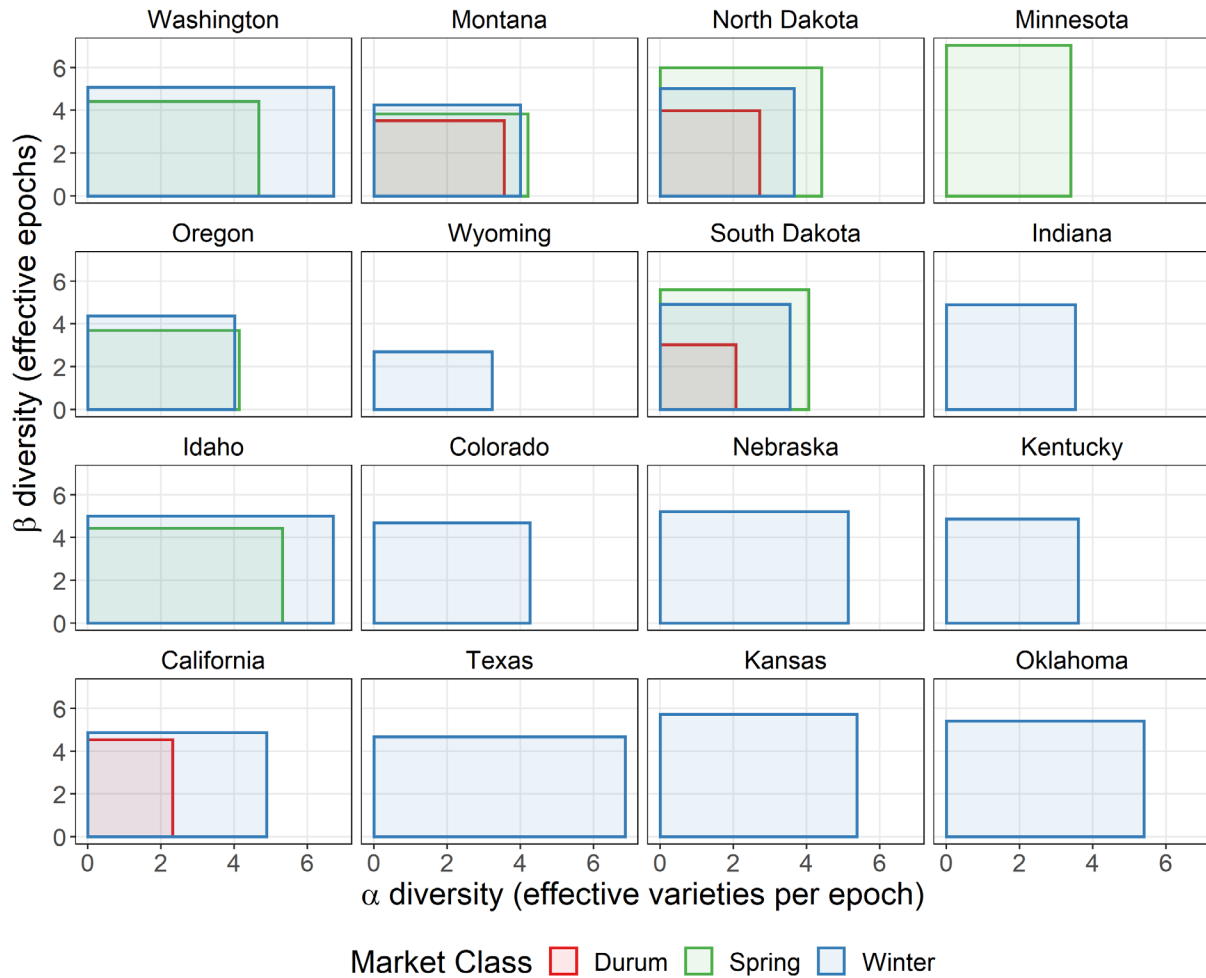


Fig. 5. Temporal decomposition of phylogenetic diversity for major U.S. wheat producing states, 1919-2019.

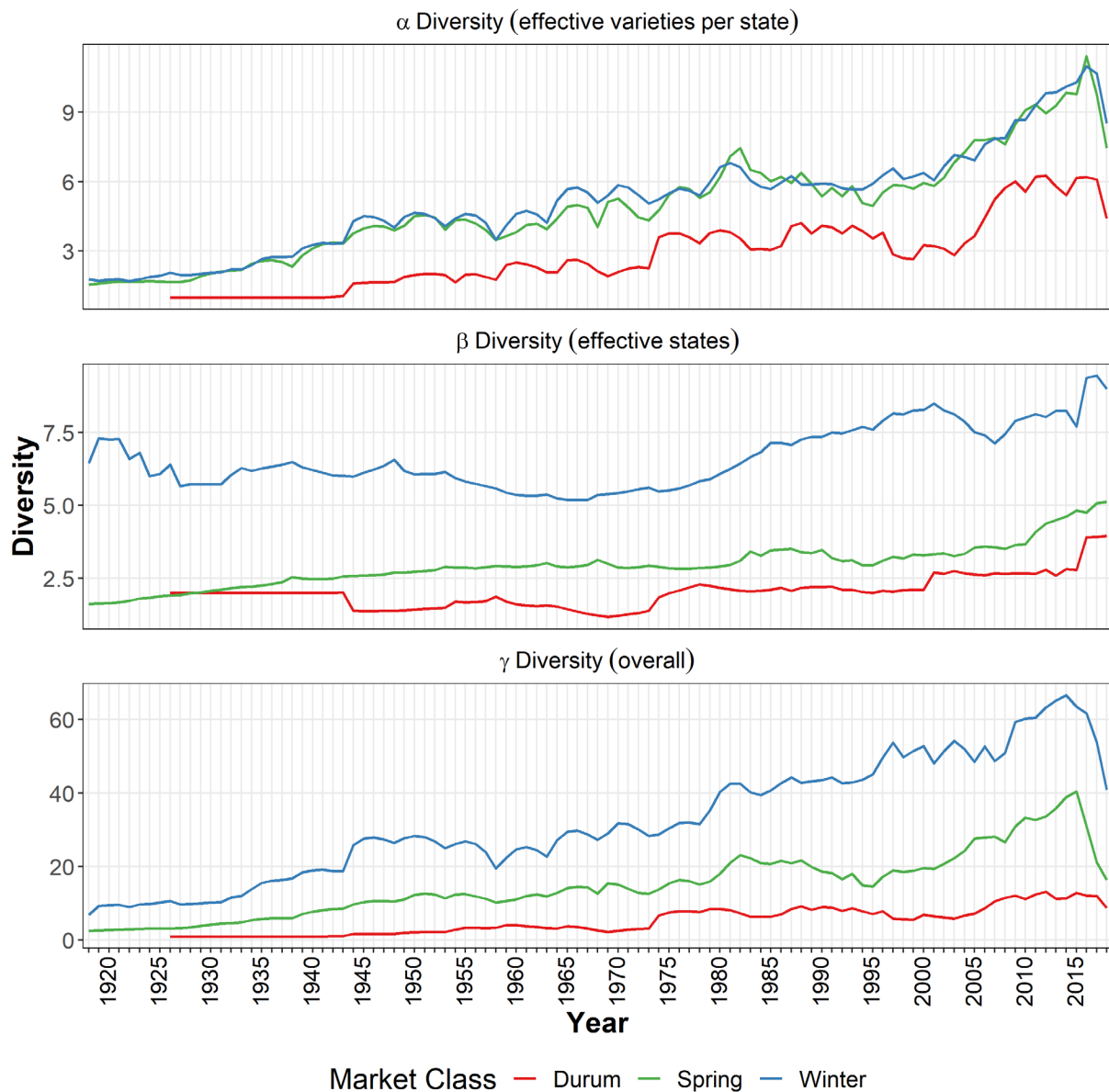


Fig. 6. Spatial decomposition of phylogenetic diversity for major U.S. wheat producing states, 1919-2019.

Notes: Limitations in reported area-by-variety data in more recent years (especially after 2016) accounts for the decline in diversity indexes for the more recent years.

References

- Almanza-Pinzon, M.I., Khairallah, M., Fox, P.N. and Warburton, M.L., 2003. Comparison of molecular markers and coefficients of parentage for the analysis of genetic diversity among spring bread wheat accessions. *Euphytica* 130(1), pp.77-86.
- Alston, J.M., P.G. Pardey. The economics of agricultural innovation, chapter in C.B. Barrett and D.R. Just, eds. *Elsevier Handbook of Agricultural Economics, Volume 5*. Amsterdam: Elsevier (2021).
- Arrow, K.J. Tinbergen on economic policy. *J. Am. Stats. Assoc.* 53(281), 89–97 (1958).
- Ball, C.R. The history of American wheat improvement. *Ag. History* 4(2), 48-7 (1930).
- Barrett, B.A., K.K. Kidwell, and P.N. Fox. 1998. Comparison of AFLP and pedigree-based genetic diversity assessment methods using wheat cultivars from the Pacific Northwest. *Crop Science* 38(5):1271–1278.
- van Beuningen, L.T., and R.H. Busch. 1997. Genetic diversity among North American spring wheat cultivars: I. analysis of the coefficient of parentage matrix. *Crop Science* 37(2):cropsci1997.0011183X003700020043x.
- Borlaug, N. Feeding a hungry world. *Science* 318, 359 (2007).
- Biffen, R.H. Mendel's law of inheritance and wheat breeding. *J. Ag. Sci.*1(1), 4-48(1905).
- Brennan, J.P. and D. Byerlee. The rate of crop varietal replacement on farms: measures and empirical results for wheat. *International Journal of Plant Varieties and Seeds* 4, 99-106 (1991).
- Brush, S. B. *Farmers' Bounty: Locating Crop Diversity in Contemporary World*. Yale University Press. (2004).
- Chao, A., C.H. Chiu, and L. Jost. 2010. Phylogenetic diversity measures based on Hill numbers. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 365(1558):3599–3609.
- Chao, A., C.H. Chiu, and L. Jost. 2014. Unifying Species Diversity, Phylogenetic Diversity, Functional Diversity, and Related Similarity and Differentiation Measures Through Hill Numbers." *Annual Review of Ecology, Evolution, and Systematics* 45(1):297–324.
- Chao, S., W. Zhang, J. Dubcovsky, and M. Sorrells. 2007. Evaluation of genetic diversity and genome-wide linkage disequilibrium among U.S. wheat (*triticum aestivum* l.) germplasm representing different market classes. *Crop Science* 47(3):1018–1030.
- Chen, H.B., J.M. Martin, M. Lavin, and L.E. Talbert. 1994. Genetic diversity in hard red spring wheat based on sequence-tagged-site PCR markers. *Crop Science* 34(6):cropsci1994.0011183X003400060037x.
- Chiu, C.H., and A. Chao. 2014. Distance-based functional diversity measures and their decomposition: a framework based on Hill numbers. *PLOS ONE* 9(7):e100014.
- CIMMYT. Genetic Resource Information System (GRIS) for Wheat and Triticale. Online at <http://wheatpedigree.net>

- Corbellini, M., Perenzin, M., Accerbi, M., Vaccino, P. and Borghi, B., 2002. Genetic diversity in bread wheat, as revealed by coefficient of parentage and molecular markers, and its relationship to hybrid performance. *Euphytica* 123(2), pp.273-285.
- Corden, W.M. *Trade Policy and Economic Welfare*. Oxford: Oxford University Press (1974).
- Cox, T.S., J.P. Murphy, D.M. Rodgers. Changes in genetic diversity in the red winter wheat regions of the United States. *PNAS*. 83, 5583-5586 (1986).
- Crozier, R.H. 1992. Genetic diversity and the agony of choice. *Biological Conservation* 61(1):11–15.
- Dainese, M., Martin, E.A., Aizen, M.A., Albrecht, M., Bartomeus, I., Bommarco, R., Carvalheiro, L.G., Chaplin-Kramer, R., Gagic, V., Garibaldi, L.A. and Ghazoul, J., 2019. A global synthesis reveals biodiversity-mediated benefits for crop production. *Science Advances* 5(10), p.eaax0121.
- Daróczy, Z. 1970. Generalized information functions. *Information and Control* 16(1):36–51.
- Delabre, I., L.O. Rodriguez, J.M. Smallwood¹, J.P.W. Scharlemann, J. Alcamo¹, A.S. Antonarakis, P. Rowhani, R.J. Hazell, D.L. Aksnes, P. Balvanera, C.J. Lundquist, C. Gresham, A.E. Alexander, N.C. Stenseth. Actions on sustainable food production and consumption for the post-2020 global biodiversity framework. *7 Sci. Adv.* (2021): eabc8259
- Duvick, D.N. Genetic diversity in major farm crops on the farm and in reserve. *Econ. Botany* 38(2), 161-178(1984).
- Faith, D.P. 1992. Conservation evaluation and phylogenetic diversity. *Biological Conservation* 61:1–10.
- FAO (Food and Agriculture of the United Nations). 2021. FAOSTAT Database. Rome: FAO. Available at <http://www.fao.org/faostat/en/#data>.
- Fisher, R.A., A.S. Corbet, and C.B. Williams. 1943. The relation between the number of species and the number of individuals in a random sample of an animal population. *Journal of Animal Ecology* 12(1):42–58.
- Fowler, C., *Unnatural Selection: Technology, Politics and Plant Evolution*. Switzerland: Gordon and Breach (1994).
- Fowler, C., and P.R. Mooney. 1990. *Shattering: Food, Politics, and the Loss of Genetic Diversity*. Tucson, AZ: University of Arizona Press.
- Fradgley, N., K.A. Gardner, J. Cockram, J. Elderfield, J.M. Hickey, P. Howell, R. Jackson, I.J. Mackay. A large-scale pedigree resource of wheat reveals evidence for adaptation and selection by breeders. *PLOS Biology* (2019): doi.org/10.1371/journal.pbio.3000071.
- Frankel, O.H. Genetic dangers in the green revolution. *World Agric.* 19(3), 9–13 (1970).
- Fufa, H., P.S. Baenziger, B.S. Beecher, I. Dweikat, R.A. Graybosch, and K.M. Eskridge. 2005. Comparison of phenotypic and molecular marker-based classifications of hard red winter wheat cultivars. *Euphytica* 145(1):133–146.
- Griliches, Z. Hybrid corn: An exploration in the economics of technological change. *Econometrica* 254: 501-522 (1957).

- Harlan, J.R. Our vanishing genetic resources. *Science* 188(4199):618-621(1975).
- Hartell, J., M. Smale, P.W. Heisey, and B. Senauer. The contribution of genetic resources and diversity to wheat productivity in the Punjab of Pakistan, In Smale, M. ed. *Farmers Gene Banks and Crop Breeding: Economic Analyses of Diversity in Wheat Maize and Rice*. Dordrecht: Springer Netherlands. (2000).
- Havrda, J., and F. Charvát. 1967. Quantification method of classification processes. Concept of structural α -entropy. *Kybernetika* 3(1):30–35
- Hawkes, J.G. *The Diversity of Crop Plants*. Cambridge: Harvard University Press (1983).
- Hill, M.O. 1973. Diversity and evenness: a unifying notation and its consequences. *Ecology* 54(2):427–432.
- Jost, L. 2006. Entropy and diversity. *Oikos* 113(2):363–375.
- Keylock, C.J. 2005. Simpson diversity and the Shannon–Wiener index as special cases of a generalized entropy. *Oikos* 109(1):203–207.
- Kim, H.S., and R.W. Ward. 1997. Genetic diversity in Eastern U.S. soft winter wheat (*Triticum aestivum* L. em. Thell.) based on RFLPs and coefficients of parentage. *Theoretical and Applied Genetics* 94(3):472–479.
- Lu, Y., J.M. Bullock. Biodiversity conservation in a changing environment beyond 2020. 7. *Sci. Adv.* (2021) eabl8162.
- Magurran, A.E., S.R. Baillie, S.T. Buckland, J.M. Dick, D.A. Elston, E.M. Scott, R.I. Smith, P.J. Somerfield, and A.D. Watt. 2010. Long-term datasets in biodiversity research and monitoring: assessing change in ecological communities through time. *Trends in Ecology & Evolution* 25(10):574–582.
- Marcon, E., I. Scotti, B. Hérault, V. Rossi, and G. Lang. 2014. “Generalization of the partitioning of Shannon diversity” *PLOS ONE* 9(3):e90289.
- Marić, S., Bolarić, S., Martinčić, J., Pejić, I. and Kozumplik, V., 2004. Genetic diversity of hexaploid wheat cultivars estimated by RAPD markers, morphological traits and coefficients of parentage. *Plant Breeding* 123(4), pp.366-369.
- Meng, E.C.H., M. Smale, M. Bellon and D. Grimaneli. Definition and measurement of crop diversity for economic analysis. Chapter in M. Smale, ed. *Farmers, Gene Banks and Crop Breeding: Economic analyses of Diversity in Wheat, Maize and Rice*. Boston: Kluwer Academic Publisher (1998).
- Miller, J. Genetic erosion: Crop plants threatened by government neglect. *Science* 182, 1231–1233(1973) <https://doi.org/10.1126/science.182.4118.1231>
- Muhammad, A., J.L. Seale, Jr., B. Meade, A. Regmi. *International Evidence on Food Consumption Patterns: An Update Using 2005 International Comparison Program Data*. Washington, D.C.: U.S. Dept. of Agriculture, Econ. Res. Serv. (2011). www.ers.usda.gov/webdocs/publications/47579/7637_tb1929.pdf?v=5058.1
- Murphy, J.P., T.S. Cox, and D.M. Rodgers. 1986. Cluster analysis of red winter wheat cultivars based upon coefficients of parentage. *Crop Science* 26(4):672–676.

- NAS (National Academy of Sciences). *Genetic Vulnerability of Major Crops*. Washington, D.C.: National Academy of Sciences (1972).
- Peng, J.H., D. Sun, E. Nevo. Domestication evolution, genetics and genomics in wheat. *Mol Breeding*. 28, 281–301 (2011). doi 10.1007/s11032-011-9608-4
- Poets, A., K. Silverstein, P. Pardey, S. Hearne, J. Stevenson. *DNA fingerprinting for crop varietal identification: fit-for-purpose protocols, their costs and analytical implications*. Rome: CGIAR Standing Panel on Impact Assessment (2020).
- Rao, C.R. 1982. Diversity and dissimilarity coefficients: A unified approach. *Theoretical Population Biology* 21(1):24–43.
- Shannon, C.E. 2001. A mathematical theory of communication. *SIGMOBILE Mob. Comput. Commun. Rev.* 5(1):3–55.
- Simpson, E.H. 1949. Measurement of diversity. *Nature* 163:688.
- Smale, M., J. Hartell, P.W. Heisey, and B. Senauer. 1998. The Contribution of Genetic Resources and Diversity to Wheat Production in the Punjab of Pakistan. *American Journal of Agricultural Economics* 80(3):482–493.
- Smolders, H. ed. 2006. *Enhancing Farmers' Role in Crop Development: Framework Information for Participatory Plant Breeding in Farmer Field Schools*. Centre for Genetic Resources, the Netherlands: PEDIGREA.
- Smýkal, P., R.K. Varshney, V.K. Singh, C.J. Coyne, C. Domoney, E. Kejnovský, T. Warkentin. From Mendel's discovery on pea to today's plant genetics and breeding *Theor. Appl. Genet.* 29, 2267–2280 (2016). doi 10.1007/s00122-016-2803-2
- Solow, A., S. Polasky, and J. Broadus. 1993. On the measurement of biological diversity. *Journal of Environmental Economics and Management* 24(1):60–68.
- Solow, A.R., and S. Polasky. 1994. Measuring biological diversity. *Environmental and Ecological Statistics* 1(2):95–103.
- Sthapit, S.R., K. Marlowe, D.C. Covarrubias, T.M. Ruff, J.D. Eagle, E.M. McGinty, M.A. Hooker, N.B. Duong, D.Z. Skinner, D.R. See. Genetic diversity in historical and modern wheat varieties of the U.S. Pacific Northwest. *Crop Science* (2020). doi: 10.1002/csc2.20299.
- J.R. Stevenson, N. Villoria, D. Byerlee, T. Kelley, M. Maredia. Green Revolution research saved an estimated 18 to 27 million hectares from being brought into agricultural production. *PNAS* 110(21), 8363-8368 (2013).
- Thrupp, L.A. 2000. Linking agricultural biodiversity and food security: the valuable role of agrobiodiversity for sustainable agriculture. *International Affairs* 76(2):283–297.
- Tilman, D. 2001. Functional diversity. *Encyclopedia of Biodiversity* 3(1):109–120.
- Tsallis, C. 1988. Possible generalization of Boltzmann-Gibbs statistics. *Journal of Statistical Physics* 52(1–2):479–487.
- Tuomisto, H. 2012. An updated consumer's guide to evenness and related indices. *Oikos* 121(8):1203–1218.
- UN (United Nations). *Sustainable Development Report 2021*. Cambridge University Press.

- UNEP (United Nations Environment Programme). *Food Systems and Natural Resources. A Report of the Working Group on Food Systems of the International Resource Panel*. H. Westhoek, J. Ingram, S. Van Berkum, L. Özay, M. Hajer Paris: UNESCO, (2016).
- UN-SCBD (Secretariat of the Convention on Biological Diversity). *Global Biodiversity Outlook 5*. Montreal, 2020. Online at <https://www.cbd.int/gbo5>.
- USDA-ARS. GrainGenes, a database for Triticeae and Avena. Accessed in 2021 online at <https://wheat.pw.usda.gov/GG3/>
- USDA-NASS. 2021. “United States Department of Agriculture National Agricultural Statistics Service: Quick Stats.” Available at: <https://quickstats.nass.usda.gov/>.
- USDA-NGRP (National Genetic Resources Program). The Germplasm Resources Information Network (GRIN). Accessed 2021 at <https://www.ars-grin.gov/>
- USDA-PVPO (Plant Variety Protection Office). Plant Variety Database. Accessed in 2021 online at <https://www.ams.usda.gov/datasets/plant-variety>
- Vane-Wright, R.I., C.J. Humphries, and P.H. Williams. 1991. What to protect?—Systematics and the agony of choice. *Biological Conservation* 55(3):235–254.
- Waggoner, P.E. How much land can ten billion people spare for nature? *Daedalus* 125(3), 73-93, (1996).
- Warwick, R.M., and K.R. Clarke. 1995. New ‘biodiversity’ measures reveal a decrease in taxonomic distinctness with increasing stress. *Marine Ecology Progress Series* 129(1/3):301–305.
- Weitzman, M.L. 1992. “On Diversity.” *The Quarterly Journal of Economics* 107(2):363–405.
- Whittaker, R.H. 1960. Vegetation of the Siskiyou Mountains, Oregon and California. *Ecological Monographs* 30(3):279–338.

Supplementary Material

Table S1. Summary of selected prior wheat diversity studies

Studies	Number of varieties	Spatial and Temporal Extent	Taxonomy-informed Diversity Measures			Taxonomy-blind Diversity Measures	Spatial / Temporal Dynamic Measures
			Pedigree-based	Phenotypic traits	Molecular marker ¹		
Cox et al. (1986)	Between 60 and 262 depending on year	U.S. winter wheat states (1919-1984)	COP (area weighted)	- ²	-	-	Time trend of COP
Murhy et al. (1986)	110	U.S. winter wheat states	COP	-	-	-	-
Kim and Ward (1987)	22	North America	COP	-	RFLP	-	-
Brennan and Byerlee (1991)	not specified	Regions in Argentina, Australia, Brazil, Mexico, Netherlands, New Zealand Pakistan, U.S. (1970-1986)	-	-	-	Varietal area and age	Time trend of varietal area and age
Chen et al. (1994)	45	25 varieties from U.S. 20 varieties from China, Pakistan, India, Iraq, Hungary, Austria, Yugoslavia	-	-	STS	-	-
Barbosa-Neto et al. (1996)	112	U.S. midwest	COP	-	RFLP	-	-
van Beuningen and Busch (1997)	270	North America	COP	-	none	-	-

Barrett et al. (1998)	43	U.S. Pacific Northwest	COP	-	AFLP	-	-
Hartell et al. (2000)	18	Pakistan	COP	-	-	Varietal area and age	Time trend of varietal area and age
Corbellini et al. (2002)	40	Central and South Europe	COP	-	RFLP, AFLP	-	-
Almanza-Pinzón et al. (2003)	70	32 accessions from CIMMYT 38 accessions from 15 countries	COP	-	AFLP, SSR	-	-
Marić et al. (2004)	14	Croatia	COP	Yes ³	RAPD	-	-
Fufa et al. (2005)	30	U.S. Northern Great Plains	COP	Yes ³	SSR, SRAP	-	-
Chao et al. (2007)	43	U.S.	none	-	SSR	-	-
Fradgley et al. (2019)	454	U.K.	Kinship	-	SNP	-	-
Sthapit et al. (2020)	320	U.S. Pacific Northwest (1900-2019)	-	-	SNP	Varietal area	Time trend of varietal area

¹Abbreviations for different molecular markers: AFLP (amplified fragment length polymorphism); RAPD (random amplified polymorphic DNA); RFLP (restriction fragment length polymorphism); SNP (single nucleotide polymorphisms); SRAP (sequence related amplified polymorphism); SSR (simple sequence repeats); STS (sequence tagged sites).

²Dash (-) indicates such measures were not used.

³Marić et al. (2004) includes 12 morphological traits; Fufa et al. (2005) includes 5 morphological traits and 4 end-use quality traits

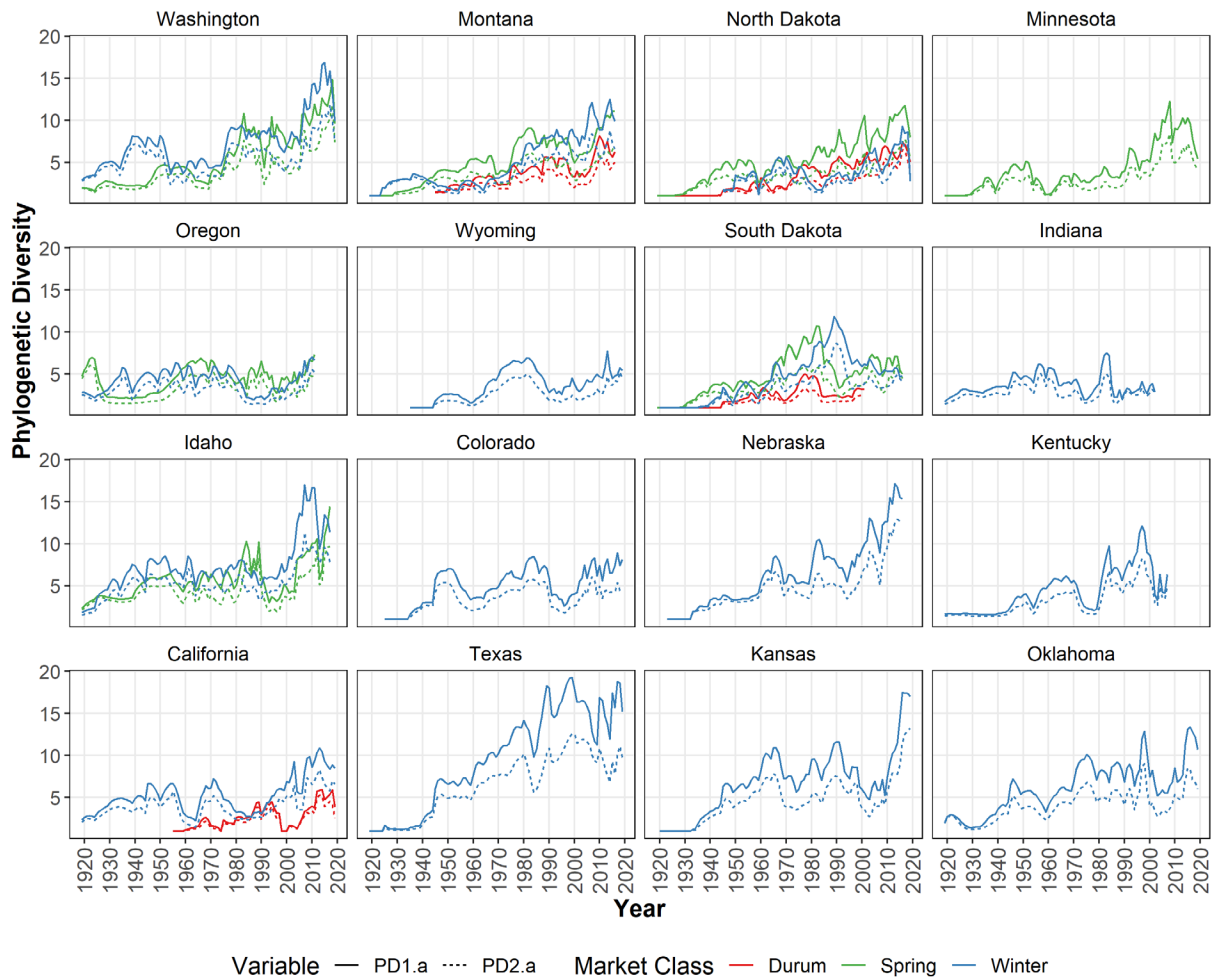


Fig S1. Comparing phylogenetic diversity indexes of order 1 and 2 for major U.S. wheat producing states, 1919-2019.