JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Article

pubs.acs.org/JAFC

Investigation of Biochemical Diversity in a Soybean Lineage Representing 35 Years of Breeding

George G. Harrigan,*[,†] Angela H. Culler,[†] Matthew Culler,[†] Matthew L. Breeze,[†] Kristina H. Berman,[†] Steven C. Halls,[†] and Jay M. Harrison[§]

[†]Monsanto Company, 800 North Lindbergh Boulevard, St. Louis, Missouri 63167, United States

[§]Maritz Holdings Inc., 1395 North Highway Drive, Fenton, Missouri 63099, United States

Supporting Information

ABSTRACT: This paper reports an assessment of seed biochemical and metabolite variability and diversity in a series of nine soybean varieties; all lines share the same genetic lineage but represent ~35 years of breeding (launch years 1972−2008) and differing yield potentials. These varieties, including six conventional and three glyphosate-tolerant lines, were grown concurrently at two replicated field sites in the United States during the 2011 growing season, and seeds were harvested at maturity. A compositional assessment included measurement of proximates, amino acids, fatty acids, tocopherols, isoflavones, saccharides, organic acids, and selected phytohormones. Statistical analysis included application of principal variance component analysis (PVCA) to investigate the interrelationships among compositional components from these soybean varieties and the impacts of location (environment) and pedigree on variability of these components. Results demonstrated that (i) some biochemical analytes showed trends (either increased or decreased) with launch year and/or yield, (ii) some analytes varied according to variety but showed no trend with launch year and/or yield, and (iii) almost all analytes showed extensive variation within and across sites. In summary, varietal development of high-yielding soybean, as represented in this study, has been accompanied by compositional changes but these are typically modest relative to environmental factors.

KEYWORDS: *soybean (Glycine max), selective breeding, compositional and metabolite variability, principal variance component analysis (PVCA)*

INTRODUCTION

Since its initial cultivation in the United States in the late 1700s,[1,2] soybean (*Glycine max* L.) has been established as a major source of food and feed products. Selective breeding of this versatile crop has now afforded hundreds of commercially available elite high-yielding cultivars.[3−8] An interest in improving the nutritional content of seed from high-yielding cultivars is also evident with numerous breeding programs established to modify concentrations of key seed components such as fatty acids and vitamins.[9−11] This activity has generated an extensive body of literature on the impact of germplasm and environment on levels of soybean metabolites.[12] Most studies to date report only on a subset of nutrients, and any systematic compositional assessments measuring multiple seed components have focused mainly on characterization of new genetically modified (GM) products.[13] However, there is value in understanding the impact of genotype and phenotype on a wide range of seed nutritional components as we seek to enhance concurrently both agronomic and nutritional qualities. Herein, we report the results of a study designed to assess seed biochemical composition in related soybean varieties of differing yield potentials. Selection of the study varieties was based on commercial launch year, providing further insights into variation in seed composition over multiple years of soybean breeding; nine varieties sharing the same genetic lineage and representing ~35 years of breeding (launch years 1972−2008) were assessed. The selection included six conventional and three glyphosate-tolerant (Roundup Ready) varieties. All varieties were grown concurrently at two replicated field sites in the United States during the 2011 growing season, and seed was harvested at maturity. The biochemical compositional assessment included measurement of proximates, amino acids, fatty acids, tocopherols, isoflavones, saccharides, organic acids, and selected phytohormones. Statistical analysis included application of principal variance component analysis (PVCA).[14,15] PVCA combines the application of two popular data analysis procedures, principal component analysis (PCA) and variance component analysis, and has three main goals: (i) to summarize large data sets with a smaller set of relevant variables; (ii) to describe the percentages of variance in the original data that are explained by the new variables; and then (iii) to describe the relative amounts of variation in those variables that can be explained by different aspects of the experimental design and other covariates. This approach was used here to investigate the impact of location (environment) and pedigree on variability and interrelationships among seed yield and biochemical components from the soybean varieties developed through multiple years of selective breeding.

MATERIALS AND METHODS

Soybean Samples and Field Production. Nine soybean varieties representing a genetic lineage from Williams (1972) to A3555 (2008)

Received: July 23, 2013
Revised: October 10, 2013
Accepted: October 14, 2013
Published: November 4, 2013

were grown at two sites in Illinois (Jerseyville and Jacksonville) during the 2011 season. Varieties included six conventional and three glyphosate-tolerant (Roundup Ready) lines. Variety and launch year are listed in Table 1. Starting seeds were planted in a randomized

Table 1. Soybean Varieties, Launch Year, Yield, and Seed Size

variety	launch year	mean yield (bu/acre)	mean seed size (g/100 seeds)
Williams	1972	65.87	14.72
A3127	1979	64.58	12.02
CX366	1986	69.26	13.56
CX375	1996	69.30	13.36
A3469	1997	76.87	12.23
AG3701	1999	71.95	12.66
AG3705	2006	78.73	14.55
AG3803	2008	77.57	14.89
A3555	2008	80.05	13.95

complete block design with six replicates. Soybean plants were treated with maintenance pesticides as necessary throughout the growing season at both sites. The three Roundup Ready varieties were not treated with glyphosate. Seeds were harvested at maturity, homogenized by grinding with dry ice to a fine powder, and stored frozen at approximately −20 °C. Samples were then lyophilized prior to compositional analysis.

Agronomic Analysis. Yield is reported in bushels per acre, and seed size is based on grams per 100 seeds.

Compositional Analyses. Components assessed included proximate (protein, ash, fat), amino acids, fatty acids, tocopherols (α-, γ-, δ-), saccharides (raffinose, stachyose, glucose, fructose, galactose, sucrose), organic acids (acetic, citric, lactic, malic, oxalic), isoflavones (daidzein, genistein, glycitein), and selected phytohormones (abscisic acid, indole-3-acetic acid, indole-3-acetyl aspartic acid, isopentenyladenosine, dihydrozeatin, dihydrozeatin-3-riboside, *trans*-zeatin riboside). Brief descriptions of the methods utilized for the analyses of proximates, fatty acids, isoflavones, and sugars appear in refs 16 and 17. Amino acid methodology is reported in ref 18. Organic acid methodology utilized an AOAC method.[19] Tocopherol analysis was based on a reversed-phase HPLC method using fluoresence detection with excitation at 290 nm and emission at 336 nm. Tocopherols were extracted from ground lyophilized seed with 0.1% pyrogallol in ethanol. The reversed-phase HPLC system comprised a Keystone Aquasil C_{18} column (100 mm × 4.6 mm × 3.0 μm, Thermo Fisher) at 40 °C and methanol as mobile phase. Flow rate was 1 mL/min.

Phytohormone analysis was based on LC-MS/MS methodology. The methods were separated into three categories: group 1 (abscisic acid, indole-3-acetic acid, indole-3-acetyl aspartic acid); group 2 (isopentenyladenosine, dihydrozeatin, dihydrozeatin-3-riboside); and *trans*-zeatin riboside. The extraction process for group 1 involved 0.2 g of ground lyophilized seed suspended in 0.9 mL of 60:40 water/ACN in 0.1% formic acid and a 0.1 mL spike of deuterated internal standard in 60% ACN. A further 0.1 mL of 60% ACN was then added to all samples. The clear supernatants (0.1 mL) of centrifuged samples were transferred after dilution with 0.2 mL of 0.1% (v/v) acetic acid in water for chromatography. *trans*-Zeatin riboside can be analyzed from the group 1 extraction. The extraction process for group 2 involved 0.4 g of ground lyophilized seed suspended in 1.4 mL of 80:20 ACN/water in 0.1% formic acid and a 0.15 mL spike of deuterated internal standard in 60% ACN. A further 0.1 mL of 60% ACN was then added to all samples. Extract solution (0.15 mL) plus 0.05 mL of ACN was evaporated to dryness and then reconstituted in 0.025−0.05 mL of 60% CAN followed by addition of 0.3 mL of 0.1% formic acid (v/v) in ACN.

Deuterated internal standards were prepared at 10 μg/mL for group 1 and at 2.5 μg/mL for group 2 and *trans*-zeatin riboside. Deuterated internal standards were purchased from ChemIM Ltd. (Czech Republic) with the exception of salicylic acid-d_4 (CDN Isotopes, Canada).

Chromatography for group 1 and *trans*-zeatin riboside used reversed-phase analysis (Supelco Acentis Phenyl, 50 mm × 2.1 mm × 3.0 μm column). Mobile phases for group 1 were (A) 0.1% (v/v) acetic acid in water and (B) 0.5% (v/v) acetic acid in methanol with the following gradient: 0 min, 5% B; 0.0−3.0 min, 30% B; 3.0−11.0 min, 30% B; 11.0−11.5 min, 80% B; 11.0−14.30 min, 80% B; 14.3−14.31 min, 5% B; 14.31−16.0 min, 5% B. The flow rate was 0.25 mL/min for the entire run except between 11.5 and 14.0 min, when the flow rate was 0.5 mL/min. The operating temperature was 45 °C. Mobile phases for *trans*-zeatin ribose were (A) 100% deionized (DI) water and (B) 5 mM ammonium formate in methanol with the following gradient: 0.0−2.0 min, 10−30% B; 2.0−6.0 min, 30−80% B; 6.0−8.0 min, 80% B; 8.0−10.0 min, 10% B. The flow rate was 0.6 mL/min. The operating temperature was ambient. Chromatography for group 2 used HILIC (Seilic Obelisc-N, 100 mm × 5 μm. Mobile phases for group 2 were (A) 0.3% (v/v) formic acid in water, pH 3, with ammonium formate and (B) 1% (v/v) formic acid, 90% (v/v) acetonitrile, and 9% (v/v) water with the following gradient: 0 min, 92% B; 0.0−7.0 min, 60% B; 7.0−10.0 min, 60% B; 10.0−14.0 min, 60% B; 14.0−14.1 min, 92% B; 14.1−17.0 min, 92% B. The flow rate was 0.15 mL/min for the entire run except between 10.5 and 15.5 min, when the flow rate was 0.6 mL/min. The operating temperature was 40 °C. All chromatography was performed on a Shimadzu (Columbia, MD, USA) Promenence 20A HPLC and autosampler system.

Mass spectrometric data analyses for group 1 were performed in negative MRM mode. Precursor and product ions for group 1 metabolites and deuterated standards are as follows: abscisic acid (263/153), abscisic acid-d_6 (269/159), indole-3-acetic acid (174/130), indole-3-acetic acid-d_5, (179/135), indole-3-acetyl aspartic acid (289/132), and salicylic acid-d_4 (137/65).

Mass spectrometric data analyses for group 2 and *trans*-zeatin riboside were performed in positive MRM mode. Precursor and product ions for group 2 metabolites, *trans*-zeatin riboside, and deuterated standards are as follows: isopentenyladenosine (336/204), isopentenyladenosine-d_6 (342/210), dihydrozeatin (222/136), dihydrozeatin-d_3 (225/149), dihydrozeatin-3-riboside (354/222), dihydrozeatin-3-riboside-d_3 (357/225), *trans*-zeatin riboside (352/136), and *trans*-zeatin riboside-d_5 (357/137).

The MS source conditions were as follows: For group 1, curtain gas (CUR) of 45 psi (N_2), ion spray voltage (IS) of −4000 V, collisionally activated dissociation gas (CAD) of 6 psi (N_2), nebulizer gas (GS1) of 45 psi, heater gas (GS2) of 45 psi, and source temperature (TEM) of 550 °C. For Group 2 and for *trans*-zeatin riboside, the MS source conditions were CUR of 35 psi (N_2), IS of 5000 V, CAD of 6 psi (N_2), nebulizer gas (GS1) of 45 psi, heater gas (GS2) of 45 psi, and TEM of 400 °C. The mass spectrometer was an Applied Biosystems API 5000 MS/MS using Analyst software.

Statistical Analysis of Composition Data. Analytes for which >50% of observations were below the assay LOQ were excluded from analysis. These included fructose, glucose, galactose, acetic acid, lactic acid, oxalic acid, dihydrozeatin, dihydrozeatin-3-riboside, and *trans*-zeatin riboside.

The procedure and rationale for applying PVCA to crop compositional data were described in detail in ref 15. Essentially, the first step of the PVCA procedure is to normalize the responses, if necessary. For example, a variance-stabilizing transformation, such as a logarithm, may be applied to individual analytes that have skewed distributions due to a few extreme values. The next step is to standardize each of the responses by subtracting the mean of all observations for each analyte and then dividing that difference by the sample standard deviation, or $Z = (X - \overline{X})/Sx$. The third step is to check for linear dependencies in the data. This can be done by computing the correlation matrix of the standardized variables and then computing the rank of that matrix. If the rank is less than the number of variables, then some of the variables must be dropped from the analyses. The fourth step is to apply PCA to the correlation matrix. Li et al.[14] continued with their development of the original PVCA procedure with the principal components, but we added an intermediate step to assist with the interpretation of the results.

Factor analysis is a method of deriving new linear combinations of variables from the principal components via rotation, a mathematical operation involving matrix multiplication. For our procedure, we chose the varimax rotation method. After application of a varimax rotation to the principal components, the next step is to apply variance components analysis to each of the derived factor variables $\{F^1, F^2, ..., F^j\}$. An ANOVA model is applied, and all of the sources of variation of interest to the researcher are modeled with random effects.

In the current study, PCA showed that 10 principal components were needed to meet a threshold of 80% (as discussed in refs 14 and 15) of the total study variance explained and were retained for subsequent examination. Each of the first 10 eigenvalues was >1, indicating that each of these components explained more than the average amount of variance that could be explained by all principal components.

Varimax rotation was applied to the first 10 principal components to form a set of 10 rotated factors as described in ref 15, which also discusses the advantages of applying varimax rotation. In our application of PVCA, the principal components and factor scores are said to be strongly correlated with individual compositional analytes if the absolute value of the linear correlation between the derived variable and the analyte exceeds 0.707, with the interpretation that at least 50% of the variation in the analyte can be explained by its linear relationship with the new variable. Absolute correlations between 0.5 and 0.707 can be described as moderately correlated, with the amount of variation in the analyte that can be explained by the factor ranging between 25 and 50%.

PVCA was conducted in JMP 9.0 (SAS Institute Inc.). Means reported in Tables 1− 4 and correlations reported in the following text were calculated in JMP.

RESULTS AND DISCUSSION

Selective plant breeding has contributed to an increase in soybean yield as reflected in the varieties selected for this study (Table 1). Seed size data are also presented in Table 1. The varieties selected here therefore represent an opportunity to evaluate compositional variability associated with decades of conventional breeding as well as its relationship to variability in yield and seed size. The compositional evaluation of these lines was conducted by analyzing the seed data within and across both field sites. Summaries of results from this combined-site analysis are presented in Tables 2−4. Overall, the data demonstrated that (i) some biochemical analytes showed trends (either increased or decreased in levels) with launch year and/or yield, (ii) levels of some analytes varied according to variety but showed no trend with launch year and/or yield, and (iii) levels of almost all analytes showed extensive variation within and across sites. PVCA was subsequently used to quantify the relative contribution of the features of the experimental design: location, pedigree (i.e., variety for a given launch year), pedigree by location interaction, biological replicate within location, and residual error on compositional variability. The statistical model also included yield and seed size to allow an assessment of their relationships with composition. Finally, this study also represented a continued evaluation of PVCA[15] as a complementary addition to traditional univariate methods in compositional analysis.

A total of 10 factors described >80% of the variance in the study. Factors F^1−F^5 are discussed in more detail below. Summaries of factors F^6−F^{10} are presented in Supplementary Tables 5−9 in the Supporting Information. The assignments of each individual compositional analyte to factors, based on the results of the varimax rotation, are listed in Tables 2−4. Coefficients for all factors are listed in Supplementary Table 1 (Supporting Information) with the information presented in heatmap and dendrogram format in Figure 1. Figure 2 shows the cumulative contributions of each variance component over all factors, whereas Figure 3 shows the same information when grouped by factor.

Factor F^1, Protein and Amino Acids, and Factor F^8, Cystine, Methionine, and Proline. The first factor, F^1, obtained by varimax rotation of the 10 principal components, explained 21.2% of the total variation in the composition study. It was strongly correlated with total protein and strongly or moderately correlated with each amino acid except proline, methionine, cystine, and tryptophan (see later). Next, analysis of variance with only random effects was used to estimate the variance in F^1 from the following sources: location, pedigree (i.e., variety for a given launch year), pedigree by location interaction, biological replicate within location, and residual error. This analysis showed that the largest source of variance in F^1 was residual error, followed by pedigree and location (Table 5). Figure 4 contains box plots to display the variation of values in F^1 due to location and pedigree.

A consistent observation in studies of soybean seed is the inverse correlation between protein values and yield;[20] our results showed a trend, albeit modest, consistent with this observation (r^2 = 0.249; F_1 correlation with yield = −0.34). Overall, there was a decline in mean protein levels (and mean F^1 values) from the older to newer launch varieties (Table 2 and Figure 4). There was also extensive within-site variability as evidenced in the range of individual replicate values (Table 2). Therefore, it may be concluded that long-term changes in seed protein composition in this lineage are less than changes attributable to real-time environmental factors, such as growing location and field position.

Data from the amino acid analysis were broadly consistent with those of protein (see Table 2) and consistent with the observation by Mahmoud et al.[20] that protein quality has not been noticeably affected by selective breeding. In contrast,[20] however, we did not observe any differential trend for the sulfur-containing amino acids, cystine and methionone. Methionine and cystine were strongly correlated with F^8, and proline was moderately correlated. This factor explained 4.51% of the variance in the study, and the largest contributor to variation was residual error.

Factor F^2, Fatty Acid Composition. Factor F^2 explained 10.41% of the total variance in this study and was strongly or moderately correlated with all of the major fatty acids measured in this study with the exception of palmitic acid. The largest source of variance in F^2 was pedigree followed by residual error (Supplementary Table 2, Supporting Information). There were no obvious trends in F^2 fatty acid levels across launch era. This observation also extended to fat, which showed a wide range of combined-site mean values extending from 16.07% dwt (A3469) to 18.37% dwt (A3733/CX329). As for protein, there was extensive within-site variability for fat and fatty acids as evidenced by the range of individual replicate values for these components (Table 3). The data suggest that breeding for enhanced agronomics traits will not have an immediate or meaningful impact on fatty acid composition.

Factor F^3, Yield Components, Palmitic Acid, Sucrose, and Isoflavone Composition. The factor F^3 explained 8.65% of the total variance in this study and was strongly correlated with daidzein and genistein and moderately correlated with glycitein, seed weight, and palmitic acid. Correlations were observed for sucrose (0.49) and yield (0.45). These values were below our stated threshold for moderate correlation (0.5) but did represent the highest values for sucrose and yield with any

Table 2. Summary of Protein and Amino Acid Values

component[a]	Williams mean (range[b])	A3127 mean (range[b])	CX366 mean (range[b])	CX375 mean (range[b])	A3469 mean (range[b])	AG3701 mean (range[b])	AG3705 mean (range[b])	AG3803 mean (range[b])	A3555 mean (range[b])	main factor[c]	effect[d]
protein	36.79 (34.4−37.9)	37.95 (36.8−38.7)	36.1 (34.3−37.5)	36.93 (35.8−37.8)	36.07 (35.1−37.1)	35.95 (34.6−37.1)	34.77 (33.6−35.8)	35.72 (35.2−36.6)	36.32 (35.3−37.7)	F^1	site
alanine	1.64 (1.57−1.72)	1.61 (1.51−1.69)	1.57 (1.38−1.66)	1.60 (1.51−1.70)	1.57 (1.47−1.64)	1.57 (1.41−1.68)	1.52 (1.35−1.60)	1.54 (1.42−1.64)	1.58 (1.46−1.68)	F^1	site
arginine	2.74 (2.61−2.86)	2.75 (2.57−2.88)	2.64 (2.38−2.79)	2.72 2.55−2.95)	2.63 (2.46−2.74)	2.63 (2.34−2.82)	2.51 (2.25−2.69)	2.61 (2.37−2.85)	2.66 (2.45−2.85)	F^1	site
aspartic acid	4.12 (3.85−4.34)	4.05 (3.68−4.43)	3.92 (3.49−4.15)	4.03 (3.69−4.59)	3.87 (3.57−4.18)	3.84 (3.36−4.18)	3.75 (3.28−4.22)	3.85 (3.57−4.21)	3.90 (3.57−4.11)	F^1	site
cystine	0.52 (0.28−0.61)	0.57 (0.49−0.66)	0.54 (0.44−0.59)	0.56 (0.47−0.65)	0.55 (0.32−0.67)	0.54 (0.31−0.62)	0.51 (0.32−0.61)	0.56 (0.35−0.66)	0.53 (0.31−0.64)	F^8	residual error
glutamic acid	7.34 (6.95−7.67)	7.36 (6.82−7.86)	7.06 6.32−7.49)	7.21 (6.82−7.81)	7.01 (6.60−7.37)	6.97 (6.41−7.45)	6.65 (5.99−7.19)	6.93 (6.41−7.40)	7.02 (6.51−7.57)	F^1	site
glycine	1.63 (1.55−1.73)	1.63 (1.53−1.73)	1.58 (1.39−1.67)	1.61 (1.53−1.70)	1.61 (1.51−1.68)	1.61 (1.44−1.71)	1.53 (1.33−1.61)	1.57 (1.42−1.72)	1.59 (1.47−1.71)	F^1	site
histidine	0.88 (0.71−0.97)	0.83 (0.65−0.95)	0.83 (0.64−0.92)	0.83 (0.71−0.92)	0.76 (0.61−0.86)	0.82 (0.78−0.87)	0.81 (0.67−0.87)	0.81 (0.67−0.94)	0.84 (0.68−0.93)	F^1	site
isoleucine	1.88 (1.71−2.00)	1.83 (1.66−2.01)	1.77 (1.62−1.89)	1.83 1.71−2.03)	1.73 (1.58−1.88)	1.73 (1.64−1.87)	1.69 (1.48−1.89)	1.74 (1.55−1.95)	1.77 (1.65−1.86)	F^1	site
leucine	2.89 (2.75−3.03)	2.83 (2.64−2.98)	2.76 (2.45−2.91)	2.82 (2.65−3.04)	2.70 (2.54−2.81)	2.70 (2.41−2.86)	2.64 (2.35−2.81)	2.69 (2.45−2.91)	2.74 (2.55−2.94)	F^1	site
lysine	2.39 (2.28−2.53)	2.35 (2.14−2.50)	2.27 (2.00−2.40)	2.33 (2.22−2.45)	2.27 (2.19−2.40)	2.28 (2.06−2.47)	2.19 (1.90−2.32)	2.25 (2.00−2.42)	2.29 (2.14−2.46)	F^1	site
methionine	0.58 (0.37−0.64)	0.61 (0.57−0.66)	0.59 (0.53−0.62)	0.59 (0.54−0.62)	0.57 (0.32−0.64)	0.58 (0.36−0.63)	0.56 (0.38−0.63)	0.58 (0.35−0.64)	0.57 (0.35−0.64)	F^8	residual error
phenylalanine	1.54 (1.46−1.62)	1.51 (1.41−1.59)	1.47 (1.30−1.55)	1.50 (1.41−1.68)	1.45 (1.34−1.51)	1.44 (1.29−1.53)	1.39 (1.25−1.52)	1.43 (1.32−1.54)	1.46 (1.34−1.57)	F^1	site
proline	1.86 (1.33−2.25)	1.99 (1.56−2.29)	1.80 (1.47−2.11)	1.85 (1.46−2.32)	1.88 (1.38−2.22)	1.74 (1.28−2.26)	1.82 (1.53−2.16)	2.08 (1.64−3.39)	1.70 (1.76−2.36)	F^8	residual error
serine	1.84 (1.65−2.05)	1.86 (1.69−2.04)	1.80 (1.58−2.04)	1.81 (1.60−2.05)	1.82 (1.56−2.01)	1.79 (1.49−2.04)	1.72 (1.56−1.88)	1.78 (1.56−1.97)	1.80 (1.52−2.01)	F^1	site
threonine	1.49 (1.42−1.57)	1.46 (1.35−1.55)	1.43 (1.28−1.52)	1.46 (1.35−1.55)	1.43 (1.33−1.52)	1.42 (1.28−1.55)	1.40 (1.28−1.49)	1.40 (1.31−1.47)	1.44 (1.34−1.53)	F^1	site
tryptophan	0.50 (0.42−0.52)	0.48 (0.37−0.54)	0.47 (0.41−0.49)	0.49 (0.44−0.50)	0.48 0.46−0.51)	0.46 (0.39−0.49)	0.46 (0.41−0.48)	0.47 (0.45−0.53)	0.49 (0.46−0.52)	none	none[e]
tyrosine	1.47 (1.41−1.53)	1.44 (1.32−1.52)	1.41 (1.25−1.48)	1.43 (1.35−1.56)	1.39 (1.32−1.46)	1.40 (1.28−1.48)	1.35 (1.23−1.46)	1.37 (1.24−1.48)	1.41 (1.30−1.50)	F^1	site
valine	1.95 (1.76−2.05)	1.89 (1.72−2.05)	1.84 (1.65−1.98)	1.91 1.77−2.19)	1.78 (1.63−1.98)	1.81 (1.72−1.95)	1.77 (1.54−2.05)	1.80 (1.63−2.02)	1.82 (1.67−1.93)	F^1	site

[a]Expressed as % dwt. [b]Minimum to maximum of individual replicate values across both sites. [c]Factor to which an analyte shows coefficient of greatest magnitude. [d]Greatest source of variance for given factor. [e]Not highly correlated with any of the factors that explained 80% of the total variation.

dx.doi.org/10.1021/jf4032102 | J. Agric. Food Chem. 2013, 61, 10807−10815

Table 3. Summary of Fat, Fatty Acid, and Tocopherol Values

component[a]	Williams mean (range[b])	A3127 mean (range[b])	CX366 mean (range[b])	CX375 mean (range[b])	A3469 mean (range[b])	AG3701 mean (range[b])	AG3705 mean (range[b])	AG3803 mean (range[b])	A3555 mean (range[b])	main factor[c]	effect[d]
total fat	18.27 (15.20−21.20)	16.58 (14.40−19.00)	17.44 (15.30−19.20)	18.37 (16.40−20.60)	16.07 (13.20−17.50)	17.55 (12.10−19.30)	17.73 (15.40−19.10)	17.67 (14.10−19.20)	17.89 (15.70−21.80)	none	none[e]
myristic	0.08 (0.07−0.08)	0.08 (0.08−0.09)	0.08 (0.07−0.09)	0.08 (0.07−0.08)	0.08 (0.07−0.09)	0.07 (0.07−0.09)	0.07 (0.07−0.09)	0.08 (0.08−0.10)	0.08 (0.08−0.09)	F^6	pedigree
pentadecylic	0.02 (0.02−0.02)	0.02 (0.02−0.03)	0.02 (0.02−0.02)	0.02 (0.02−0.03)	0.02 (0.02−0.03)	0.02 (0.02−0.03)	0.02 (0.02−0.02)	0.02 (0.02−0.02)	0.02 (0.02−0.02)	F^3	pedigree
palmitic	10.54 (10.35−10.81)	10.89 (10.58−11.24)	10.70 (10.20−11.02)	10.63 (10.32−10.87)	10.32 (9.95−11.30)	10.60 (10.42−10.68)	11.26 (11.09−11.71)	11.77 (11.50−12.26)	11.84 (11.55−12.02)	F^3	pedigree
palmitoleic	0.08 (0.07−0.08)	0.09 (0.09−0.11)	0.08 (0.08−0.09)	0.09 (0.08−0.10)	0.10 (0.09−0.11)	0.10 (0.10−0.11)	0.08 (0.07−0.08)	0.10 (0.09−0.10)	0.09 (0.09−0.10)	F^4	pedigree
heptadecanoic	0.11 (0.10−0.11)	0.10 (0.09−0.10)	0.10 (0.09−0.11)	0.10 (0.09−0.11)	0.09 (0.08−0.11)	0.10 (0.09−0.11)	0.11 (0.10−0.12)	0.11 (0.10−0.12)	0.10 (0.10−0.11)	none	none[e]
heptadecenoic	0.05 (0.05−0.05)	0.05 (0.04−0.05)	0.05 (0.05−0.05)	0.05 (0.05−0.06)	0.05 (0.05−0.06)	0.05 (0.05−0.06)	0.05 (0.04−0.06)	0.05 (0.05−0.06)	0.05 (0.04−0.06)	F^6	pedigree
stearic	4.09 (3.81−4.42)	4.09 (3.87−4.83)	4.01 (3.69−4.66)	4.21 (3.98−4.45)	3.68 (3.38−4.43)	3.80 (3.32−4.13)	4.35 (4.05−4.87)	4.24 (3.96−5.08)	3.99 (3.78−4.18)	F^2	pedigree
oleic	21.49 (20.36−22.61)	20.09 (18.82−22.38)	20.97 (19.76−23.23)	21.39 (20.10−22.30)	20.59 (19.97−21.85)	21.07 (19.90−22.88)	21.06 (19.60−22.65)	19.72 (18.81−20.49)	19.32 (18.51−20.11)	F^2	pedigree
vaccenic	1.35 (1.29−1.40)	1.52 (1.48−1.60)	1.47 (1.43−1.58)	1.45 (1.39−1.51)	1.62 (1.59−1.67)	1.59 (1.52−1.66)	1.38 (1.30−1.42)	1.48 (1.41−1.55)	1.57 (1.51−1.63)	F^4	pedigree
linoleic	52.40 (51.61−53.15)	52.55 (49.99−53.87)	52.65 (50.01−53.88)	51.96 (51.52−52.43)	52.58 (51.54−53.56)	52.37 (50.83−53.27)	51.49 (49.86−52.26)	52.09 (50.67−52.83)	52.56 (51.78−53.30)	F^2	pedigree
linolenic	7.08 (6.39−7.59)	7.53 (6.66−8.08)	7.06 (6.14−7.52	7.08 6.72−7.54)	7.90 (7.43−8.37)	7.45 (6.65−8.13)	7.23 (6.58−7.83)	7.13 (6.86−7.41)	7.72 (7.23−8.20)	F^2	pedigree
arachidic	0.30 (0.29−0.33)	0.32 (0.28−0.41)	0.31 (0.28−0.38)	0.32 (0.29−0.36)	0.28 (0.26−0.32)	0.28 (0.23−0.31)	0.34 (0.31−0.41)	0.32 (0.29−0.41)	0.29 (0.11−0.32)	F^2	pedigree
eicosenoic	0.20 (0.18−0.24)	0.20 (0.18−0.24)	0.21 (0.20−0.26)	0.21 (0.18−0.22)	0.20 (0.19−0.24)	0.19 (0.18−0.21)	0.21 (0.19−0.24)	0.19 (0.17−0.21)	0.20 (0.18−0.20)	F^2	pedigree
eicosdienoic	0.06 (0.06−0.07)	0.06 (0.06−0.07)	0.07 (0.06−0.07)	0.07 (0.06−0.07)	0.07 (0.06−0.08)	0.07 (0.06−0.07)	0.07 (0.06−0.08)	0.06 (0.06−0.08)	0.07 (0.06−0.07)	F^{10}	residual error
behenic	0.32 (0.29−0.37)	0.36 (0.31−0.47)	0.34 (0.31−0.43)	0.36 (0.31−0.40)	0.33 (0.29−0.39)	0.34 (0.25−0.38)	0.38 (0.35−0.48)	0.32 (0.29−0.42)	0.32 (0.30−0.35)	F^2	pedigree
lignoceric	0.14 (0.12−0.16)	0.15 (0.13−0.20)	0.15 (0.14−0.19)	0.14 (0.12−0.16)	0.14 (0.12−0.16)	0.13 (0.09−0.15)	0.15 (0.14−0.19)	0.14 (0.13−0.19)	0.15 (0.14−0.16)	F^2	pedigree
α-tocopherol	9.92 (8.30−12.40)	11.88 (9.60−14.50)	11.98 (8.50−14.60)	11.13 (9.00−13.30)	15.00 (12.40−17.30)	8.54 (7.00−10.60)	11.43 (9.10−14.20)	9.14 (7.10−10.90)	12.48 (10.50−15.60)	F^5	residual error
γ-tocopherol	173.28 (151.30−201.80)	162.94 (143.70−184.00)	174.53 (146.10−200.90)	161.78 (142.20−190.50)	165.23 (145.60−186.60)	147.33 (127.90−174.70)	157.45 (133.90−183.60)	156.55 (134.70−176.50)	169.82 (153.10−183.10)	F^5	residual error
δ-tocopherol	51.97 (45.70−59.20)	41.43 (35.40−47.50)	53.33 (46.70−61.50)	46.44 (41.80−57.10)	49.60 (41.50−68.50)	42.38 (36.40−51.40)	49.98 (37.40−79.20)	48.25 (38.90−53.20)	55.05 (51.50−58.50)	F^5	residual error

[a]Fat expressed as % dwt; fatty acids (FA) expressed as % total FA; tocopherols expressed as mg/100g dwt. [b]Minimum to maximum of individual replicate values across both sites. [c]Factor to which an analyte shows coefficient of greatest magnitude. [d]Greatest source of variance for given factor. [e]Not highly correlated with any of the factors that explained 80% of the total variation.

dx.doi.org/10.1021/jf4032102 | J. Agric. Food Chem. 2013, 61, 10807−10815

Table 4. Summary of Isoflavone, Sacharide, Organic Acid, Phytohormone, and Ash Values

component[a]	Williams mean (range[b])	A3127 mean (range[b])	CX366 mean (range[b])	CX375 mean (range[b])	A3469 mean (range[b])	AG3701 mean (range[b])	AG3705 mean (range[b])	AG3803 mean range[b])	A3555 mean (range[b])	main factor[c]	effect[d]
isoflavones											
daidzein	632 (544−716)	362 (295−422)	521 (420−622)	569 (476−622)	633 (561−723)	700 (628−765)	744 (639−849)	701 (619−764)	1081 (930−1210)	F^3	pedigree
genstein	922 (815−1030	555 (472−610)	724 (559−866)	771 (673−849)	764 (662−841)	888 (812−1020)	878.1 (785−949)	837 (773−903)	910 (835−998)	F^3	pedigree
glycitein	146.7 (104−190)	177.7 126−231)	155 (94.5−219)	129.9 (105−167)	109.6 (71.7−139)	163.2 (109−217)	83.2 (60.8−121)	90 (68.1−127)	101 (73.2−145)	F^3	pedigree
saccharides[e]											
sucrose	5.44 (4.30−5.90)	4.73 (4.20−5.10)	5.62 (4.90−7.00)	5.46 (5.10−6.20)	6.15 (5.80−6.70)	6.11 (5.50−6.70)	6.15 (5.60−6.70)	6.46 (5.80−6.90)	5.59 (5.20−6.00)	F^{3i}	pedigree
raffinose	0.58 (0.50−0.60)	0.63 (0.60−0.70)	0.59 (0.50−0.70)	0.51 (0.50−0.60)	0.63 (0.60−0.70)	0.64 (0.60−0.70)	0.75 (0.60−0.80)	0.66 (0.60−0.80)	0.69 (0.60−0.80)	F^{7i}	residual error
stachyose	3.80 (3.60−4.00)	4.42 (4.10−5.00)	4.28 (4.00−4.50)	4.22 (4.00−4.60)	4.58 (4.20−4.80)	4.21 (3.80−4.60)	4.42 (4.00−4.90)	4.35 (3.80−4.80)	4.50 (4.20−4.90)	F^6	pedigree
organic acids[f]											
citric	1366 (853−1660)	1647 (1460−1950)	1687 (1470−1900)	1602 (1390−1760)	1489 (1380−1610)	1678 (1280−1940)	1653 (932−1820)	1540 (1100−1840)	1492 (1310−1600)	none	none[h]
malic	95 (57−133)	101 (84−123)	111 (96−137)	101 (75−139)	177 (150−242)	152 (117−197)	110 (62−138)	140 (97−199)	173 (127−200)	F^4	pedigree
phytohormones[g]											
abscisic acid	0.11 (0.09−0.13)	0.10 (0.08−0.12)	0.10 (0.09−0.10)	0.08 (0.07−0.10)	0.08 (0.06−0.09)	0.09 (0.07−0.11)	0.09 (0.07−0.11)	0.10 (0.07−0.16)	0.10 (0.08−0.13)	F^7	residual error
indole-3-acetic acid	0.03 (0.03−0.04)	0.06 (0.05−0.09)	0.04 (0.03−0.06)	0.03 (0.02−0.04)	0.04 (0.02−0.06)	0.04 (0.02−0.06)	0.04 (0.03−0.06)	0.11 (0.03−0.41)	0.04 (0.03−0.05)	F^7	residual error
indole-3-acetyl-aspartic acid	0.17 (0.11−0.26)	0.53 (0.36−0.91)	0.18 (0.13−0.27)	0.12 (0.10−0.17)	0.18 (0.14−0.24)	0.18 (0.15−0.22)	0.17 (0.14−0.18)	0.49 (0.39−0.60)	0.27 (0.22−0.35)	F^{10}	residual error
isopentenyladenosine	0.25 (0.21−0.31)	0.26 (0.19−0.31)	0.25 (0.21−0.32)	0.22 (0.17−0.27)	0.36 (0.30−0.46)	0.45 (0.35−0.54)	0.27 (0.21−0.32)	0.40 (0.19−0.50)	0.34 (0.28−0.37)	F^4	pedigree
ash	5.13 (4.90−5.46)	5.26 (5.06−5.54)	5.20 (5.00−5.45)	5.23 (4.93−5.70)	4.97 (4.80−5.20)	5.12 (4.90−5.41)	5.28 (5.07−5.72)	5.23 (4.92−5.66)	5.00 (4.84−5.20)	none	none[h]

[a]Isoflavones and organic acids expressed as ppm dwt; saccharides and ash expressed as % dwt; phytohormones expressed as ppm dwt with the exception of isopentenyladenosine, which was expessed as ppb dwt. [b]Minimum to maximum of individual replicate values across both sites. [c]Factor to which an analyte shows coefficient of greatest magnitude. [d]Greatest source of variance for given factor. [e]Glucose, fructose, and galactose had >50% of values below assay LOQ. [f]Acetic, lactic, and oxalic acid had >50% of values below assay LOQ. [g]Dihydrozeatin, dihydrozeatin riboside, and *trans*-zeatin riboside had >50% of values below assay LOQ. [h]Not highly correlated with any of the factors that explained 80% of the total variation. [i]Coefficient scores were <0.5 but >0.48.

dx.doi.org/10.1021/jf4032102 |*J. Agric. Food Chem.* 2013, 61, 10807−10815

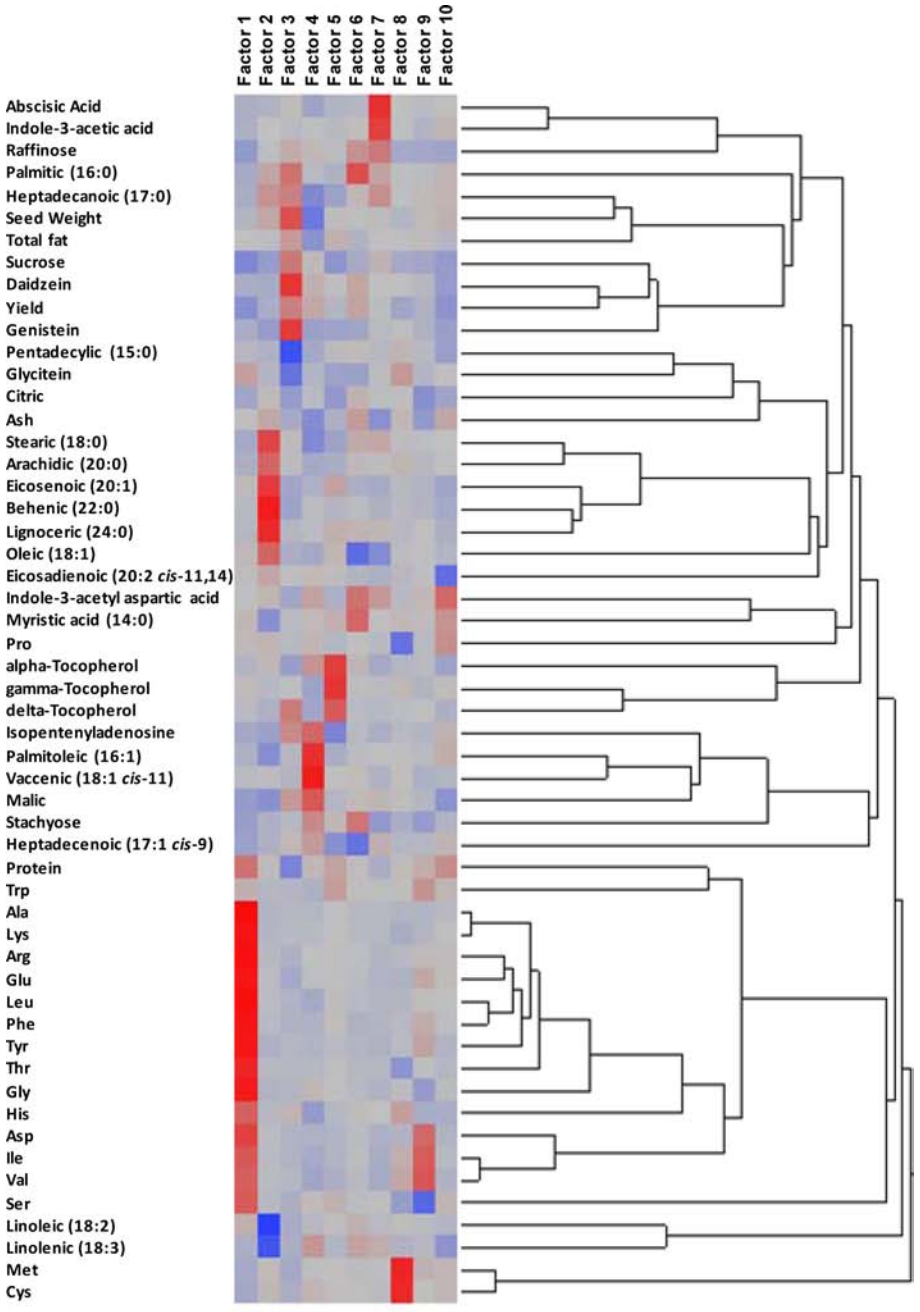

Figure 1. Heatmap and dendrogram of varimax coefficient scores. Dendrograms are schematic representations of multivariate distances with joined segments indicating similarity.

other factor. The largest source of variance in F^3 was pedigree followed by residual error (Table 6). Figure 5 contains box plots to display the variation of values in F^3 due to location and pedigree. This figure also illustrates the launch era-related change in factor F^3 (and consequently of the compositional analytes correlated with F^3). Isoflavone data were consistent with those of Yin and Vyn,[21] who showed daidzein and genistein to be positively correlated with soybean yields. Our study also suggested a negative correlation between yield and glycitein and that genistein may have a stronger correlation to seed size than yield. Overall, on a concentration basis, isoflavone levels varied with soybean yield to a much greater magnitude than other analytes. An inverse relationship between isoflavone content and soybean seed size has been reported.[22] There is little information on the relationship of palmitic acid with soybean yield components. Sucrose has been reported to correlate positively with yield.[23]

Factor F^4: Isopentenyladenosine, Palmitoleic Acid, Vaccenic Acid, and Malic Acid. Factor F^4 explained 7.64% of the total variance in this study and was strongly correlated with palmitoleic acid and vaccenic acid and moderately correlated with isopentenyladenosine and malic acid. A moderate negative correlation was observed for seed weight. The largest source of variance in F^4 was pedigree (Supporting Information, Supplementary Table 3). The unexpected association of these analytes illustrates the value of PVCA in hypothesis generation on potential relationships between different metabolites.

Factor F^5: Tocopherols. Factor F^5 explained 5.15% of the total variance in this study and was strongly or moderately

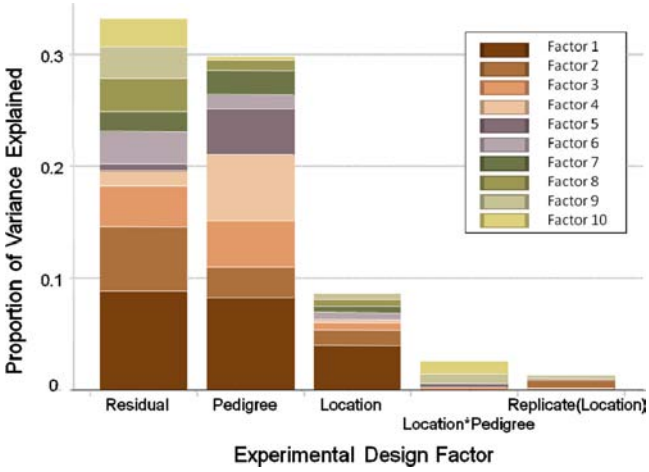

Figure 2. Proportions of variance explained by variance components for each factor. Numbers in the plot represent the cumulative proportions of variation in the data that are attributed to each variance component, totaled over factors.

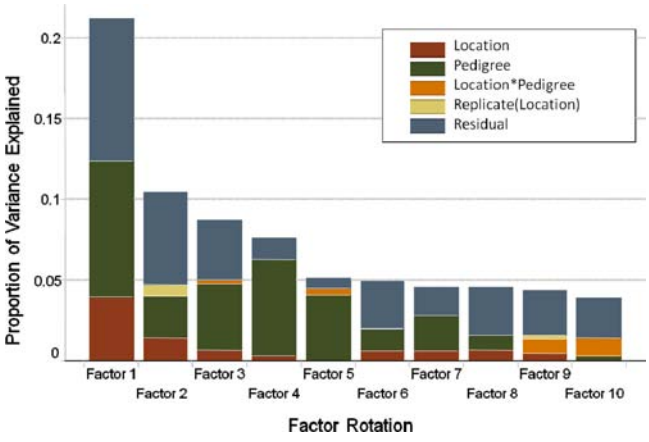

Figure 3. Proportions of variance explained by each factor with relative contributions from each variance component. Numbers in the plot represent the cumulative proportions of variance in the data that are attributed to each factor.

Table 5. Summary of Factor 1 Sources of Variation (21.17%)

source of variance	fraction of variance due to source	PVC contribution
residual error	41.725	8.83
pedigree	39.244	8.31
location	18.760	3.97
rep (location)	0.272	0.06
location × pedigree	0.000	0.00

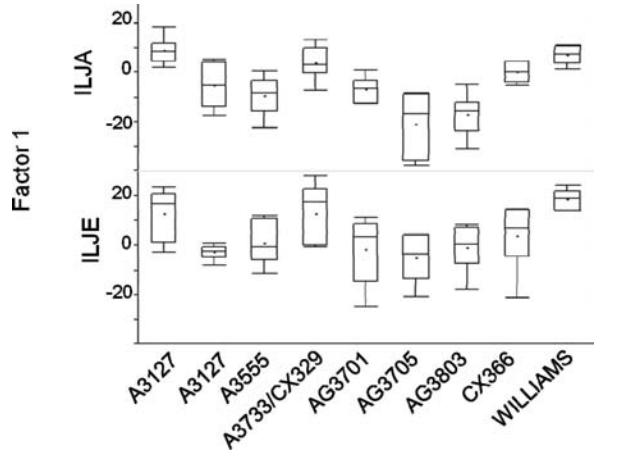

Figure 4. Variation of the first varimax factor F^1 at each site. Filled circles indicate means, and horizontal lines on a box plot, from top to bottom, represent the maximum, third quartile, median, first quartile, and minimum. ILJA and ILJE are site codes for the Jacksonville and Jerseyville sites, respectively.

Table 6. Summary of Factor 3 Sources of Variation (8.65%)

source of variance	fraction of variance due to source	PVC contribution
pedigree	77.296	6.69
residual error	17.851	1.54
location	3.913	0.34
rep (location)	0.606	0.05
location × pedigree	0.335	0.03

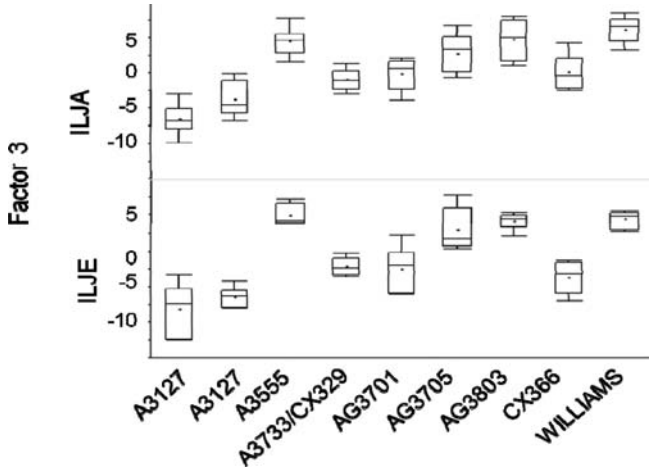

Figure 5. Variation of varimax factor F^3 among varieties at each site. ILJA and ILJE are site codes for the Jacksonville and Jerseyville sites, respectively.

correlated with the tocopherols. The largest source of variance in F^5 was residual error followed by pedigree (Supporting Information, Supplementary Table 4). Levels of tocopherols in soybean seed are known to be affected by environment and germplasm;[24−26] however, in this study, no meaningful trends across launch era were observed.

Concluding Remarks. The continued development of new higher yielding soybean varieties is a major requirement for sustainable agriculture globally. This study demonstrated that decades of varietal development to achieve desired agronomic changes such as improved yield can be associated with compositional changes. These changes appear modest with no nutritional consequences and are much less than variation due to location effects. The use of multivariate analyses such as PVCA allowed a more unified interpretation of the effects of environment, germplasm, and desired phenotypic variation on biochemical systems than can be obtained from customary univariate ANOVA for individual analytes. For example, PVCA showed that several components such as isoflavones, palmitic acid, and sucrose were correlated with each other and could be considered as measurements of one phenomenon that was most strongly associated with changes in yield and soybean launch year. An attractive option of PVCA is that the final results can be phrased in term of percentages and plotted as simple graphics; as such, they are easy to convey to a broad audience.

In summary, varietal development of high-yielding soybean, as represented in this study, has been accompanied by compositional changes, but these are typically modest relative to environmental factors.

■ ASSOCIATED CONTENT

Supporting Information

Additional tables as detailed in the text. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*(G.G.H.) Phone: (314) 694-1268. Fax: (314) 694-8575. E-mail: george.g.harrigan@monsanto.com.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank the Monsanto Agronomy and Sample Management groups for generation and preparation of the samples used in this study and the Monsanto Product Characterization group for the molecular analysis of the seed. We are also grateful to the excellent analytical scientists at the Monsanto Pipeline Analytics and Monsanto Environmental Sciences Technology Center.

■ REFERENCES

(1) Hymowitz, T.; Shurtleff, W. R. Debunking soybean myths and legends in the historical and popular literature. *Crop Sci.* **2005**, *45*, 473−476.

(2) Hymowitz, T.; Harlan, J. R. The introduction of the soybean to North America by Samuel Bowen in 1765. *Econ. Bot.* **1983**, *37*, 371−379.

(3) Ustun, A.; Allen, F. L.; English, B. C. Genetic progress in soybean of the U.S. Midsouth. *Crop Sci.* **2001**, *41*, 993−998.

(4) Durham, D. The United Soybean Board's Better Bean Initiative: building United States soybean competitiveness from the inside out. *AgBioForum* **2003**, *6*, 23−26.

(5) Brown-Guedira, G. L.; Thompson, J. A.; Nelson, R. L.; Warburton, M. L. Evaluation of genetic diversity of soybean introductions and North American ancestors using RAPD and SSR Markers. *Crop Sci.* **2000**, *40*, 815−823.

(6) Thompson, J. A.; Nelson, R. L. Utilization of diverse germplasm for soybean yield improvement. *Crop Sci.* **1998**, *38*, 1362−1368.

(7) Wilcox, J. R. Sixty years of improvement in publicly developed elite soybean lines. *Crop Sci.* **2001**, *41*, 1711−1716.

(8) Specht, J. E.; Hume, D. J.; Kumudini, S. V. Soybean yield potential. A genetic and physiological perspective. *Crop Sci.* **1999**, *39*, 1560−1570.

(9) Wilson, R. F. Seed composition. In *Soybeans: Improvement, Production and Uses*, 3rd ed.; Boerma, H. R., Specht, J. E., Eds.; ASA, CSSA, SSA: Madison, WI, 2004.

(10) Fehr, W. R. Breeding for modified fatty acid composition in soybean. *Crop Sci.* **2007**, *47* (S3), S72−S87.

(11) Clemente, T. E.; Cahoon, E. B. Soybean oil: genetic approaches for modification of functionality and total content. *Plant Physiol.* **2009**, *151*, 1030−1040.

(12) Rotundo, J. L.; Westgate, M. E. Meta-analysis of environmental effects on soybean seed composition. *Field Crops Res.* **2009**, *110*, 147−156.

(13) Harrigan, G. G.; Lundry, D.; Drury, S.; Ridley, W. P.; Riordan, S. G.; Nemeth, M. A. Natural variation in crop composition and the impact of biotechnology. *Nat. Biotechnol.* **2010**, *28*, 402−404.

(14) Li, J.; Bushel, P. R.; Chu, T.-M.; Wolfinger, R. D. Principal variance components analysis: estimating batch effects in microarray gene expression data. In *Batch Effect and Experimental Noise in Microarray Studies: Sources and Solutions*; Scherer, A., Ed.; Wiley: West Sussex, UK, 2009.

(15) Harrison, J. M.; Howard, D.; Malven, M.; Halls, S. C.; Culler, A. H.; Harrigan, G. G.; Wolfinger, R. D. Principal variance component analysis of crop composition data: a case study on herbicide-tolerant cotton. *J. Agric. Food Chem.* **2013**, *61*, 6412−6422.

(16) Berman, K. H.; Harrigan, G. G.; Riordan, S. G.; Nemeth, M. A.; Hanson, C.; Smith, M.; Sorbet, R.; Zhu, E.; Ridley, W. P. Compositions of seed, forage, and processed fractions from insect-protected soybean MON 87701 are equivalent to those of conventional soybean. *J. Agric. Food Chem.* **2009**, *57*, 11360−11369.

(17) Lundry, D. R.; Ridley, W. P.; Meyer, J. J.; Riordan, S. G.; Nemeth, M. A.; Trujillo, W. A.; Breeze, M. L.; Sorbet, R. Composition of grain, forage, and processed fractions from second-generation glyphosate-tolerant soybean, MON 89788, is equivalent to that of conventional soybean (*Glycine max* L.). *J. Agric. Food Chem.* **2008**, *56*, 4611−4622.

(18) Harrigan, G. G.; Stork, L. G.; Riordan, S. G.; Reynolds, T. L.; Ridley, W. P.; Masucci, J. D.; MacIsaac, S.; Halls, S.; Orth, R.; Smith, R. G.; Brown, W. E.; Welsch, M.; Riley, R.; McFarland, D.; Pandravanda, A.; Glenn, K. C. Impact of genetics and environment on nutritional and metabolite components of maize grain. *J. Agric. Food Chem.* **2007**, *55*, 6177−6185.

(19) AOAC. *Official Methods of Analysis of AOAC International*; AOAC International: Gaithersburg, MD, XXXX; Method 986.13.

(20) Mahmoud, A. A.; Natarajan, S. S.; Bennett, J. O.; MacWhinney, T. P.; William, W. J.; Krishnan, H. B. Effect of six decades of selective breeding on soybean protein composition and quality; a biochemical and molecular analysis. *J. Agric. Food Chem.* **2006**, *54*, 3916−3922.

(21) Yin, X.; Vyn, T. J. Relationships of isoflavone, oil, and protein in seed with yield of soybean. *Agron. J.* **2005**, *97*, 1314−1321.

(22) Kim, E.-H.; Ro, H.-M.; Kim, S.-L.; Kim, H.-S.; Chung, I.-M. Analysis of isoflavone, phenolic, soyasapogenol, and tocopherol compounds in soybean [*Glycine max* (L.) Merrill] germplasms of different seed weights and origins. *J. Agric. Food Chem.* **2012**, *60*, 6045−6055.

(23) Wilcox, J. R.; Shibles, R. M. Interrelationships among seed quality attributes in soybean. *Crop Sci.* **2001**, *41*, 11−14.

(24) Britz, S. J.; Kremer, D. F.; Kenworthy, W. Tocopherols in soybean seeds; genetic variation and environmental effects in field-grown crops. *J. Am. Oil Chem. Soc.* **2008**, *85*, 931−936.

(25) Seguin, P.; Turcotte, P.; Tremblay, G.; Pageau, D.; Liu, W. Tocopherols concentration and stability in early maturing soybean genotypes. *Agron. J.* **2009**, *101*, 1153−1159.

(26) Seguin, P.; Tremblay, G.; Pageau, D.; Liu, W. Soybean tocopherol concentrations are affected by crop management. *J. Agric. Food Chem.* **2010**, *58*, 5495−5501.