RESEARCH ARTICLE

The *Phaseolus vulgaris* L. Yellow Bean Collection: genetic diversity and characterization for cooking time

Rie Sadohara · Paulo Izquierdo · Filipe Couto Alves · Timothy Porch · James Beaver · Carlos A. Urrea · Karen Cichy

Received: 6 May 2021 / Accepted: 14 December 2021 / Published online: 24 January 2022

This is a U.S. government work and not under copyright protection in the U.S.; foreign copyright protection may apply 2022

Abstract Common bean (*Phaseolus vulgaris* L.) is a nutrient-rich food, but its long cooking times hinder its wider utilization. The Yellow Bean Collection (YBC) was assembled with 295 genotypes from global sources to assess the genetic and phenotypic diversity for end-use quality traits in yellow beans. The panel was genotyped with over 2,000 SNPs identified via Genotyping-By-Sequencing (GBS). Through population structure analyses with the GBS markers, the YBC was determined to be 69% Andean, 26% Middle American, and 5% admixture. The YBC was grown in two major bean production regions in the U.S.,

Michigan (MI) and Nebraska (NE) over two years. The genotypes exhibited a wide diversity in days to flower, seed weight, water uptake, and cooking time. The cooking times of the YBC ranged from 17–123 min. The cooking time were longer and varied more widely in NE with many more genotypes exhibiting hardshell than in MI. Fast-cooking genotypes were identified with various yellow colors; 20 genotypes cooked within 20 min in MI, and eight genotypes cooked within 31 min in NE. Water uptake and cooking time were significantly affected by the environment, which included both the growing and cooking environment, and notably in relation to cooking, NE is higher elevation than MI. SNPs associated with cooking time were identified with genome-wide association analyses and a polygalac-

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10722-021-01323-0>.

R. Sadohara · P. Izquierdo · K. Cichy (✉)
Department of Plant, Soil, and Microbial Sciences,
Michigan State University, East Lansing,
MI, USA
e-mail: karen.cichy@usda.gov

F. Couto Alves
Department of Epidemiology and Biostatistics, Michigan
State University, East Lansing, MI, USA

T. Porch
USDA-ARS, Tropical Agriculture Research Station,
Mayaguez, PR, USA

J. Beaver
Department of Agroenvironmental Sciences, University of
Puerto Rico, Mayaguez, PR, USA

C. A. Urrea
Panhandle Research & Extension Center, Scottsbluff,
NE, USA

K. Cichy
USDA-ARS, Sugarbeet and Bean Research Unit,
East Lansing, MI, USA

turonase gene on Pv04 was considered to be a candidate gene. The genotypic and phenotypic variability, fast-cooking genotypes, and the associated SNPs of the YBC will lay the foundation for utilizing yellow beans for breeding and genetic analyses.

Keywords *Phaseolus vulgaris* L. · Yellow beans · Cooking time · Diversity panel · GWA

Introduction

Common bean (*Phaseolus vulgaris* L., $2n = 2x = 22$) is a predominantly self-pollinating annual leguminous species with a relatively small genome size of 587 Mbp (Schmutz et al. 2014). Common bean is the world's most produced grain legume (Siddiq and Uebersax 2012) and serves as a staple food and a source of protein in many parts of Latin America and Africa (Broughton et al. 2003). Wild common bean is considered to originate in the Mesoamerica region and to have diverged into two distinct gene pools, Middle America in Mexico and the Andean in South America, where domestication occurred independently (Gepts et al. 1986; Mamidi et al. 2013; Bitocchi et al. 2013). Within each gene pool, beans can be further classified into races based on their characteristics including morphological traits and DNA marker patterns: Mesoamerica, Jalisco/Durango, Guatemala races in the Middle American gene pool and Chile, Peru, and Nueva Granada in the Andean gene pool (Singh et al. 1991; Beebe et al. 2000; Díaz and Blair 2006).

Beans are diverse in seed sizes, shapes, colors, and patterns. One interesting color class within *P. vulgaris*, is yellow. Yellow beans occur in many shades and sizes (Fig. 1) (Sadohara et al., 2021), and they make up important market classes in Latin America, the Caribbean, and Africa (Wortmann et al. 1998; Kilimo Trust 2012; Voysest 2012). The premier source of yellow colored beans is the Peruvian coast, where they have been grown since ancient times (Voysest 2012). Including some unique classes only found in Peru, there are at least a dozen market classes of yellow beans produced in Latin America. Yellow beans are also important in African countries including Angola, Tanzania, and Kenya (Buruchara et al. 2011). In fact, yellow beans fetch higher prices than other seed types in Zambia (Sichilima et al. 2016), are some of the most

traded types in Tanzania (Sperling et al., 2021), and they are the most preferred type in Uganda (Kilimo Trust 2012). A few of the most important yellow market classes include Manteca, Canary, and Mayocoba in Latin America, and Njano in Tanzania (Sones 2015).

Manteca is a market class with pale yellow seed coat and medium-sized seeds. Mantecas have been preferred by consumers in Chile because they are easier to digest and are more expensive due to this added benefit (Leakey 1992; Bassett 1999). The consumers' perception of low flatulence may be explained by the observations that Manteca beans are free of tannins that can decrease starch and protein digestibility (Beninger et al. 1998), and low in indigestible proteins and starch, and insoluble dietary fiber (Hosfield et al. 1998; Engleright et al. 1999; Hooper et al. 2016). Manteca beans have other advantages such as good taste and texture (Hosfield et al. 1998; Beninger et al. 1998). Moreover, some Manteca beans have been found to be fast-cooking and high in bioavailable iron (Cichy et al. 2015b; Wiesinger et al. 2018) due to their unique polyphenolic profiles with high levels of iron absorption promoters and low levels of inhibitors (Hart et al. 2020). Taken together, Manteca beans have offered multiple end-use and nutritional attributes that benefit consumers.

Peruvian Canary (Canario) has a bright yellow seed coat similar to a canary bird with no hilum ring color and has a long history of cultivation in Peru (Voysest 2012). It was introduced to Mexico in the 1970s and was given the name Peruano in Mexico to distinguish it from the Canary type that existed in Mexico, which has a duller yellow seed coat and a dark hilum ring and corona. Peruano was crossed with Mexican Canary to produce varieties that have the same seed coat color as the bright yellow Peruano. One of the resultant varieties, 'Azufrado Pimono 78', is also called Mayocoba, and Mayocoba became the name for the Peruano seed type in the U.S. (Bassett 2002; Voysest 2012). Mayocoba beans have been involved in a highly controversial U.S. patent granted to a Peruano type yellow bean called Enola in 1999, which claimed exclusive monopoly rights for the bright yellow seed color (Proctor 1999). DNA fingerprinting concluded that Enola was most likely a selection of yellow beans brought back from Mexico, involving no 'development' of the color, and genetically very similar to Azufrado Peruano 87 (Pallottini et al. 2004). The

Fig. 1 Examples of various shapes and colors of yellow beans

patent was cancelled after a protracted legal battle, but this incident discouraged U.S. bean breeders to invest as much effort in Mayocoba bean improvements as compared to other market classes. However, Mayocoba beans have been included in varietal trials for many years, and some Mayocoba-type varieties were recently released including Yellowstone and Patron (US Department of Agriculture, Agricultural Marketing Service 2019; Kelly et al. 2021).

Yellow beans are important market classes in Eastern and Southern Africa, where beans are produced and consumed as a dietary staple (Wortmann et al. 1998; Broughton et al. 2003). Yellow beans in Africa vary in shape, size, and color with regional preferences by consumers (Wortmann et al. 1998; Sichilima et al. 2016; Tumeo et al. 2017). The Pan-Africa Bean Research Alliance (PABRA) and the Alliance of Bioversity International and International Centre for Tropical Agriculture (ABC) surveyed yellow bean trading and market size in Eastern and Southern Africa regions including Tanzania, Burundi, DRC, Kenya, Rwanda, Uganda, and Zambia. The study revealed the growing popularity of yellow beans in the region, a vibrant informal seed system adept at incorporating new varieties, and that consumer preferences for yellow beans are especially associated with palatability and fast cooking times (Birachi et al., 2020; Sperling et al., 2021). Green-yellow shaded beans are referred to as Njano Kijani and Kijivu in Tanzania (Dr. Susan Nchimbi-Msolla, Sokoine University of Agriculture, Tanzania, personal communication), and Njano is a preferred type for home consumption in Tanzania (Sones 2015). Yellow, tan, and brown seed types are particularly popular

in Angola, where Manteiga is a popular seed class composed of several seed coat colors, ranging from cream to yellow, often grown or marketed in mixtures (Wortmann et al. 1998).

One of the quality characteristics associated with some market classes of yellow beans is fast cooking times (Mishili et al. 2011). Cooking time is the length of time beans take to be cooked and palatable. Beans typically take 30 min–2 h to be cooked, but beans with shorter cooking times are demanded by consumers and could lead to increased consumption in countries where beans are eaten less frequently (Brouwer et al. 1996; Smith et al. 2013; Wiesinger et al. 2018; Aseete et al. 2018). In order to improve any trait including cooking time through plant breeding, a wide genetic diversity and a high heritability for the trait are desirable.

A wide genetic variability in cooking time has been observed in common beans, and yellow beans were one of the fastest cooking types identified in an Andean bean diversity panel (ADP) (Cichy et al. 2015b). In addition, yellow beans are found in both of the two gene pools of common bean, Andean and Middle American (Voysest 2012), making inter- or intra- gene pool crosses an option for plant breeders. The heritability of cooking time is estimated to be moderate to high (Elia et al. 1997; Jacinto-hernandez et al. 2003; Garcia et al. 2012; Cichy et al. 2019; Katuuramu et al. 2020). The high genetic variability and heritability highlights the possibility of improving (shortening) cooking times of beans, and yellow beans will be a useful material to study and exploit for the fast-cooking trait. In addition to genetic architecture, the production environment can impact cooking time of beans (Cichy et al. 2019; Berry et al. 2020; Bassett

et al. 2021a). Therefore, it is necessary to assess the environmental effect and genotype $\times$ environment interaction on the cooking time performance of bean germplasm.

The objectives of this research were to 1) describe the genetic diversity of a Yellow Bean Collection (YBC) of 295 yellow beans (*P. vulgaris*) assembled with germplasm from the Americas, Europe, and Africa; 2) evaluate bean water uptake and cooking time when the YBC was grown and cooked in two major bean producing regions in the U.S.; and 3) investigate the genetic control and G $\times$ E effects of water uptake and cooking time.

Materials and methods

Plant material

A collection of 295 *P. vulgaris* L. accessions with various yellow seed colors were assembled from global sources (Table 1, Fig. S1). This diversity panel is called the Yellow Bean Collection (YBC). A set of 14 non-yellow genotypes and commercial varieties (white kidney, navy, cranberry, red mottled, light red kidney, and dark red kidney) were evaluated as controls for this panel: some with good agronomic adaptation to Michigan growing conditions (navy, dark red kidney, and light red kidney), and some previously characterized for cooking time G $\times$ E (Cichy et al. 2019). The characteristics of the control genotypes are shown in Table S1.

Field design

The YBC lines and 14 control varieties were grown at Michigan State University Montcalm Research Station (43°21.2'N, 85°10.6'W, 284 m elevation) in 2018 and 2019, and at the Scottsbluff (41°53.6'N, 103°40.7'W, 1200 m elevation) and Mitchell, Nebraska, Ag Labs (41°56.6'N, 103°41.9'W, 1240 m elevation), in 2018 and 2019, respectively. The local standard agricultural practices were followed in terms of fertilizer, herbicide, and pesticide applications as described by Kelly et al. (2018) and Kelly et al. (2019) in all the environments. Sets of 210–265 genotypes were planted and phenotyped for days to flower, 100-seed weight, water uptake, and cooking time (Table S2). Fewer genotypes were

Table 1 Origin of the Yellow Bean Collection entries

Region	Number of entries
Africa	89
Uganda	36
Tanzania	28
Angola collection	11
Kenya	5
Burundi	3
Other	3
Unknown	3
South America	51
Colombia	38
Ecuador	5
Brazil	3
Peru	3
Bolivia	1
Chile	1
Central America & the Caribbean	32
Haiti	27
University of Puerto Rico	3
Caribbean collection	1
Nicaragua	1
North America	83
Mexico	41
United States	35
Canada	7
Europe	35
Netherlands	11
Bulgaria	5
Hungary	6
Ukraine	3
Macedonia	2
Georgia	1
Other	5
Unknown	2
Middle East	3
Iran	2
Turkey	1
East Asia	2
India	2
Total	295

planted in Nebraska due to the initial shortage of seed of the 27 lines from Haiti. Each genotype was planted in a randomized complete block design with two field replications. At the Montcalm Research Station in

Michigan, each plot consisted of four 6.1-m rows with a 0.51-m row width. The central two and the outer rows of each plot were sown with seeds of the YBC genotypes and a dark red kidney bean, respectively, with sowing density of around 13 seeds per linear meter. Forty-five lines were sown at a reduced rate (25 seeds/row) in 2018 due to a shortage of seed. Rainfall was supplemented with an overhead irrigation applied 10 times throughout the growing season with 13.7 mm (5.4 inches) in total in 2018 and 12 times with 18.4 mm (7.25 inches) in total in 2019. At maturity, all bean plants in the middle 4.6-m section of the center two rows were hand-pulled, and the seeds were harvested by using a Hege plot thresher. Harvested seeds were dried at room temperature, cleaned to remove chaff, stones, and plant debris by using a seed cleaner (Clipper Office Tester, AT Ferrell Company Inc., Bluffton IN, USA), hand-cleaned to remove gravel and damaged or foreign seeds, and stored at a low-temperature and low-humidity storage chamber until water uptake and cooking time were measured. At the Scottsbluff and Mitchell Ag Labs, each plot consisted of two 3.7-m rows with 1.11-m row width, with a sowing density of around 16 seeds per linear meter. Rainfall was supplemented with an overhead and furrow irrigation applied 10 times throughout the growing season with 412.0 mm (16.24 inches) in total in 2018 and 10 times with 409.5 mm (16.12 inches) in total in 2019. At maturity, all bean plants in the center 3.0-m section of the two rows were hand-pulled, and the seeds were harvested by using an Almaco plot thresher. Harvested seeds were dried at room temperature, cleaned to remove chaff, stones, and plant debris by using a seed cleaner (Clipper Office Tester, AT Ferrell Company Inc., Bluffton IN, USA), hand-cleaned to remove gravel and damaged or foreign seeds, and stored at a room-temperature until water uptake and cooking time were measured.

Water uptake and cooking time measurement

An automated Mattson cooker system was used to measure cooking times according to the method described by Cichy et al. (2015b). Beans grown in MI were cooked at Michigan State University in East Lansing, MI and beans grown in NE were cooked at the Panhandle Research & Extension Center in Scottsbluff, NE. In MI seed moisture was equilibrated in a cold storage room at 4 °C and 75% relative humidity and

confirmed to be in the range of 10–14% before bean samples were subject to water uptake and cooking time measurements. Thirty seeds were soaked in distilled water for 12 h for genotypes grown in MI and 16 h for genotypes grown in NE, and the soaked weight was measured after blotting soaked seeds with a paper towel. The water uptake was calculated as water absorbed by the seeds (g) {dry seed weight before soaking (g)}⁻¹ × 100. Twenty-five soaked seeds were placed on an automated Mattson cooker apparatus (Wang and Daun 2005), cooked in boiling distilled water, and the 80% (20th) pin drop time was recorded as the cooking time. The cooking time of each field replication was measured. In Michigan, nine lines in 2018 and two lines in 2019 had only one field replication planted or harvested. For these lines, two samples from the harvested plot were taken as technical replications. The mean of the two field or technical replications were used as the cooking time for each environment and year.

Statistical analyses

The water uptake and cooking time data for MI and NE for 2018 and 2019 were transformed according to Box-Cox transformation (Box and Cox 1964) by using the MASS package (Venables and Ripley 2002) in R (R Core Team 2017). The statistical analyses of the phenotypes collected on all trials were performed with the following mixed liner model: $Y_{ijkl} = \mu + G_i + E_j + Y_k + EY_{jk} + GE_{ij} + GY_{ik} + GEY_{ijk} + B(EY)_{jkl} + \varepsilon_{ijkl}$, where: Y_{ijkl} is the phenotypic value of the i th YBC genotype grown in the l th block of the j th environment in the k th year; μ is the grand mean; and G_i is a random effect of the i th genotype. E_j and Y_k are fixed effects of the j th environment and the k th year, respectively; EY_{jk} , GE_{ij} , and GY_{ik} are two-way interaction terms; GEY_{ijk} is a three-way interaction term; $B(EY)_{jkl}$ is a random effect of the l th block nested in the j th environment and in the k th year; and ε_{ijkl} is an error term. The variance components of the model were used to calculate the broad-sense heritability of the traits (Fehr 1987). The variance components were also used to partition the interaction terms into two parts, V and L. V represents variation due to heterogeneity of genotypes among various environments, and L represents the lack of correlation between genotypes among various environments (Dickerson 1962; Cooper and DeLacy 1994). If V accounts for a large portion of

genotype $\times$ environment, selection would be more straightforward, and genotypes with the highest mean could be selected. If L is the more important portion of genotype $\times$ environment effect, the interaction would complicate selection due to changes in rank of genotypes in various environments. Only significant interactions of interest (i.e. genotype $\times$ environment and genotype $\times$ year) were examined for each trait.

The broad-sense heritability was also estimated within each environment by using the following mixed linear model: $Y_{ijk} = \mu + G_i + Y_j + GY_{ij} + B(Y)_{jk} + \varepsilon_{ijk}$, where Y_{ijk} is the phenotypic value of the i th genotype grown in j th year in the k th block; μ is the grand mean; G_i is the random effect of the i th genotype; Y_j is the fixed effect of the j th year; GY_{ij} is the interaction term of the i th genotype and the j th year; $B(Y)_{jk}$ is the random effect of the k th block nested in the j th year; and ε_{ijk} is the error term. The Best Linear Unbiased Estimators (BLUE) were calculated for the 260 and 217 phenotyped individuals for cooking time grown in MI and NE, respectively, by setting the genotype as a fixed effect and all other terms as random effects using the emmeans package (Lenth 2021) in R for genome-wide association (GWA). All models were fitted using the lme4 (Bates et al. 2015) and the lmerTest (Kuznetsova et al. 2017) packages in R.

Genotyping of the YBC accessions and genetic identity controls

Five seeds of each YBC line were planted in the Michigan State University research greenhouse. Five trifoliate leaves per line were collected, frozen at -80°C , lyophilized, and milled by using a benchtop mill (Geno Grinder 2000, Spex Certiprep, Metuchen, NJ, USA). DNA was extracted from the milled leaves using NucleoSpin Plant II Kit (Macherey–Nagel, Duren, Germany) following the ‘Genomic DNA from plant’ protocol. DNA was eluted in 50 μL of PE buffer. The purity of DNA was checked by running random samples of DNA on 1% agarose gels visualized using ethidium bromide stain. DNA concentration was measured by using Quant-iTTM PicoGreenTM dsDNA Assay Kit (Invitrogen, Waltham, MA, USA), and 10 ng/ μL of DNA was used for Genotyping-By-Sequencing (GBS) library preparation according to Elshire et al. (2011) with a single restriction enzyme, ApeKI. Barcoded DNA from 96 lines were pooled into

one sequencing sample. Of the 96 adapter sequences, 11 were modified to increase specificity according to Thomas P. van Gurp (<http://www.deenabio.com/services/gbs-adapters>, accessed on April 20, 2018). The libraries were cleaned by using NucleoSpin Gel and PCR Clean-up kit (Macherey–Nagel, Duren, Germany) and sequenced with an Illumina standard HiSeq 4000, generating 150 bp single-end reads at RTSF Genomics Core at Michigan State University. Cutadapt ver. 1.16 was used to remove the adaptor sequence ‘GAGATCGGAAGAG’ with a minimum base quality score of 20, a minimum read length of 36, a maximum adapter sequence repeat of 3 times, and with the ‘remove flanking N bases from each read’ option. FastQC ver.0.11.7 was used to confirm that there was no overrepresented sequences or adapter sequences in the libraries. An additional set of four control varieties with known genepool origins were sequenced: 1) Puebla 152 (black, MA race Mesoamerica), 2) Stampede (pinto, MA race Durango), 3) Montcalm, and 4) Talon (dark red kidneys, Andean race Nueva Granada) by using the same method described above except that 50 bp single-end reads were generated.

SNP calling

SNP calling with the 18 control varieties and 295 YBC lines was carried out by using NGSEP software (Duitama et al. 2014). The libraries were de-multiplexed, and the reads of each line were aligned to the *P. vulgaris* genome ver.2.0 (Phytozome, <https://phytozome.jgi.doe.gov/>). Approximately 2 million SNPs and in-dels were called among the 295 YBC lines and 18 control varieties. SNPs were removed if they contained heterozygous calls, were non-biallelic, had a quality score < 40 , were on scaffolds, or were on the predicted repetitive regions of the common bean genome (BeanWGS_RefV2_CIAT_UCDA_VIS_repeatsNGSEPMergedMasked.txt, downloaded on December 18, 2018 from <https://datadryad.org/stash/dataset/doi:10.5061/dryad.46pk7> (Lobaton et al. 2018)), after which 417,142 SNPs were retained. Heterozygous loci were removed because they are potentially genotyping errors as common bean is a self-pollinating species and is expected to have almost all homozygous genotypes. These 417,142-SNPs were the base SNP set, and it was filtered in different ways for the following uses.

Population diversity analyses

The base SNP set was further filtered for a minimum minor allele frequency (MAF) of 5% and < 20% missing genotyping data, which resulted in 2,234 SNPs. This SNP set was used for a principal component analysis (PCA) using PLINK ver. 1.9 (Purcell et al. 2007). The first 10 principal components (PCs) were calculated, and the ggplot2 package (Wickham 2016) in R (R Core Team 2017) was used to visualize the diversity of the panel and the control varieties.

The base SNP set was filtered to remove SNPs in linkage disequilibrium (LD) to comply with the assumption by the STRUCTURE program that markers are at LD within subpopulations (Pritchard et al. 2000). Pairwise LD was calculated with a window size of 50 kb and a shift size of 5, and SNPs with $R^2 > 0.8$ were removed using PLINK ver. 1.9 (Purcell et al. 2007). The resulting 862 SNPs were used to conduct population structure analysis by STRUCTURE ver. 2.3.4 using the admixture model with independent allele frequencies and no prior population information for 10,000 burn-in periods followed by 50,000 Markov chain Monte Carlo (MCMC) replications (Pritchard et al. 2000). The first significant population split was determined by using delta K, a value related to the second order rate of change in the likelihood of the number of clusters (Evanno et al. 2005). Structure Harvester (Earl and VonHoldt 2012) was used to determine delta K from the STRUCTURE output, which returned $K = 2$ as the strongest population division. Common bean diversity studies have shown that two distinct genepools are present in common bean: Andean and Middle American (Gepts and Bliss 1985; Gepts et al. 1986; Koenig and Gepts 1989). The control lines of known genepool origin were specified in the second run of STRUCTURE analysis with subpopulation numbers (K) from 2 to 4 for 10,000 burn-ins and 50,000 MCMC replications using prior genepool information of the 15 control varieties to assign individuals to clusters (Pritchard et al. 2009).

The base SNP set was filtered by using SNPhylo ver. 20140701 (Lee et al. 2014) with the default filtering conditions to remove SNPs with > 20% missing data, < 5% MAF, and/or pairwise LD of $R^2 > 0.8$ with a sliding window size of 500 kb. As a result, 965 SNPs were retained and used to construct a phylogenetic tree. The phylogenetic tree generated by SNPhylo was visualized by using FigTree ver. 1.4.3

(<http://tree.bio.ed.ac.uk/software/figtree/>, downloaded on December 16, 2020).

Genome-wide association study

The base 417,142-SNP set was filtered to include the 260 and 217 YBC lines grown and phenotyped for water uptake and cooking time in MI and NE, respectively, and to exclude SNPs with > 20% missing data or < 5% MAF. As a result, 2,285 and 2,317 SNPs for MI and NE, respectively, were selected for GWAS. GWAS for cooking time in MI and NE were conducted by using the GAPIT package in R (Lipka et al. 2012) using the least squares means, which were adjusted for the year effects. Bayesian-information and linkage-disequilibrium iteratively nested keyway (BLINK) model was then used to detect associations between the SNPs and the traits (Huang et al. 2019). BLINK runs two fixed effect models: the first one tests associations between markers and trait values with pseudo quantitative trait nucleotides (previously associated SNPs with the trait; QTNs) as covariates; and the second one optimizes the selection of QTNs to be included as covariates in the first model using Bayesian Information Criteria. BLINK has higher statistical power than its preceding model, Farm-CPU. A false discovery rate at $\alpha = 0.05$ was used to call associations significant. The CMplot package (Yin 2020) in R was used to generate QQ plots and Manhattan plots.

Results

Genetic diversity

The YBC is comprised of 295 *P. vulgaris* accessions with yellow-colored seeds primarily represented by major bean production regions of the world including Africa (89 accessions), followed by North America (83), South America (51), Europe (35), Central America and the Caribbean (32), and Middle East and East Asia (5) (Table 1). The color types of the YBC accessions included Amarillo dark and light, beige, brown, Canary (Mexican), Manteca, Mayocoba, and green-yellow (Njano) (Fig. 1).

The genetic diversity of the YBC was evaluated through PCA, analysis with STRUCTURE, and the generation of a phylogenetic tree. With PCA, most of

Fig. 2 Biplots of principal component analysis of the YBC. **A:** PC1 and PC2 by the region of origin. The numbers in the legend indicate the number of accessions originated from each region.

Arrows indicate varieties with a known genepool or race origin; **B:** PC1 and PC3 by the region of origin

the variance was explained by the first three PCs, with PC1 explaining 63.8%, PC2 explaining 16.1% and PC3 explaining 4.5% of the variation (**Fig. S2**). PC1 separated the Andean and Middle American YBC lines, and PC2 separated the Middle Americans into two groups: defined by Mesoamerica and Jalisco-Durango races (**Fig. 2A**). The accessions scattered in the middle of the biplot were considered as admixtures between genepools. Many of the accessions in the Jalisco-Durango cluster are from North America, specifically Mexico, while the accessions in the Mesoamerica group are from Central America and

the Caribbean, mainly Haiti. PC3 separated Andeans into subgroups; The Andean lines from South America had higher values of PC3 while those from Europe had lower values of PC3 (**Fig. 2B**).

The STRUCTURE analysis at $K = 2$ classified the YBC lines and the non-yellow 18 control genotypes into two clusters: Andean and Middle American similar to PC1 (**Fig. 3A**). The YBC lines were classified as Andeans if the proportion of K1 was 70% or higher, and as Middle Americans if the proportion of K2 was 70% or higher. The lines that had less than 70% of both K1 and K2 were classified as

admixtures. With the 70% cutoff, 203 YBC lines were classified as Andeans, 16 as admixtures, and 76 as Middle Americans. The membership proportions of the YBC in the order of their geographical origin are shown in Fig. 3B. $K = 2$ shows that the accessions from Mexico contain both Andean and Middle American genotypes. At $K = 3$, the Andeans were separated into two subgroups; many of the U.S., Canadian, and CIAT lines from South America had high $K3$ membership. At $K = 4$, both Andeans and Middle Americans were divided into two subgroups as expected from the results of $K = 2$ and $K = 3$.

The phylogenetic tree generated by SNPhylo is shown in Fig. S3. As observed with PCA and STRUCTURE analyses, the YBCs were separated into two major groups: Andean and Middle American, and the lines that were considered as admixtures according to the population structure analysis at $K = 2$ were placed between the two genepool clusters or with a longer distance from the Andean or Middle American clusters. Andeans formed a denser cluster than Middle Americans, which separated into

Mesoamerica (leftmost) and Jalisco-Durango groups. In summary, the PCA, STRUCTURE, and SNPhylo analyses all agreed that the YBC contains Andean, Middle American, and admixed accessions.

Phenotypic diversity

The YBC was evaluated for plant growth habit (MI 2018), days to flower, seed size, water uptake, and cooking time (MI and NE in 2018 and 2019). The plant growth habit of the YBC by genepool is shown in Table S3. The Andean genotypes were 63% determinate and 37% indeterminate, whereas most of the admixtures and Middle American lines were indeterminate.

Days to flower

The YBC lines in each genepool exhibited variability for days to flower. Andeans had a higher proportion of genotypes that flowered early (about 37 days) than Middle Americans and admixture genotypes in all the

Fig. 3 Population membership of the YBC lines and the controls determined by STRUCTURE. A: Classification by the K1 membership ($K = 2$); B: Classification by the geographical and institutional origin ($K = 2-4$)

Table 2 The mean and the range of days to flower, 100-seed weight, water uptake and cooking times of the YBC grown in MI and NE in 2018 and 2019

	Days to flower (day)	100-seed weight (g)	Water uptake (%)	Cooking time (min)
Mean	46.0	35.1	85	40
SD	8.7	7.9	13	14
Range	34.8–77.0	13.6–57.4	41–116	17–123
Fold-difference	2.2	4.2	2.8	7.2

environments and years (Fig. S4). Out of the 58 Middle American genotypes, 40 took 50 days or more to flower in MI 2018, but when almost the same set of genotypes were grown in MI 2019, the most frequently observed days to flower of the Middle Americans was between 45–50. Many of the Jalisco-Durango lines from Mexico were unadapted in Michigan and Nebraska (i.e., did not flower or flowered/matured too late) due to photoperiod sensitivity. Days to flower measured in 2018 and 2019 were highly correlated within each environment (Fig. S5).

100-seed weight

On average, the Andean YBC lines had the largest seed weights, the admixtures were intermeditate, and the Middle American genotypes were the lowest. This trend was observed in both environments and across years (Fig. S4). Andeans and admixed genotypes produced smaller seeds in NE than in MI in both years, while Middle American genotypes produced seeds of similar size in all the environments. There should not

have been drought stress since the plots were irrigated in NE; however, the Andeans and admixed genotypes might have been heat-stressed, flowered early, and produced smaller seeds under the NE environment. The numbers of days with temperatures above 30 °C were 35 in 2018 and 32 in 2019. Temperatures above 30 °C cause excessive flower drop and abortion of fertilized ovules (Rainey and Griffiths 2005). The seed weights measured in 2018 and 2019 were highly correlated within each environment (Fig. S6).

Water uptake and cooking time

The YBC seeds exhibited a large variability in water uptake and cooking time, ranging from 41 to 116% and from 17 to 123 min, respectively (Table 2). Beans typically double in weight after being fully hydrated, which is equal to 100% water uptake. The YBC lines grown in MI showed normal distributions around 90–100% for water uptake in both years with a small number of lowly hydrating lines (Fig. 4A). In contrast, the YBC grown in NE showed much flatter

Fig. 4 Phenotypic distribution of **A**: mean water uptake and **B**: mean cooking time of the YBC grown in MI and NE in 2018 and 2019. The black bars represent the median in each environment and year

Fig. 5 Density plots of cooking time by genepool of the YBC measured in MI and NE in 2018–2019. MA: Middle American

distributions with the medians lower than those grown in MI. The distributions of cooking time were similar to those of water uptake except that the cooking time distribution for NE 2018 was similar to a normal distribution but skewed towards longer cooking times (Fig. 4B). In NE, the lines that had low water uptake had longer cooking times, whereas in MI, there was not a strong relationship between water uptake and cooking time (Fig. S7). Despite the difference between the environments, the 2-year 2-environment mean of water uptake and cooking time were correlated (Fig. S8).

The cooking times measured in each environment were moderately consistent across the two years, although genotypes grown in NE showed higher year-to-year variability (Fig. S9). The cooking times of the genotypes grown in MI and NE were lowly correlated for either 2018 or 2019 (Fig. S10). The cooking time density plots by genepool is shown in Fig. 5. In MI, there was no obvious difference by genepool in the cooking time distributions, but the Middle Americans had a peak at a slightly shorter time in 2019. In NE, Andeans tended to be the faster cooking than admixtures and Middle Americans. There was no large difference in cooking time between seed size categories in MI for both years, whereas small-seeded

Table 3 The *p*-values of the factor effects and broad-sense heritability for water uptake and cooking time

Factor	<i>p</i> -value	
	Water uptake	Cooking time
Genotype (G)	0.12	0.0043
Year (Y)	< 0.0001	< 0.0001
Environment (E)	< 0.0001	< 0.0001
G × E	< 0.0001	< 0.0001
V ^a (%)	47	1
L ^b (%)	53	99
G × Y	1.0	1.0
E × Y	< 0.0001	< 0.0001
G × E × Y	< 0.0001	< 0.0001
H ² ^c - overall	0.18	0.34
H ² —MI	0.76	0.80
H ² – NE	0.77	0.63

^aV: G×E component due to heterogeneity of genotypes among environments

^bL: G×E component due to lack of correlation among genotypes among environments

^cH²: Broad-sense heritability

beans took longer to cook in NE, especially in 2019 (Fig. S11), and this is considered as a result of the lower water uptake by small-seeded beans, mainly for

Table 4 The five fastest cooking YBC lines phenotyped in MI and NE. The cooking times and 100 seed weights are the means of the two years in each environment

MI					
ID	YBC231	YBC281	YBC297	YBC192	YBC301
Genotype name	CAN 44	SSIN 533	F7500	Roba-1	M0181
Cooking time (min)	15	16	17	17	19
Part of ADP ^a	No	No	No	No	No
Country of Origin	Colombia	Uganda	Haiti	Uganda	Haiti
Source	CIAT	CIAT-Uganda	Farmer	CIAT-Uganda	Market
Genepool ^b	Andean	MA	MA	MA	MA
100-seed weight (g)	27.1	29.4	22.0	21.8	23.7

NE					
ID	YBC231	YBC071	YBC239	YBC129	YBC214
Genotype name	CAN 44	PI433611	CAN 79	Ervilha	NE-17-7
Cooking time (min)	23	24	27	30	31
Part of ADP ^a	No	No	No	Yes (ADP0512)	No
Country of Origin	Colombia	United States	Colombia	Angola	United States
Source	CIAT	US GRIN	CIAT	Angola collection	University of Nebraska
Genepool ^b	Andean	Andean	Andean	Andean	Andean
100-seed weight (g)	19.8	29.1	37.7	45.3	35.4

^aADP: Andean Diversity Panel; ^bGenepool is based on the STRUCTURE classification at K = 2; ^cMA: Middle American

the Middle Americans (Fig. S12). The genotype effects were only significant for cooking time while the year, environment, genotype $\times$ environment, environment $\times$ year, and genotype $\times$ environment $\times$ year were all significant for water uptake and cooking time (Table 3). The lack of significance for the genotype effect for water uptake may be related with the flat distribution of the phenotypes observed in NE in both years (Fig. 4A). For water uptake, the genotype $\times$ environment was partitioned into 47% V and 53% L. For cooking time, genotype $\times$ environment was partitioned into 1% V and 99% L (Table 3).

The higher L for cooking time corresponded with the low correlation of cooking time among genotypes between the two environments (Fig. S10). The broad sense heritability estimate of cooking time was low (0.34) when both environments were combined but was higher when estimated for each environment (0.80 for MI and 0.63 for NE, Table 3).

Table 4 shows the five fastest cooking beans grown in MI and NE. Four of the five fastest cooking beans in MI were classified as Middle Americans, and none of them were part of the Andean Diversity Panel. They were small-seeded, and their seed colors varied from

Fig. 6 Biplot of principal component analysis based on phenotype data of the YBC (mean of 2 environments and 2 years of days to flower, 100-seed weight, water uptake, and cooking time)

beige to dark orange. All the five fastest cooking beans in NE were Andean and had similar seed coat colors of lemon- yellow or yellow but differed in seed size ranging from 19.8 to 45.3 g per 100 seeds. YBC231 was the fastest cooking in both environments; moreover, the other four fastest cooking beans in NE were also fast cooking (28 min or less) in MI.

PCA was conducted with the four phenotypic values: days to flower, 100-seed weight, water uptake, and cooking time. PC1 and PC2 explained 50% and 30% of the variation, respectively. PC1 separated genotypes by water uptake and somewhat by cooking time, and PC2 separated genotypes by days to flower and seed weight (Fig. 6). Central American and Caribbean materials were short-cooking, whilst U.S. and Canadian materials were longer-cooking. In contrast, genotypes were clustered primarily by their gene pools based on SNPs (Fig. 2A).

GWAS for water uptake and cooking time

The QQ plots and Manhattan plots generated from the GWAS analyses for water uptake and cooking time are shown in Fig. 7. For water uptake, significant SNPs were identified on Pv01, Pv02, Pv04, Pv08, and Pv09 (Table 5). For cooking time, significant SNPs were identified on Pv01, Pv03, Pv04, Pv06, Pv08, and Pv10 (Table 5). The phenotypic effect was generally larger

for SNPs detected in NE than in MI that corresponds to the wider distribution of both water uptake and cooking time in NE (Fig. 4A, B). There was no overlap of significant SNPs between MI and NE for either water uptake or cooking time, thus associations were environment dependent. However, some SNPs overlapped for both traits within each environment: the two significant SNPs detected on Pv08 for water uptake and cooking time in MI were 677 kb apart; and the SNP Chr04pos4548944 identified in NE was significant for both water uptake and cooking time.

The major allele groups had shorter cooking times than the minor allele groups for all the significant SNPs in NE. For MI, the results were mixed: two out of four significant SNPs had the major allele group having shorter cooking times (Table 5). These findings lend some degree of support to our original idea that yellow beans are a source of fast cooking time, with the growing and cooking environments playing an important part. The cooking time ranges of the five fast-cooking control varieties were 18–29 min in MI and 26–83 min in NE, and those of the three slow-cooking control varieties were 22–41 min in MI and 34–117 min in NE. The cooking time ranges of the YBC genotypes were larger than those of the control varieties, and some YBC genotypes cooked faster than the fast-cooking control varieties in each combination

Water uptake

Cooking time

Fig. 7 Genome-wide association analysis of water uptake and cooking time in MI and in NE. The QQ plots on the left show the model fit, and the Manhattan plots show the p -values of SNPs for associations. The gray dotted lines indicate the false discovery rate-adjusted threshold for p -values ($\alpha = 0.05$). CT.6.1 and CT.8.2 are QTL reported by Bassett et al. (2021b).

CT.10.1 is a QTL reported by Berry et al. (2020). S08_62659170 is a SNP significantly associated with cooking time (604 kb apart from the significant SNP in this study) (Bassett et al. 2021a). The significant SNP on Pv04 detected in NE is in the coding region of a polygalacturonase gene (Phvul.004G038700).

Table 5 QTL identified by the genome-wide association analysis for water uptake and cooking time in environments MI and NE

SNP	Chr	Position	FDR-adjusted <i>p</i> -values	Major allele			Minor allele			MAF	SNP effect ^a
				Allele	n	Mean phenotypic value	Allele	n	Mean phenotypic value		
<i>Water uptake (%)</i>											
MI											
Chr02pos24852892	2	24,852,892	1.4E-04	AA	225	95.2	CC	19	97.7	0.08	− 2.5
Chr04pos25006646	4	25,006,646	2.5E-07	TT	156	97.5	AA	62	91.3	0.28	6.1
Chr08pos1429025	8	1,429,025	3.8E-12	AA	236	96.4	GG	14	82.1	0.06	14.3
Chr08pos62731875	8	62,731,875	3.4E-07	AA	160	97.1	GG	75	91.4	0.32	5.7
NE											
Chr01pos45386764	1	45,386,764	4.5E-03	AA	166	75.5	GG	22	50.6	0.12	24.9
Chr02pos49223533	2	49,223,533	3.1E-02	GG	157	67.7	TT	28	84.8	0.15	− 17.1
Chr04pos4548944	4	4,548,944	4.5E-03	CC	141	77.2	AA	37	55.5	0.21	21.7
Chr09pos35530354	9	35,530,354	2.4E-03	GG	194	68.5	AA	19	96.2	0.09	− 27.7
<i>Cooking time (minutes)</i>											
MI											
Chr03pos3019420	3	3,019,420	1.2E-03	GG	190	26.0	AA	34	30.5	0.15	− 4.5
Chr06pos27715283	6	27,715,283	1.1E-03	GG	222	27.0	AA	14	22.6	0.06	4.4
Chr08pos62054742	8	62,054,742	3.1E-02	GG	152	25.7	CC	68	28.7	0.31	− 3.0
Chr10pos43615359	10	43,615,359	2.1E-10	CC	224	27.2	GG	25	21.8	0.10	5.5
NE											
Chr01pos1528388	1	1,528,388	2.8E-02	CC	178	51.8	AA	25	66.7	0.12	− 14.9
Chr03pos50288381	3	50,288,381	4.9E-02	AA	181	49.7	CC	12	87.3	0.06	− 37.6
Chr04pos4548944	4	4,548,944	3.5E-05	CC	141	47.8	AA	37	69.5	0.21	− 21.7
Chr08pos511826	8	511,826	1.3E-04	AA	180	49.6	TT	19	83.4	0.10	− 33.8
Chr10pos9440210	10	9,440,210	1.4E-05	AA	151	48.8	TT	24	78.3	0.14	− 29.5

^aSNP effect is the difference between the mean cooking times of the YBC lines with the major and minor alleles

of environment and year, indicating a rich genetic diversity in cooking time of the YBC.

Discussion

There have been interesting findings on yellow beans for consumer-focused traits such as fast cooking times and high iron bioavailability (Cichy et al. 2015b; Wiesinger et al. 2016, 2018). Given the diversity in shape, size, and colors within yellow beans and their importance in markets across the world, the YBC was assembled to facilitate the improvement and use of yellow beans. Genotypic and phenotypic characterization of the YBC for important traits will help assess

the potential of the panel as a resource for breeding and genetic studies.

The YBC lines were genotyped by using the GBS technology, and the genetic diversity was assessed using PCA, STRUCTURE, and phylogenetic analyses. All the analyses indicated that the YBC lines can be grouped as Andean, Middle American, or admixture lines. The first major split of the panel separated Andeans and Middle Americans and was consistent with previous studies (Díaz and Blair 2006; Cichy et al. 2015a; Oladzad et al. 2019). The existence of yellow beans in both genepools is advantageous in bringing in beneficial traits from yellow beans to other market classes because of the difficulty in inter-

genepool hybridization in common bean (Gepts and Bliss 1985).

There were 62 genotypes in the YBC classified as Mayocoba type, 32 of which were from South America. Among the 32 genotypes from South America, 27 were from CIAT, Colombia. Most of the accessions originating from South America were in the Andean cluster, which is expected considering that the Andean gene pool includes accessions from this region. The Mayocoba type also included 19 U.S. breeding lines and one variety (YBC116, Myasi). Our STRUCTURE analysis showed that one of the Andean subgroups at $K = 3$ consisted of genotypes from U.S. breeding lines and CIAT lines (Fig. 3B). K3 in this case could be separating race Peru from other Andean races because many of these lines were collected in Peru, Ecuador, Chile, or Colombia, where race Peru beans are known to exist, especially as yellow beans.

The Mesoamerica group contained many lines from Central America and the Caribbean (Fig. 2A), and 27 of them were from Haiti. Haitian growers seemed to prefer Middle American landraces because of their high agronomic yield (Mainviel 2019). Black beans are the most important seed type for Haitian consumers, but yellow beans are also consumed (Beaver et al. 2012). To our knowledge, this is the first study to report the genotypic and phenotypic diversity of these yellow beans from Haiti. Half the Mexican accessions were Middle Americans, and the other half were Andeans when classified using STRUCTURE (Fig. 3B). This highlights the diversity of yellow beans present in Mexico: Mexico is the center of origin for Middle American germplasm, but local consumers accepted and grow Andean yellow beans that must have been introduced from South America (Voysest 2012).

The YBC accessions from Africa were found in all the three clusters in the PCA, indicating a wide genetic diversity of African germplasm (Fig. 2A). Similarly, whole-genome sequencing of selected beans and related species found that African accessions distributed into various races and gene pool groups (Lobaton et al. 2018). The diversity of African germplasm likely resulted from the importation of germplasm developed in other parts of the world and the relatively lenient market class specification requirements (Lobaton et al. 2018).

The YBC showed a 7.2-fold variation in cooking time (Table 2), and the differences were more

pronounced in NE (Table S4, Fig. 4). The fast-cooking beans in MI were not necessarily fast-cooking in NE in this study (Fig. S10), indicating strong environmental effects. Some studies reported limited environmental effects on cooking time (Cichy et al. 2019; Katuuramu et al. 2020), while a two-year study using a biparental RIL population grown in contrasting environments showed a large environmental effect on cooking time (Berry et al. 2020). Likewise, the strong environmental effect found in this study is probably due to the two contrasting environments: MI with high humidity and NE with low humidity. Indeed, the heritability estimates for those traits were higher when estimated for each environment (Table 3) and were consistent with previous studies which reported moderate to high heritability (Elia et al. 1997; Cichy et al. 2019; Katuuramu et al. 2020; Bassett et al. 2021a). In addition, the beans were cooked in the environment rather than being shipped to MI unlike some of the previous studies (Cichy et al. 2019; Berry et al. 2020), and the environmental differences in handling and storage conditions may have been a part of the environmental effect. The NE production environment has the tendency to induce hardshell more so than the MI environment. Hardshell occurs in environments where the weather is dry during seed filling (Castellanos et al. 1995). Seeds with hardshell do not absorb water during soaking and take longer times to cook, which is a hydrothermal process. Hardshell can increase variability within genotypes because it does not necessarily impact all seeds of a genotype but just a few. The NE cooking location was also 1000 masl higher than the MI cooking location. Bean cooking times tend to increase incrementally with incremental increases in elevation (Bressani and Chon 1996). Overall, these two environments, MI and NE, are valuable for testing the effect of humid and arid growing conditions and elevation on cooking time.

For water uptake, over half of the genotype $\times$ environment was accounted for by L, the lack of correlation between genotypes in different environments. For cooking time, the L was essentially the sole source of genotype $\times$ environment interaction (Table 3), which is in accordance with the low correlation between the environments (Fig. S10). A high L component impedes gain from selection due to changes in rank of genotypes among environments; therefore, selection in specific environments or testing

in a large number of environments may be necessary to select high-performing lines for these traits. However, this estimate is based on only two environments; a larger number of environments should be used in order to estimate more generalizable V and L. Previous studies with large numbers of environments (9–15) have concluded that the genotype $\times$ environment interaction effect on water uptake and cooking time is small and therefore a small number of environments are needed for evaluation of those traits (Cichy et al. 2019; Katuuramu et al. 2020). It would be more practical for breeding programs to estimate V and L with this type of multi-environment study.

This study identified fast cooking yellow beans including Manteca and other types that cooked in 15–31 min in MI and NE (Table 4). In addition, 15 more YBC lines grown in MI cooked within 20 min. Several YBC lines including YBC231 showed constantly short cooking times in both environments; they will be valuable genetic resources for developing fast-cooking bean varieties with stable performance in different environments. The cooking times of promising lines should be evaluated in the target production environments because YBC281 (SSIN 533 from CIAT-Uganda) and YBC192 (Roba-1, CIAT-Uganda) were two of the five fastest cooking in MI but took 80 and 97 min, respectively, in NE.

Some Manteca beans were reported to be fast-cooking (Wiesinger et al. 2016, 2018; Cichy et al. 2019). Tannin content and cooking times were positively correlated (Elia et al. 1997), and the lack of polyphenols such as proanthocyanidins (condensed tannins) in Manteca beans (Beninger et al. 1998) may be one of the reasons for fast cooking times of Manteca beans. It is interesting, however, that the YBC contained some fast-cooking beans that are not Manteca type. The YBC lines were held in ambient conditions and cooked relatively soon after harvest. Thus, tannins present in the seeds may not have impacted cooking times as much as they may have if beans were stored longer and under unfavorable conditions, which can induce tannin polymerization leading to long cooking times (Stanley 1992; Martín-Cabrejas et al. 1997). In addition, there seems to be more mechanisms for the short cooking trait as varying polyphenol profiles have been reported in some yellow beans (Hart et al. 2020). Beans in MI were stored in cold storage until cooking time measurement, and this may be another reason for the

differences in the top five fastest cooking beans in MI and NE. The cooking times in MI were likely the shortest possible cooking time of the genotypes because the seeds were stored in cold storage, cooking time was measured within 4–6 months after harvest, and the moisture was equilibrated to 10–14% before soaking. Consumers are expected to experience longer cooking times than the potential of genotypes because beans are typically stored under room temperature in the supply chain and could be 1–2 years old before they are purchased and cooked. More attention should be paid to the age and storage conditions of beans in order to deliver the benefit of short-cooking beans to consumers.

All the significant SNPs for water uptake were novel when compared with previous studies using markers with known physical positions (Cichy et al. 2015b; Berry et al. 2020; Bassett et al. 2021a, b; Diaz et al. 2021; Soltani et al. 2021). Several SNPs on Pv02 and Pv04 were significant in this study, and QTLs were detected on the same chromosomes in some previous studies with no physical position available (Pérez-Vega et al. 2010; Sandhu et al. 2018). For one of the significant SNPs for water uptake on Pv08 in MI, Chr08pos1429025, the major allele group had a mean water uptake that was 14 percentage points higher than the minor allele group (Table 5). Since the minor allele frequency is 0.06 for this SNP, the major allele with higher mean water uptake of this SNP is expected to be found more commonly. Therefore, this SNP could be useful in filtering out genotypes with lower water uptake potential. For all the SNPs associated with water uptake in NE, the allele groups that had higher water uptake contained only Andeans or admixtures, except for Chr02pos49223533, of which the minor allele (higher water uptake) group had one Middle American and 27 Andean or admixtures. Andeans having the higher water uptake alleles corresponded with the observation that larger seeded beans had higher water uptake in NE (Fig. S12).

Some of the SNPs associated with cooking time were supported by previous studies. The SNP on Pv06, Chr06pos27715283, in MI was within a QTL found by Berry et al. (2020): CT.6.1, 24,511,850–27,381,730 bp on Pv06, which was detected at two locations in Tanzania. They used a RIL population derived from a cross of two brown beans with contrasting cooking times. The parental lines of that population are part of the YBC: TZ-

27 (aka YBC133) and TZ-37 (aka YBC 134). The significant SNP on Pv08, Chr08pos62054742 in MI and S08_62659170, the SNP associated with cooking time in the Andean Diversity Panel reported by Bassett et al. (2021a), were close (604 kb apart), and they were within the range of a QTL, CT.8.2, for cooking time detected with a yellow bean RIL population (Bassett et al. 2021b). The QTL study used a yellow bean RIL population generated by a cross between a fast-cooking Manteca bean (Ervilha, YBC129) and a slow cooking Njano bean (PI527538, YBC130). In this study, the major allele group of Chr08pos62054742 had 3-min shorter mean cooking time than the minor allele group (Table 5), and YBC129 was in the minor allele (long-cooking) group. This was consistent with the QTL on Pv08, where the fast-cooking YBC129 conferred the longer-cooking allele (Bassett et al. 2021b). Overall, one QTL and three significantly associated SNPs for cooking time were found at 53.03–62.66 Mbp on Pv08 in three studies with Andean and/or yellow beans, suggesting that this region is worth further investigation of fast cooking time. The QTL and the SNPs could be informative markers for selection, especially yellow beans for cooking time.

The significant SNP on Pv10, Chr10pos43615359, in MI had the lowest *p*-values among all the SNPs associated with cooking time (Table 5). This SNP was in the range of a QTL, CT10.1 (37.83–43.84 Mb), detected with the highest LOD score among the QTLs for cooking time in the aforementioned yellow bean RIL population (Bassett et al. 2021b). Considering the strong and consistent association between the two studies, this region on Pv10 will be another region to focus on to investigate the mechanism of cooking time of yellow beans. Interestingly, YBC129 conferred the short cooking allele at CT10.1, while both YBC129 and YBC130 were in the major allele group in this study, which had a longer cooking time than the minor allele group at the SNP (Table 5). In the CT10.1 genomic region resides the *J* gene that controls non-darkening of seed coat (Erfatpour et al. 2018; Erfatpour and Pauls 2020). Post-harvest darkening is an important consumer-facing trait as darkened seeds are considered as old and long-cooking by consumers, leading to reduced economic values (Felicetti et al. 2012; Erfatpour et al. 2018). Some yellow genotypes are susceptible to post-harvest darkening while others are not (Sadohara et al. 2021).

The significant SNP on Pv01, Chr01pos1528388, in NE was at a distal region of Pv01. Three QTLs for cooking time were previously detected on Pv01 (Garcia et al. 2012); however, their physical locations were unknown. On Pv04, the marker Chr04pos4548944 was significantly associated with water uptake and cooking time in NE. Chr04pos4548944 is located in the coding region of Phvul.004G038700, which encodes a polygalacturonase (EC 3.2.1.15). Cotyledon cell separation is a pivotal process that allows bean softening, and cell separation is preceded by pectin solubilization in the middle lamella (Rockland and Jones 1974; Chigwedere et al. 2019). Pectins are complex polysaccharides made up of polygalacturonans as a backbone (Ridley et al. 2001). Polygalacturonase activities were increased during soaking and were associated with increased polygalacturonan extractability during cooking and decreased cooking time as compared to unsoaked beans (Martínez-Manrique et al. 2011). Therefore, a polymorphism in Phvul.004G038700 encoding a polygalacturonase may result in changes in the degree of polygalacturonan degradation in pectin, resulting in longer cooking times. The minor allele group of this SNP had 21.7 min longer cooking time than the major group. Moreover, a significantly associated SNP for cooking time with the ADP was also reported by Bassett et al. (2021a) 592 kbp upstream of this SNP.

Conclusions

This study evaluated the genotypic and phenotypic diversity of a yellow bean diversity panel in two environments in the U.S. for two years. The genotypic data for the YBC indicated that the YBC contains Andean, Middle American, and admixed individuals. A wide diversity was observed for growth habit, days to flower, seed weight, cooking time, and water uptake. Cooking times ranged from 17 to 123 min across all the environments and years, and the fastest cooking beans were of various yellow colors and sizes. The environmental effect including growing and cooking locations on water uptake and cooking time was significant, and no SNPs associated with those traits overlapped between the two environments, indicating the importance of testing multiple production environments. SNPs on Pv04, Pv06, Pv08, and Pv10 significantly associated with cooking time were

supported by previous studies, which calls for further investigation. A significant SNP on Pv04 was in the coding region of a polygalacturonase gene, and it was considered as a candidate gene for the fast-cooking trait. Given the diversity and the significantly associated SNPs for cooking time, the YBC will serve as a valuable resource for bean breeding for consumer-focused qualities and for genetic studies to elucidate the mechanism of the fast-cooking trait.

Acknowledgements The authors thank Anna Akariza, Amber Bassett, Gasana Ingabiregasana, Miranda Haus, Sharon Hooper, Queen Iribagiza, Hannah Jeffery, Hannah Peplinski, Scott Shaw, Jason Wiesinger, and Evan Wright for assistance with cooking, collecting plant tissues, and field operations. The authors express gratitude to Dr. Clare Mukankusi of Alliance Bioversity-CIAT Uganda and Bodo Raatz formerly of Alliance Bioversity-CIAT-Colombia for contributing the CIAT germplasm for this study and Phil McClean at North Dakota State University for his advice in analyzing the genetic diversity of the YBC.

Funding This work was supported by the USDA National Institute of Food and Agriculture AFRI (Award Number #: 2017–67013-26212), U.S. Department of Agriculture, Agricultural Research Service (Project 5050–21430-010-00D), Michigan State University, the National Science Foundation Research Traineeship Program (DGE-1828149) awarded to Rie Sadohara, and the Nebraska Hatch Project (NEB43-116).

Declarations

Conflicts of interest There are no competing interests to be disclosed.

References

- Aseete P, Katungi E, Bonabana-Wabbi J et al (2018) Consumer demand heterogeneity and valuation of value-added pulse products: a case of precooked beans in Uganda. *Agric Food Secur* 7:51. <https://doi.org/10.1186/s40066-018-0203-3>
- Bassett MJ (1999) The seedcoat color genotype of “Prim” and the Manteca and Coscorrón market classes of common bean. *HortScience* 34:336–337
- Bassett MJ (2002) Classical and Molecular Genetic Studies of the Strong Greenish Yellow Seedcoat Color in ‘Wagenaar’ and ‘Enola’ Common Bean. *J Am Soc Hortic Sci* 127:50–55
- Bassett A, Kamfwa K, Ambachew D, Cichy K (2021a) Genetic variability and genome-wide association analysis of flavor and texture in cooked beans (*Phaseolus vulgaris* L.). *Theor Appl Genet*. <https://doi.org/10.1007/s00122-020-03745-3>
- Bassett A, Katuuramu DN, Song Q, Cichy K (2021b) QTL mapping of seed quality traits including cooking time, flavor, and texture in a yellow dry bean (*Phaseolus vulgaris* L.) population. *Front Plant Sci* 12:670284. <https://doi.org/10.3389/fpls.2021.670284>
- Bates D, Mächler M, Bolker B, Walker S (2015) Fitting linear mixed effects models using lme4. *J Stat Softw*. 67:1–8. <https://doi.org/10.18637/jss.v067.i01>
- Beaver J, Zapata M, Alameda M, et al (2012) Common bean improvement in the Caribbean. In: 2012 Global Pulse Researchers Meeting. Kigali, Rwanda
- Beebe S, Skroch PW, Tohme J et al (2000) Structure of genetic diversity among common bean landraces of middle American origin based on correspondence analysis of RAPD. *Crop Sci* 40:264–273. <https://doi.org/10.2135/cropsci2000.401264x>
- Beninger CW, Hosfield GL, Nair MG (1998) Flavonol glycosides from the seed coat of a new Manteca-type dry bean (*Phaseolus vulgaris* L.). *J Agric Food Chem* 46:2906–2910. <https://doi.org/10.1021/jf9801522>
- Berry M, Izquierdo P, Jeffery H et al (2020) QTL analysis of cooking time and quality traits in dry bean (*Phaseolus vulgaris* L.). *Theor Appl Genet* 133:2291–2305. <https://doi.org/10.1007/s00122-020-03598-w>
- Birachi EA, Sperling L, Kadege E, Mdachi M, Upendo T, Radegunda K, Mutua M, Mbiu J, Raya N, Ndunguru A, William M, Kabungo C, Mcharo D, Shida N, Kilango M, Magelanga A, Maganga R, Kalempera S, Katungi E, Mukankusi C, Malle S, Dey B, Templer N, Rubyogo JC, Onyango P, and Buruchara R (2020) Analysis of the yellow bean corridor in Tanzania. A Feed the Future global supporting seed systems for development activity (S34D) report
- Bitocchi E, Bellucci E, Giardini A et al (2013) Molecular analysis of the parallel domestication of the common bean (*Phaseolus vulgaris*) in Mesoamerica and the Andes. *New Phytol* 197:300–313. <https://doi.org/10.1111/j.1469-8137.2012.04377.x>
- Box GE, Cox DR (1964) An analysis of transformations. *J R Stat Soc Ser B* 26:211–243. <https://doi.org/10.1111/j.2517-6161.1964.tb00553.x>
- Bressani R, Chon C (1996) Effects of altitude above sea level on the cooking time and nutritional value of common beans. *Plant Foods Hum Nutr* 49(1) 53–61. <https://doi.org/10.1007/BF01092522>
- Broughton WJ, Hernández G, Blair M et al (2003) Beans (*Phaseolus* spp.) – model food legumes. *Plant Soil* 252:55–128. <https://doi.org/10.1023/A:1024146710611>
- Brouwer I, den Hartog AP, Kamwendo M, Heldens M (1996) Wood quality and wood preferences in relation to food preparation and diet composition in Central Malawi. *Ecol Food Nutr* 35:1–13. <https://doi.org/10.1080/03670244.1996.9991471>
- Buruchara R, Chirwa R, Sperling L et al (2011) Development and delivery of bean varieties in Africa: the Pan-Africa bean research alliance (PABRA) model. *African Crop Sci J* 19:227–245
- Castellanos JZ, Guzmán-Maldonado H, Acosta-Gallegos JA, Kelly JD (1995) Effects of hardshell character on cooking time of common beans grown in the semiarid highlands of Mexico. *J Sci Food Agric* 69:437–443. <https://doi.org/10.1002/jsfa.2740690406>
- Chigwedere CM, Nkankola CM, Rai S et al (2019) Cotyledon pectin molecular interconversions explain pectin

- solubilization during cooking of common beans (*Phaseolus vulgaris*). Food Res Int 116:462–470. <https://doi.org/10.1016/j.foodres.2018.08.062>
- Cichy KA, Porch TG, Beaver JS et al (2015a) A *Phaseolus vulgaris* diversity panel for Andean bean improvement. Crop Sci 55:2149–2160. <https://doi.org/10.2135/cropsci2014.09.0653>
- Cichy KA, Wiesinger JA, Mendoza FA (2015b) Genetic diversity and genome-wide association analysis of cooking time in dry bean (*Phaseolus vulgaris* L.). Theor Appl Genet 128:1555–1567. <https://doi.org/10.1007/s00122-015-2531-z>
- Cichy KA, Wiesinger JA, Berry M et al (2019) The role of genotype and production environment in determining the cooking time of dry beans (*Phaseolus vulgaris* L.). Legum Sci. <https://doi.org/10.1002/leg3.13>
- Cooper M, DeLacy IH (1994) Relationships among analytical methods used to study genotypic variation and genotype-by-environment interaction in plant breeding multi-environment experiments. Theor Appl Genet 88:561–572
- Diaz S, Ariza-Suarez D, Ramdeen R et al (2021) Genetic architecture and genomic prediction of cooking time in common bean. Front Plant Sci 11:622213. <https://doi.org/10.3389/fpls.2020.622213>
- Díaz LM, Blair MW (2006) Race structure within the Mesoamerican gene pool of common bean (*Phaseolus vulgaris* L.) as determined by microsatellite markers. Theor Appl Genet 114:143–154. <https://doi.org/10.1007/s00122-006-0417-9>
- Dickerson GE (1962) Implications of genetic-environmental interaction in animal breeding. Anim Sci 4:47–63. <https://doi.org/10.1017/S0003356100034395>
- Duitama J, Quintero JC, Cruz DF et al (2014) An integrated framework for discovery and genotyping of genomic variants from high-throughput sequencing experiments. Nucleic Acids Res 42:e44. <https://doi.org/10.1093/nar/gkt1381>
- Earl DA, VonHoldt BM (2012) Structure harvester: a website and program for visualizing structure output and implementing the Evanno method. Conserv Genet Resour 4:359–361. <https://doi.org/10.1007/s12686-011-9548-7>
- Elia FM, Hosfield GL, Kelly JD, Uebersax M (1997) Genetic analysis and interrelationships between traits for cooking time, water absorption, and protein and tannin content of Andean dry beans. J Amer Soc Hort Sci 122:512–518
- Elshire RJ, Glaubitz JC, Sun Q et al (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity Species. PLoS ONE 6:e19379. <https://doi.org/10.1371/journal.pone.0019379>
- Engleright R, Beimiriki M, Hosfield G (1999) Determination of total dietary fiber, indigestible starch, and indigestible protein in dry bean (*Phaseolus vulgaris* L.). Annu Rep Bean Improv Coop 42:123–124
- Erfatpour M, Pauls KP (2020) A R2R3-MYB gene-based marker for the non-darkening seed coat trait in pinto and cranberry beans (*Phaseolus vulgaris* L.) derived from ‘Wit-rood boontje.’ Theor Appl Genet 133:1977–1994. <https://doi.org/10.1007/s00122-020-03571-7>
- Erfatpour M, Navabi A, Pauls KP (2018) Mapping the non-darkening trait from ‘Wit-rood boontje’ in bean (*Phaseolus vulgaris*). Theor Appl Genet 131:1331–1343. <https://doi.org/10.1007/s00122-018-3081-y>
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. Mol Ecol 14:2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Fehr WR (1987) Principles of Cultivar Development. Macmillan USA, New York City
- Felicetti E, Song Q, Jia G et al (2012) Simple sequence repeats linked with slow darkening trait in pinto bean discovered by single nucleotide polymorphism assay and whole genome sequencing. Crop Sci 52:1600–1608. <https://doi.org/10.2135/cropsci2011.12.0655>
- Garcia RAV, Rangel PN, Bassinello PZ et al (2012) QTL mapping for the cooking time of common beans. Euphytica 186:779–792. <https://doi.org/10.1007/s10681-011-0587-7>
- Gepts P, Bliss FA (1985) F1 hybrid weakness in the common bean: differential geographic origin suggests two gene pools in cultivated bean germplasm. J Hered 76:447–450. <https://doi.org/10.1093/oxfordjournals.jhered.a110142>
- Gepts P, Osborn TC, Rashka K, Bliss FA (1986) Phaseolin-protein variability in wild forms and landraces of the common bean (*Phaseolus vulgaris*): evidence for multiple centers of domestication. Econ Bot 40:451–468. <https://doi.org/10.1007/BF02859659>
- Hart JJ, Tako E, Wiesinger J, Glahn RP (2020) Polyphenolic profiles of yellow bean seed coats and their relationship with iron bioavailability. J Agric Food Chem 68:769–778. <https://doi.org/10.1021/acs.jafc.9b05663>
- Hooper S, Wiesinger JA, Echeverria D et al (2016) Carbohydrate profile of a dry bean (*Phaseolus vulgaris* L.) panel encompassing broad genetic variability for cooking time. Cereal Chem J 94:135–141. <https://doi.org/10.1094/CCHEM-04-16-0126-FI>
- Hosfield G, Bennink M, Beninger C et al (1998) Variability for starch digestibility in dry bean (*Phaseolus vulgaris*). HortScience 33:472
- Huang M, Liu X, Zhou Y et al (2019) BLINK: a package for the next level of genome-wide association studies with both individuals and markers in the millions. Gigascience. <https://doi.org/10.1093/gigascience/giy154>
- Jacinto-hernandez C, Azpiroz-rivero S, Acosta-gallegos JA et al (2003) Genetic analysis and random amplified polymorphic DNA markers associated with cooking time in common bean. Crop Sci 43:329–332
- Katuramu DN, Luyima GB, Nkalubo ST et al (2020) On-farm multi-location evaluation of genotype by environment interactions for seed yield and cooking time in common bean. Sci Rep 10:3628. <https://doi.org/10.1038/s41598-020-60087-2>
- Kelly JD, Awale HE, Wiersma AT et al (2021) Registration of ‘Yellowstone’ yellow bean. J Plant Regist. <https://doi.org/10.1002/plr2.20075>
- Kelly JD, Wright EM, Wiersma A (2018) Dry bean yield trials. In: Michigan State University AgBioResearch (ed) 2018 Research Report Saginaw Valley Research & Extension. p 15–51
- Kelly JD, Wright EM, Wiersma A (2019) 2019 Dry bean yield trials. In: Michigan State University AgBioResearch (ed) 2019 Research Report Saginaw Valley Research & Extension. p 16–52

- Kilimo Trust (2012) Development of Inclusive Markets in Agriculture and Trade (DIMAT). Undp 1–48
- Koenig R, Gepts P (1989) Allozyme diversity in wild *Phaseolus vulgaris*: further evidence for two major centers of genetic diversity. *Theor Appl Genet* 78:809–817. <https://doi.org/10.1007/BF00266663>
- Kuznetsova A, Brockhoff PB, Christensen RHB (2017) LmerTest package: tests in linear mixed effects models. *J Stat Softw* 82:1–26. <https://doi.org/10.18637/jss.v082.i13>
- Leakey CLA (1992) Breeding on the C and J and B loci for modification of bean seedcoat flavonoids with the objective of improving food acceptability. *Annu Rep Bean Improv Coop* 35:xiii–xvii
- Lee T-H, Guo H, Wang X et al (2014) SNPhylo: a pipeline to construct a phylogenetic tree from huge SNP data. *BMC Genomics* 15:162. <https://doi.org/10.1186/1471-2164-15-162>
- Lenth R V (2021) emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.5.5–1. <https://cran.r-project.org/package=emmeans>
- Lipka AE, Tian F, Wang Q et al (2012) GAPIT: genome association and prediction integrated tool. *Bioinformatics* 28:2397–2399. <https://doi.org/10.1093/bioinformatics/bts444>
- Lobaton JD, Miller T, Gil J et al (2018) Resequencing of common bean identifies regions of inter-gene pool introgression and provides comprehensive resources for molecular breeding. *Plant Genome*. <https://doi.org/10.3835/plantgenome2017.08.0068>
- Mainvieu R (2019) Agronomic performance and genetic diversity of common bean (*Phaseolus vulgaris*) varieties in Haiti. The University of Florida
- Mamidi S, Rossi M, Moghaddam SM et al (2013) Demographic factors shaped diversity in the two gene pools of wild common bean (*Phaseolus vulgaris* L.). *Heredity* (edinb) 110:267–276. <https://doi.org/10.1038/hdy.2012.82>
- Martín-Cabrejas MA, Esteban RM, Perez P et al (1997) Changes in physicochemical properties of dry beans (*Phaseolus vulgaris* L.) during long-term storage. *J Agric Food Chem* 45:3223–3227. <https://doi.org/10.1021/jf970069z>
- Martínez-Manrique E, Jacinto-Hernández C, Garza-García R et al (2011) Enzymatic changes in pectic polysaccharides related to the beneficial effect of soaking on bean cooking time. *J Sci Food Agric* 91:2394–2398. <https://doi.org/10.1002/jsfa.4474>
- Mishili FJ, Temu A, Fulton J, Lowenberg-DeBoer J (2011) Consumer preferences as drivers of the common bean trade in Tanzania: a marketing perspective. *J Int Food Agribus Mark* 23:110–127. <https://doi.org/10.1080/08974438.2011.558761>
- Oladzad A, Porch T, Rosas JC et al (2019) Single and multi-trait GWAS identify genetic factors associated with production traits in common bean under abiotic stress environments. *G3 Genes Genomes Genetics*. 9(6):1881–1892. <https://doi.org/10.1534/g3.119.400072>
- Pallottini L, Garcia E, Kami J et al (2004) The genetic anatomy of a patented yellow bean. *Crop Sci* 44:968–977. <https://doi.org/10.2135/cropsci2004.9680>
- Pérez-Vega E, Pañeda A, Rodríguez-Suárez C et al (2010) Mapping of QTLs for morpho-agronomic and seed quality traits in a RIL population of common bean (*Phaseolus vulgaris* L.). *Theor Appl Genet* 120:1367–1380. <https://doi.org/10.1007/s00122-010-1261-5>
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959
- Pritchard JK, Wen X, Falush D (2009) Documentation for structure software: Version 2.3. 1–39
- Proctor L (1999) Field bean cultivar called Enola. 1–8
- Purcell S, Neale B, Todd-Brown K et al (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81:559–575. <https://doi.org/10.1086/519795>
- R Core Team (2017) R: A language and environment for statistical computing.
- Rainey KM, Griffiths PD (2005) Differential response of common bean genotypes to high temperature. *J Am Soc Hortic Sci* 130(1):18–23. <https://doi.org/10.21273/JASHS.130.1.18>
- Ridley BL, O'Neill MA, Mohnen D (2001) Pectins: structure, biosynthesis, and oligogalacturonide-related signaling. *Phytochemistry* 57:929–967. [https://doi.org/10.1016/S0031-9422\(01\)00113-3](https://doi.org/10.1016/S0031-9422(01)00113-3)
- Rockland LB, Jones FT (1974) Scanning electron microscope studies on dry beans. effects of cooking on the cellular structure of cotyledons in rehydrated large lima beans. *J Food Sci* 39:342–346. <https://doi.org/10.1111/j.1365-2621.1974.tb02890.x>
- Sadohara R, Long Y, Izquierdo P, et al (2021) Seed coat color genetics and G × E in a Yellow Bean Collection via image analysis paired with machine-learning and GWAS. *Plant Genome* e20173. <https://doi.org/10.1002/tpg2.20173>
- Sandhu KS, You FM, Conner RL et al (2018) Genetic analysis and QTL mapping of the seed hardness trait in a black common bean (*Phaseolus vulgaris*) recombinant inbred line (RIL) population. *Mol Breed* 38:34. <https://doi.org/10.1007/s11032-018-0789-y>
- Schmutz J, McClean PE, Mamidi S et al (2014) A reference genome for common bean and genome-wide analysis of dual domestications. *Nat Genet* 46:707–713. <https://doi.org/10.1038/ng.3008>
- Sichilima T, Mapemba L, Tembo G (2016) Drivers of dry common beans trade in Lusaka, Zambia: a trader's perspective. *Sustain Agric Res* 5:15–26. <https://doi.org/10.5539/sar.v5n2p15>
- Siddiq M, Uebersax MA (2012) Dry beans and pulses production and consumption-an overview. *Dry Beans and Pulses Production. Processing and Nutrition*. Blackwell Publishing Ltd., Oxford, pp 1–22
- Singh SP, Gepts P, Debouck DG (1991) Races of common bean (*Phaseolus vulgaris*, Fabaceae). *Econ Bot* 45:379–396
- Smith LP, Ng SW, Popkin BM (2013) Trends in US home food preparation and consumption: analysis of national nutrition surveys and time use studies from 1965–1966 to 2007–2008. *Nutr J* 12:45. <https://doi.org/10.1186/1475-2891-12-45>
- Soltani A, Walter KA, Wiersma AT et al (2021) The genetics and physiology of seed dormancy, a crucial trait in common bean domestication. *BMC Plant Biol* 21:58. <https://doi.org/10.1186/s12870-021-02837-6>
- Sones D (2015) Soya Njano is the bean for home consumption. Our Blog Insid. Story; Africa Soil Heal. Consort.

- Sperling L, Birachi E, Kalemra S, Mutua M, Templer N, Mukankusi C, Radegunda K, William M, Gallagher P, Kadege E, Rubyogo JC (2021) The informal seed business: focus on yellow bean in Tanzania. *Sustainability* 13(16):8897. <https://doi.org/10.3390/su13168897>
- Stanley DW (1992) A possible role for condensed tannins in bean hardening. *Food Res Int* 25:187–192. [https://doi.org/10.1016/0963-9969\(92\)90136-S](https://doi.org/10.1016/0963-9969(92)90136-S)
- Tumeo M, Mapemba L, Edriss AK, Phiri H (2017) Consumer choice of dry common beans in Malawi: the case of Lilongwe City. Washington, D.C, USA
- US Department of Agriculture Agricultural Marketing Service (2019) “Patron” yellow bean. Plant Var Prot Certif 1–27
- Venables W, Ripley B (2002) Modern applied statistics with S, Fourth. Springer, New York
- Voysey O (2012) Yellow beans in Latin America. *Annu Rep Bean Improv Coop* xii–xviii
- Wang N, Daun JK (2005) Determination of cooking times of pulses using an automated Mattson cooker apparatus. *J Sci Food Agric* 85:1631–1635. <https://doi.org/10.1002/jsfa.2134>
- Wickham H (2016) *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag, New York
- Wiesinger JA, Cichy KA, Glahn RP et al (2016) Demonstrating a nutritional advantage to the fast-cooking dry bean (*Phaseolus vulgaris* L.). *J Agric Food Chem* 64:8592–8603. <https://doi.org/10.1021/acs.jafc.6b03100>
- Wiesinger J, Cichy K, Tako E, Glahn R (2018) The fast cooking and enhanced iron bioavailability properties of the Manteca yellow bean (*Phaseolus vulgaris* L.). *Nutrients* 10(11):1609. <https://doi.org/10.3390/nu10111609>
- Wortmann CS, Kirkby RA, Eledu CA, Allen DJ (1998) Atlas of common bean (*Phaseolus vulgaris* L.) production in Africa. CIAT Publ 297:131
- Yin L (2020) CMplot: Circle Manhattan Plot. R package version 3.6.2. <https://cran.r-project.org/package=CMplot>

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.