DOI:10.1021/jf903705y

JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY
ARTICLE

Impact of Genetics and Environment on the Metabolite Composition of Maize Grain

Kirsten Skogerson,[△,‖] George G. Harrigan,*[,†,‖] Tracey L. Reynolds,[§] Steven C. Halls,[#] Martin Ruebelt,[⊥] Alberto Iandolino,[⊥] Anand Pandravada,[⊗] Kevin C. Glenn,[†] and Oliver Fiehn[△,‖]

[†]Product Safety Center, [§]Global Regulatory Pipeline, [#]Biotechnology, Monsanto Company, 800 North Lindbergh Boulevard, St. Louis, Missouri 63167, [⊥]Calgene Campus, Monsanto Company, 1920 Fifth Street, Davis, California 95616, [⊗]Corn Breeding, 12849 Gorman Lane, Woodland, California 95695, and [△]Metabolomics Research Laboratory, 451 Health Sciences Drive, Davis, California 95616. [‖]Authors contributed equally to the work.

This study sought to assess genetic and environmental impacts on the metabolite composition of maize grain. Gas chromatography coupled to time-of-flight mass spectrometry (GC-TOF-MS) measured 119 identified metabolites including free amino acids, free fatty acids, sugars, organic acids, and other small molecules in a range of hybrids derived from 48 inbred lines crossed against two different tester lines (from the C103 and Iodent heterotic groups) and grown at three locations in Iowa. It was reasoned that expanded metabolite coverage would contribute to a comprehensive evaluation of the grain metabolome, its degree of variability, and, in principle, its relationship to other compositional and agronomic features. The metabolic profiling results established that the small molecule metabolite pool is highly dependent on genotypic variation and that levels of certain metabolite classes may have an inverse genotypic relationship to each other. Different metabolic phenotypes were clearly associated with the two distinct tester populations. Overall, grain from the C103 lines contained higher levels of free fatty acids and organic acids, whereas grain from the Iodent lines were associated with higher levels of amino acids and carbohydrates. In addition, the fold-range of genotype mean values [composed of six samples each (two tester crosses per inbred × three field sites)] for identified metabolites ranged from ∼1.5- to 93-fold. Interestingly, some grain metabolites showed a non-normal distribution over the entire corn population, which could, at least in part, be attributed to large differences in metabolite values within specific inbred crosses relative to other inbred sets. This study suggests a potential role for metabolic profiling in assisting the process of selecting elite germplasm in biotechnology development, or marker-assisted breeding.

KEYWORDS: Maize; metabolic profiling; metabolomics; natural variation; *Zea mays*

INTRODUCTION

Maize represents one of the most important crops for the production of food, feed, and biofuel. As a major grower, the United States accounted for 13.1 million bushels or approximately 40% of world maize production from 2007 to 2008 at an aggregate value of over $50 billion (*1*). As a safe and wholesome commodity, corn remains a focus of further intensive selective breeding to improve quality traits. Key nutritional components in maize grain include starch, protein, fiber, and oil (*2*). However, due to its economic value, numerous projects to enhance the food and feed quality of corn grain have also centered on modulating levels of small molecule metabolites. These include attempts to increase α-tocopherol content (*3*, *4*) or decrease levels of the antinutrient phytic acid (*5*−*7*). The metabolite pool size in grain is of low abundance (∼5%) relative to the accumulated biomass of starch, protein, fiber, and oil (*8*, *9*). [For the purposes of this discussion, oil (triglycerides) is arbitrarily considered separate from the remainder of the small molecule metabolite pool of corn grain.] We reasoned that technical developments in nontargeted metabolic profiling would facilitate a comprehensive evaluation of the grain metabolome, its degree of variability, and its relationship to other compositional and agronomic features. This study therefore sought to assess genetic and environmental impacts on the small molecule metabolite composition of maize grain by surveying a range of diverse hybrids grown at three different locations in Iowa. The goal was to (i) assess the range of variation in different metabolites and metabolite classes across the study population at all sites, (ii) identify location and genotypic (tester) effects on relative levels of metabolites, and (iii) determine whether metabolites or metabolite classes were more variable within certain hybrid lines and/or across all samples.

*Author to whom correspondence should be addressed.

 Published on Web 02/16/2010

MATERIALS AND METHODS

Biological Material. Seed of the various maize (*Zea mays* L.) hybrids were planted at three different locations in Iowa and grown under normal regional agronomic practice. These were Cambridge, Huxley, and South Amana. Planting at Cambridge was on April 29, 2005, and planting at Huxley and South Amana on May 5. Ninety-six plots representing 48 inbred lines (with two independent meiotic lines per inbred) crossed with the two different tester lines, representing the C103 and Iodent heterotic groups, were planted per location. Mature grain was collected from five plants for each line selected at random and bulked for metabolite analyses. After coarse-grinding, the samples were stored at −80 °C until required for metabolite analyses.

Sample Preprocessing. Coarsely ground bulk corn samples were subsampled (100 mg) into 2 mL Eppendorf tubes and lyophilized for 24 h at 30−70 mTorr. Lyophilized samples were homogenized to a fine powder with a Retsch ball mill (3 mm diameter steel ball, 4 × 2 min, 25 s^{-1}) and subaliquoted (2 mg) for extraction.

Extraction Method Optimization. Preliminary tests were performed to optimize the extraction protocol. Four extraction solvents were each tested at two temperatures (85 and 4 °C): (1) isopropanol/acetonitrile/water (3:3:2 volume ratio), (2) 50% aqueous acetonitrile, (3) 50% aqueous isopropanol, and (4) methanol/chloroform/water (5:2:2 volume ratio). All extraction solvents yielded reproducible results with median precisions around 20% RSD; however, the hot extractions (85 °C) yielded better recoveries of free fatty acids, lower glucose and fructose levels (due to presumed faster enzyme inactivation), and unchanged or higher amino acid levels (including nonproteinogenic amino acids such as GABA). Given the slightly better recoveries for critical compounds and generally better extraction of organic acids for solvents with higher water concentration, 50% aqueous isopropanol was used in the final extractions.

Sample Extraction. Extraction solvent (50% aqueous isopropanol) was degassed with nitrogen prior to use and then preheated for 15 min at 85 °C; 1.2 mL of hot solvent was added per 2 mg of lyophilized, ground material, and the samples were vortexed vigorously for 10 s and then shaken for 10 min at 85 °C. Samples were centrifuged at 12200*g* for 2 min to pellet debris, and 600 μL of supernatant was removed to a clean Eppendorf tube. Supernatants were concentrated to dryness in a Centrivap cold trap vacuum concentrator at room temperature for 6 h.

Sample Derivatization. Derivatization was performed as described previously (*10*, *11*). In summary, 2 μL of a C8−C30 FAME mixture was used to convert retention times to retention indices. Carbonyl groups were protected by 10 μL of a solution of 20 mg/mL methoxyamine in pyridine at 30 °C for 90 min. Ninety microliters of MSTFA + 1% TMCS was added for trimethylsilylation of acidic protons at 37 °C for 30 min.

GC-MS Analysis and Operating Conditions. GC-MS was performed as described previously (*10*, *11*). In summary, 0.5 μL samples were injected in randomized sequence into a Gerstel CIS cold injection system (Gerstel, Muehlheim, Germany) in splitless mode (vent after 25 s) at 50 °C, which was ramped at 12 °C/s to a final temperature of 250 °C and held for 3 min. Liners were exchanged automatically every 10 samples. Gas chromatography was performed by a 10 m integrated guard column and a 30 m × 0.25 mm i.d. Rtx-5Sil MS separation column with 0.25 μm film at a constant flow of 1 mL/min helium and an oven program ramping from 50 to 330 °C at 20 °C/min. A Leco Pegasus IV time-of-flight mass spectrometer was operated at a transfer line temperature of 280 °C and 70 V electron impact ionization at 250 °C. Mass spectra were acquired at mass resolving power $R = 600$ from m/z 85 to 500 at 20 spectra s^{-1} and 1850 V detector voltage.

Data Processing. After methoximation/trimethylsilylation, GC-TOF MS analysis, and data annotation using the BinBase database, a total of 329 metabolic peaks were reported, after exclusion of artifact peaks and internal standards. At a threshold of at least 50% positive detection in any of the tester classes, 125 of these peaks were annotated as identified metabolites, which accounted for 119 unique metabolites after removal of peaks that are known to occur in two forms after derivatization (e.g., glucose). These 119 metabolites were classified as 26 amino acids, 42 carbohydrates, 17 free fatty acids and related metabolites, 17 hydroxyl acids, and 17 sterols, amines, and miscellaneous compounds (**Tables 1**−**8**). On a cautionary note, quantitation data for sucrose are presented with the caveat that it was generally found to be overloaded due to its high abundance relative to that of other small molecule metabolites. Methionine did not pass the Binbase thresholds but was manually integrated using the instrument's software ChromaTOF. Missing values were replaced using the raw ion trace data and subtracting local noise levels (*10*, *11*). In addition to the known compounds, 204 unidentified metabolic peaks were positively detected in at least 50% of all samples and, of these, 111 were detected in >80% of all samples and can thus be regarded as consistently present in maize grain. All identified metabolites were associated with a PubChem CID (see the Supporting Information).

Table 1. Summary of Statistical Differences in Grain Metabolite Levels

		significantly different ($p < 0.0001$)		
metabolite class	no. of analytes	tester	location	higher tester[a]
free amino acids	26	14	2	C103
sterols, amines, and others	17	6	1	—[b]
organic acids	17	6	0	C103
free fatty acids and related metabolites	17	5	0	Iodent
sugar alchohols	18	5	0	—
mono-, di-, and tri-saccharides	16	1	0	—
sugar acids	8	0	1	—

[a] This refers to the tester where mean levels of analytes within a given metabolite class were generally highest. [b] There was no consistent association of higher mean levels for this metabolite class with a specific tester.

Statistical Analysis. Data obtained for the 116 identified metabolites were analyzed across locations for each metabolite separately using a model of the form

$$y_{ijk} = \mu + s_i + t_j + (st)_{ij} + \varepsilon_{ijk}$$

where y_{ijk} is the kth response for the jth tester at the ith location, μ is the overall mean, s_i is the effect of the ith location, t_j is the effect of the jth tester, $(st)_{ij}$ is the effect of the interaction between the jth tester and the ith location, and ε_{ijk} is the random error.

Observations with studentized residuals ≤ −6 or ≥6 were removed from the analysis data set; this resulted in the removal of 0.3% of the measurements. The statistical model was fit to the data using SAS PROC MIXED. Sample means and ranges for each analyte within a given tester population (C103 or Iodent) at a specific location were calculated on the data set following outlier removal (**Tables 2**−**8**). Means combining the field samples of each genotype (six samples composed of two tester crosses per inbred × three field sites) were also determined. Partial least-squares (PLS) regression analysis was used to examine the relationships between the metabolite profiles and the tester and location data.

RESULTS AND DISCUSSION

Forty-eight inbred maize lines were crossed with two different tester lines, C103 and Iodent, to generate a total of 96 hybrids. These hybrids were grown at three different field locations in Iowa, as described under Materials and Methods. Harvested grain was subjected to GC-MS based metabolite profiling using an optimized extraction method. Ten-fold replicate extraction tests were performed to determine the technical precision of the optimized method. For the identified metabolites, an overall median precision of 18% relative standard deviation (RSD) was determined, validating that the extraction and analysis protocol was suitable for GC-TOF-MS-based metabolite profiling (*11*). As expected, the technical reproducibility was better for more abundant compounds compared to metabolites with lower concentrations. The technical reproducibility was determined as 11%

Table 2. Free Amino Acid Composition[a] of Grain from Corn Grown in the United States in 2005

	Cambridge site		Huxley site		South Amana site		*p* value	
component	C103 mean (range)	Iodent mean (range)	C103 mean (range)	Iodent mean (range)	C103 mean (range)	Iodent mean (range)	genotype	location
alanine	131736 (76443−247193)	158228 (79995−232752)	156892 (71735−255024)	192804 (101521−345656)	153446 (75549−302107)	183543 (108030−360044)	<0.0001	<0.0001
β-alanine	2938 (1271−7810)	3400 (643−7793)	3035 (1015−8336)	3412 (863−8075)	2896 (566−9141)	3462 (742−9487)	0.0269	0.9752
Arg + Orn	8049 (2761−21793)	6891 (2268−13705)	7722 (1914−14954)	6851 (2949−17248)	7752 (3884−13882)	6458 (3236−14086)	0.001	0.671
asparagine	87926 (36095−133837)	117553 (42824−186367)	84452 (36895−151194)	97963 (45313−141392)	95411 (33698−181132)	109083 (42302−169116)	<0.0001	0.0095
aspartate	157517 (50334−253411)	141745 (64780−217432)	146951 (21091−227634)	146991 (65437−199048)	162869 (87734−246610)	149446 (71798−208559)	0.0191	0.1675
citrulline	790 (253−1449)	872 (327−2032)	753 (200−1370)	823 (380−1946)	829 (421−1349)	778 (400−2071)	0.3471	0.6155
cysteine	959 (393−1521)	987 (234−2682)	1038 (265−2039)	1038 (540−1652)	1067 (481−2163)	1132 (388−1825)	0.4402	0.0364
glutamate	91071 (55310−1561001)	91409 (49238−125196)	77363 (38749−114968)	80021 (45637−121175)	86696 (49638−122796)	83194 (56158−115801)	0.9331	<0.0001
glutamine	13517 (4203−39386)	16474 (6581−34096)	11682 (3217−35014)	14760 (6331−32127)	11370 (4905−29816)	16760 (9107−34974)	<0.0001	0.1517
glycine	25911 (14975−47316)	28245 (11488−44557)	27920 (11773−44032)	27554 (10214−43918)	27757 (15939−56258)	28223 (14058−63009)	0.3689	0.699
histidine	10283 (4054−20754)	12441 (4409−23727)	8825 (1050−13838)	9721 (2884−17041)	10321 (1979−20332)	10445 (2068−23973)	0.0248	0.0018
isoleucine	8326 (3925−20531)	10532 (4120−19396)	8654 (3798−16241)	11132 (5428−22887)	8899 (3473−19580)	11842 (5510−21888)	<0.0001	0.1412
leucine	10982 (4525−25251)	15868 (5893−46110)	10247 (4553−19962)	14910 (6864−27576)	11426 (4842−38777)	17510 (7212−39416)	<0.0001	0.0684
lysine	43110 (17708−124565)	48461 (8512−128425)	37674 (7617−99928)	40635 (14845−123552)	37323 (15709−104054)	37360 (13681−113212)	0.2955	0.0265
methionine	4916 (1216−9059)	6390 (2462−11209)	5317 (1195−10370)	6719 (1814−12382)	4938 (1838−14834)	6963 (2633−12161)	<0.0001	0.2121
ornithine	6633 (1788−14929)	8015 (1924−16312)	5716 (1175−12733)	6861 (905−16031)	6889 (1609−14231)	7534 (2903−18624)	0.0041	0.0418
oxoproline	77493 (49155−114056)	76786 (54058−99401)	74139 (35953−116832)	72868 (56415−98515)	74258 (49810−98978)	76988 (54242−99778)	0.8692	0.154
phenylalanine	5037 (2537−15594)	7834 (3500−16987)	4686 (863−9296)	6819 (2689−10405)	5308 (2845−11952)	7771 (3783−15014)	<0.0001	0.0245
proline	702565 (189983−1148654)	712636 (218842−1309576)	635009 (133431−1197325)	719467 (237330−1226485)	535140 (141156−1111180)	698761 (265378−1254883)	0.00014	0.0192
serine	23008 (7592−52346)	26117 (7593−52284)	23295 (7322−46321)	28890 (9745−55976)	20067 (4774−40135)	28740 (12230−79259)	<0.0001	0.407
threonine	7845 (2591−17701)	9563 (3717−19432)	7644 (2585−17156)	10622 (4139−23029)	7442 (2058−16596)	10660 (3595−18331)	<0.0001	0.634
tryptophan	11308 (5070−18381)	14965 (8871−28508)	12554 (5448−22101)	13987 (4059−24806)	13229 (5730−22978)	14847 (6586−24130)	<0.0001	0.2005
tyrosine	41736 (31110−68558)	48802 (34764−87090)	40763 (21127−56214)	44491 (28593−62169)	44236 (31748−67235)	47480 (34124−63081)	<0.0001	0.0121
valine	27084 (14936−46489)	30710 (18587−45817)	28153 (13866−41389)	31880 (23653−49787)	29633 (15210−51710)	33323 (21594−47133)	<0.0001	0.0279
GABA	39259 (6007−91899)	51671 (21741−103770)	40839 (5861−79140)	52627 (15960−108629)	30319 (11211−56113)	49178 (17811−119586)	<0.0001	0.0374
5-hydroxynorvaline	8314 (1860−19372)	6511 (1564−15130)	7623 (1557−18489)	6322 (2212−11107)	8199 (3144−21842)	6003 (2904−11446)	<0.0001	0.6405

[a] Units are ion counts based on data from GC-TOF-MS based metabolic profiling.

RSD for the 44 more abundant compounds (>2000 counts per second) and 21% RSD for the 78 less abundant metabolites (< 2000 counts per second).

The overall data set of identified metabolites was examined with the goal of determining the range, distribution, and determinants of variation in levels of small molecule metabolites in grain from the study corn population. This involved (i) assessing the range of variation in different metabolites and metabolite classes across all lines at all sites, (ii) identifying location and genotypic (tester) effects on relative levels of metabolites, and (iii) determining whether metabolites or metabolite classes were more variable within certain lines (i.e., inbred × tester) and/or across all samples. An evaluation of the identified metabolites (**Tables 2−8**) is presented below.

Free Amino Acids. The metabolic profiling platform allowed measurement of 20 proteinogenic amino acids and 6 other identified amino acids [β-alanine, citrulline, γ-aminobutyric acid (GABA), 5-hydroxynorvaline, ornthine, oxoproline]. The data revealed that the fold-range of values recorded for each free amino acid measured across all samples ranged

Table 3. Free Fatty Acid and Related Metabolite Composition[a] of Grain from Corn Grown in the United States in 2005

	Cambridge site		Huxley site		South Amana site		*p* value	
component	C103 mean (range)	Iodent mean (range)	C103 mean (range)	Iodent mean (range)	C103 mean (range)	Iodent mean (range)	genotype	location
1-monoolein	3911 (1431−9371)	4493 (1509−16435)	3978 (1923−11182)	3819 (1429−13724)	4469 (2597−14744)	3897 (1848−12712)	0.8517	0.5781
1-monostearin	327 (183−573)	295 (185−588)	329 (188−645)	302 (169−536)	370 (179−572)	329 (153−584)	0.0031	0.0067
arachidic acid	13060 (8811−19916)	11943 (5803−19233)	13936 (7397−26849)	12957 (8021−20356)	15020 (8941−23173)	13247 (7140−24899)	0.0008	0.0025
capric acid	9128 (4017−17475)	9114 (5139−17576)	9662 (4163-21028)	9275 (4331−15816)	9851 (2361−18967)	9169 (5009−18326)	0.3162	0.6309
heptadecanoic acid	10841 (7199−15732)	10233 (6191−15828)	11747 (7133−17861)	10849 (7391−17859)	13097 (7801−19369)	10757 (6845−17841)	<0.0001	0.0011
lauric acid	24207 (8853−58752)	25608 (5758−81721)	23091 (4812−52217)	22053 (6399−46217)	25971 (6123−74153)	21593 (8621−66238)	0.4176	0.5199
lignoceric acid	4470 (1750−8687)	3940 (2140−8214)	4688 (1917−9753)	4130 (2020−8313)	5103 (2620−9698)	4359 (1983−7531)	0.0003	0.0362
linoleic acid	31823 (20379−42686)	29925 (14919−49734)	34596 (21224−62126)	30610 (21577−46021)	37926 (20096−57130)	31640 (19253−47765)	<0.0001	0.0011
linolenic acid	8306 (5303−11795)	8321 (4652−13022)	9030 (5182−13836)	8425 (4701−15939)	9938 (4735−15693)	8915 (5810−14209)	0.0279	0.0009
1-monopalmitin	1344 (720−2464)	1142 (567−2400)	1422 (654−2822)	1135 (665−2029)	1444 (821−2742)	1185 (509−1924)	<0.0001	0.4295
myristic acid	3725 (2047−9006)	3468 (2157−6802)	4211 (1766−12200)	3370 (1971−7431)	4345 (644−10716)	3562 (1759−9601)	0.0006	0.2703
octadecanol	1524 (550−6108)	1401 (517−4946)	1576 (430−4138)	1331 (525−3165)	1892 (587−6369)	1495 (377−2767)	0.011	0.0814
oleic acid	13070 (8697−20946)	12771 (7122−31749)	13974 (8615−23705)	13429 (8102−25397)	15352 (8985−23333)	13254 (8134−25218)	0.019	0.0261
palmitic acid	103660 (75486−142679)	99430 (59262−152931)	114044 (68292−193319)	101934 (72594−164874)	122941 (72097−192137)	104187 (71944−162606)	<0.0001	0.0038
pelargonic acid	28085 (11660−51684)	27362 (17895−55149)	30510 (15865−60304)	29289 (16024−60492)	32676 (7864−61580)	29200 (13737−51859)	0.0961	0.0498
stearic acid	783806 (510352−1196727)	737570 (434177−1229199)	857494 (488688−1494742)	768698 (497228−1181785)	915421 (562822−1347358)	773244 (446365−1390347)	<0.0001	0.0087
palmitoleic acid	1439 (256−5689)	1472 (257−4014)	1550 (523−6147)	1480 (457−4735)	1826 (548−6736)	1664 (334−5699)	0.6034	0.1478

[a] Units are ion counts based on data from GC-TOF-MS based metabolic profiling.

from 3.3 (oxoproline) to 22.8 (histidine) (**Table 2**). When range analysis was conducted on the means of the 48 main genotype groups composed of 6 samples each (2 tester crosses per inbred × 3 field sites), ranges fell between 1.5- and 2.6-fold. Statistical analyses revealed that there was a significant ($p < 0.0001$) tester effect on levels of 14 of the free amino acids (**Tables 1** and **2**). The observation of a tester effect implies that, in this study population, levels of free amino acids were influenced by genetic background. A conservative significance value ($p < 0.0001$) was adopted to address concerns of high false discovery rates in profiling experiments measuring multiple variables (*12*, *13*). When a significant ($p < 0.0001$) tester effect was observed, higher mean values were recorded for all free amino acids with the exception of 5-hydroxynorvaline in grain from the Iodent group at all sites. Differences in growing location (admittedly geographically similar) appeared to be a less significant contributor to variation in the levels of free amino acids in this study population, and only two amino acids (alanine and glutamate) showed a significant ($p < 0.0001$) location effect.

Overall, levels of free amino acids were higher in the Iodent group than in the C103 group at all three sites, and location effects were less pronounced than genotypic effects. It is noteworthy that levels of aspartate and glutamate, two of the most abundant free amino acids in maize grain and both directly derived from the TCA cycle, were present at equivalent levels, respectively, across all tester populations and sites.

Free Fatty Acids and Related Metabolites. The metabolic profiling platform allowed measurement of 17 identified free fatty acids and related metabolites. Inspection of the data set revealed that the fold-change of values recorded for each free fatty acid or related metabolites ranged from 1.5 to 3.7 times when measured across the means of the 48 main genotypes. Statistical analyses revealed that there was a significant ($p < 0.0001$) tester effect on the levels of five of these metabolites [heptadecanoic acid, linoleic acid, palmitic acid, stearic acid, and 1-monopalmitin (**Tables 1** and **3**)]. When a significant tester effect was observed, lower mean values were recorded for this fatty acid metabolite class in grain from hybrids in the Iodent group at all sites. Even when statistical significance was not achieved, mean values for most of this fatty acid metabolite class were generally lower in grain from the Iodent group. Differences in growing location appeared to be a less significant contributor to variation in free fatty acid content in this study population, and no free fatty acids showed a significant ($p < 0.0001$) location effect.

Thus, overall, levels of free fatty acids were lower in the Iodent group than in the C103 group at all three sites. There was little impact of the growing location on changes in free fatty acid levels. Free fatty acid levels were found to be uncorrelated to dry mass oil contents (data not shown).

Table 4. Organic Acid Composition[a] of Grain from Corn Grown in the United States in 2005

	Cambridge site		Huxley site		South Amana site		*p* value	
component	C103 mean (range)	Iodent mean (range)	C103 mean (range)	Iodent mean (range)	C103 mean (range)	Iodent mean (range)	genotype	location
α-ketoglutaric acid	260 (44−762)	240 (20−423)	249 (38−697)	271 (79−439)	260 (45−676)	223 (14−486)	0.4340	0.5851
citric acid	199665 (2078−333013)	168058 (99359−331213)	242106 (107442−497070)	182615 (105811−354940)	265070 (141067−496558)	166882 (86838−283185)	<0.0001	0.0043
fumaric acid	7232 (4017−14678)	7301 (4185−13012)	8234 (3277−13646)	8212 (3417−14970)	7670 (3682−13717)	7730 (2847−12379)	0.8949	0.0169
isocitric acid	4382 (0−8614)	3671 (2474−6940)	5058 (2400−7967)	3853 (2283−6435)	5201 (57−9085)	3550 (2153−6015)	<0.0001	0.0678
lactic acid	38051 (19112−71390)	36592 (22372−68669)	42175 (16090−98682)	43673 (960−118422)	41566 (18145−112253)	42560 (19421−67884)	0.855	0.0366
malate	38468 (17480−75280)	41804 (24975−79102)	46275 (17934−104926)	45672 (20096−83870)	44335 (19929−86748)	43748 (18400−69603)	0.6799	0.0228
succinic acid	66851 (5511−406116)	84419 (22341−334279)	93968 (2380−494605)	95758 (28943−621204)	53276 (2613-219098)	70619 (23630−322976)	0.2594	0.0433
2-hydroxyglutaric acid	2905 (1434−4991)	2641 (1665−4525)	3318 (1409−5072)	2839 (1753−6384)	3263 (1881−4981)	2723 (1538−4397)	<0.0001	0.0165
aconitic acid	2263 (481−4504)	1699 (893−3686)	2542 (1232−5499)	1903 (1089−3592)	2841 (1392−5082)	1949 (982−3400)	<0.0001	0.0009
benzoic acid	8358 (3972−15643)	8297 (2862−17718)	9363 (5210−21005)	8870 (4097−17759)	10151 (4667−18529)	8881 (4831−14573)	0.1007	0.0303
dihydroabietic acid	5753 (1777−28296)	6142 (1316−30475)	7714 (1298−33992)	6222 (1283−27649)	7864 (1124−32788)	7691 (1247−25827)	0.6231	0.2262
oxalic acid	275933 (108274−473673)	250311 (161664−494188)	266082 (100387−466536)	247664 (82539−419667)	265878 (112604−530659)	235458 (85655−370741)	0.0085	0.5551
pipecolic acid	21840 (2280−50017)	9633 (1286−26817)	20559 (3465−57966)	11782 (1095−67084)	22026 (4340−61804)	10715 (1707−26901)	<0.0001	0.9175
suberyl glycine	5388 (244−22997)	9557 (1097−38098)	4390 (289−16094)	6327 (92−17060)	5650 (449−16421)	7625 (522−25698)	<0.0001	0.0144
vanillic acid	564 (193−2202)	644 (163−2882)	719 (181−2661)	517 (161−1478)	767 (165−3422)	712 (161−2902)	0.4031	0.2155
4-hydroxycinnamic acid	1705 (376−3301)	1554 (592−2353)	1573 (172−2603)	1489 (711−2557)	1673 (352−2896)	1515 (575−2503)	0.0163	0.3351
quinic acid	931 (428−1512)	924 (476−1955)	960 (234−2676)	931 (400−1427)	939 (232−1584)	888 (288−1960)	0.3961	0.7563

[a] Units are ion counts based on data from GC-TOF-MS based metabolic profiling.

Organic Acids. The metabolic profiling platform allowed measurement of 17 organic acids. Inspection of the data set revealed that the fold-change of values recorded for each organic acid ranged from 1.6 to 12.1 times when measured across the 48 main genotypes. Ranges for 4 of the 17 metabolites exceeded 150-fold when analyzed by single values [citric acid, isocitric acid, lactic acid, and succinic acid (**Table 4**)]. Statistical analyses revealed that there was a significant ($p < 0.0001$) tester effect on the levels of six of the organic acids (**Tables 1** and **4**). Five of the significantly different organic acids were higher in grain from the C103 group. These included citric acid, isocitric acid, aconitic acid, 2-hydroxyglutaric acid, and pipecolic acid. Mean levels of suberyl glycine were higher in the Iodent group. Differences in growing location appeared to be a less significant contributor to variation in organic acid levels, and no metabolite showed a significant ($p < 0.0001$) location effect.

Levels of citric acid, isocitric acid, and aconitic acid were all correlated at $r > 0.80$ when evaluated across the entire sample population, an observation consistent with their close biosynthetic relationship. Other TCA cycle intermediates were not coregulated with isocitric acid. This was further supported by the observation of higher mean levels of succinic acid in the Iodent group, although this did not represent a statistically significant difference ($p > 0.0001$).

The statistically significantly ($p < 0.0001$) increased levels of pipecolic acid in the C103 group implied that, whereas levels of aspartate and lysine show no tester effect, lysine catabolism may be differentially regulated in the two tester populations. There is, to our knowledge, little published literature on levels of pipecolic acid in maize, although at least one report has correlated lysine-ketoglutarate reductase−saccharopine dehydrogenase activity with osmotically induced formation of pipecolic acid in leaf tissue of *Brassica napus* (*14*).

Mono-, Di-, and Trisaccharides. The metabolic profiling platform allowed measurement of 16 mono-, di-, and trisaccharides. Inspection of the data set revealed that the fold-range of values recorded for each mono-, di-, and trisaccharide ranged from 1.7 to 93.4 times when measured across the 48 main genotypes. Overall, there was no consistent pattern observed for levels of the mono-, di-, and trisaccharides in this study population. Statistical analyses revealed that there was a significant ($p < 0.0001$) tester effect for only one of these metabolites (glucoheptulose) and a significant location effect for none (**Tables 1** and **5**). Mean levels of trehalose at each site were markedly higher in the grain of the Iodent group, but did not reach statistical significance ($p > 0.0001$), an observation most likely attributable to the high degree of natural variation observed for levels of this metabolite. The maximum range difference for all sugars when measured across the 48 main genotypes was found for trehalose at 93.4-fold. A closer

Table 5. Mono-, Di-, and Trisaccharide Composition[a] of Grain from Corn Grown in the United States in 2005

	Cambridge site		Huxley site		South Amana site		*p* value	
component	C103 mean (range)	Iodent mean (range)	C103 mean (range)	Iodent mean (range)	C103 mean (range)	Iodent mean (range)	genotype	location
arabinose	15903 (9226−22851)	16626 (10188−26069)	16816 (7445−30003)	15649 (6359−33832)	18967 (11428−29595)	16275 (10688−22823)	0.0379	0.0345
β-gentiobiose	2108 (204−4582)	1960 (635−5522)	1856 (602−6205)	1886 (361−3766)	1779 (456−4786)	1886 (188−6693)	0.9769	0.3955
cellobiose	1864 (536−5123)	2062 (576−7219)	1875 (296−6917)	1881 (493−3889)	1837 (607−5489)	1901 (276−4028)	0.4217	0.7532
fructose	141478 (42184−289682)	135316 (20676−256059)	119375 (15427−259287)	126815 (45546−215545)	108434 (12256−254268)	126444 (39028−244740)	0.3048	0.0202
glucoheptulose	360 (126−606)	437 (130−943)	330 (155−709)	422 (123−758)	396 (178−986)	431 (178−673)	<0.0001	0.1810
glucose	136765 (57554−268375)	140469 (60271−253590)	116950 (34424−231687)	132470 (63820−236729)	114780 (25413−234830)	142281 (57181−237810)	0.0036	0.0951
inulotriose	1731 (482−3907)	1786 (385−3397)	1856 (368−3669)	1832 (380−4507)	2198 (829−6200)	2353 (489−7935)	0.8246	0.0008
levanbiose	2329 (497−8244)	2558 (826−6516)	2543 (956−6266)	2388 (98−5582)	2484 (609−6172)	2638 (1021−5848)	0.6591	0.8393
melezitose	931 (337−2214)	809 (189−2165)	770 (210−2563)	792 (293−2609)	936 (256−3068)	989 (369−3793)	0.8028	0.0606
melibiose	4506 (716−10037)	4102 (666−9988)	4371 (822−7905)	4027 (380−9341)	4352 (2196−8582)	4704 (933−13805)	0.5482	0.4496
palatinose	1471 (671−2943)	1325 (286−2499)	1302 (481−3656)	1343 (98−3764)	1353 (470−2465)	1380 (576−2756)	0.6756	0.6190
raffinose	42699 (11241−82106)	46328 (5597−177228)	48276 (17043−89547)	38893 (13907−152902)	53596 (22906−116039)	43230 (16826−82195)	0.0506	0.3069
sucrose	416858 (514−778058)	418374 (185294−726533)	465066 (113−796088)	418363 (53883−792995)	495265 (77793−1299953)	416156 (111*-749673)	0.0258	0.2382
trehalose	2381 (141−16342)	8981 (302−82495)	4626 (141−77826)	11070 (281−105057)	6171 (153−93698)	16119 (183−118803)	0.0011	0.1536
xylose	3326 (1615−7912)	3436 (1262−6343)	3139 (1277−6198)	3825 (2002−8107)	3410 (1681−6215)	3431 (1123−7373)	0.0266	0.7989
maltose	1211 (319−4296)	1277 (495−6108)	1122 (139−4944)	1374 (289−7299)	1308 (170−7099)	1266 (394−6143)	0.4932	0.9583

[a] Units are ion counts based on data from GC-TOF-MS based metabolic profiling.

inspection, however, showed dramatic differences (> 800-fold) for trehalose levels across all samples (**Figure 1**). This variability appears to be primarily due to extreme differences within a limited number of sample groups (4 of 48 groups, see **Figure 1**), whereas most sample groups show minimal and consistent levels of trehalose. Although considered to be an osmoprotectant with a potential role in facilitating drought tolerance, trehalose is typically present in only very low amounts in most plants (*15*, *16*). There is a growing body of evidence that links trehalose and its intermediates to sugar sensing, and it is, thus, an important metabolite that is potentially involved in regulating overall growth rates (*17*−*19*).

Sugar Alcohols. The metabolic profiling platform allowed measurement of 18 sugar alcohols. Inspection of the data set revealed that the fold-range of values recorded for each sugar alcohol ranged from 1.7 to 78.9 when measured across the means of the 48 main genotypes. The fold-range of individual samples for mannitol was 842 and that for ribitol, 6234. No immediate explanation is offered for these values, but it was noted that levels of these sugar alcohols are strongly correlated with trehalose levels at $r_{xy} > 0.8$ (see also **Figure 1**), similarly as previously reported in potato leaves (*20*). These data suggest that there are physiological or biochemical factors that coregulate levels of sugar alcohols and trehalose in different plant species and plant organs. These factors could include the role of sugar alcohols as osmoprotectants, although there was no specific stress condition noted during the field trials or the former potato study (*20*).

Statistical analyses revealed that there was a significant tester effect for five of the sugar alcohols (**Tables 1** and **6**). For the five metabolites, where a significant ($p < 0.0001$) tester effect was observed, higher mean values were recorded for threitol, erythritol, and conduritol-β-epoxide in grain from the Iodent group at all sites, and higher mean values were recorded for galactinol and isogalactinol 2 in the C103 group.

Sugar Acids. The metabolic profiling platform allowed measurement of eight sugar acids. Inspection of the data set revealed that the fold-range of values recorded for each sugar acid ranged from 1.7 to 3.3 times when measured across the 48 main genotype means. Statistical analyses revealed that there was a significant ($p = 0.0002$) tester effect for only one of these metabolites (glucuronic acid) and a significant ($p < 0.0001$) location effect for one other (threonic acid) (**Tables 1** and **7**). Mean levels of glucuronic acid were higher in hybrids derived from the Iodent group at all three sites.

Sterols, Amines, and Other Compounds. The metabolic profiling platform allowed measurement of 17 sterols, amines, and other compounds. Inspection of the data set revealed that the fold-range of values recorded for each of these miscellaneous metabolites ranged from 1.7 to 11.3 times when determined across the 48 main genotypes means. Statistical analyses revealed that there was a significant ($p < 0.0001$) tester effect for six of these metabolites (**Tables 1** and **8**). For the metabolites where a significant tester effect was observed, higher mean values were recorded for allantoin, putrescine, guanine,

Table 6. Sugar Alcohol Composition[a] of Grain from Corn Grown in the United States in 2005

	Cambridge site		Huxley site		South Amana site		*p* value	
component	C103 mean (range)	Iodent mean (range)	C103 mean (range)	Iodent mean (range)	C103 mean (range)	Iodent mean (range)	genotype	location
arabitol	2396 (1011−4113)	2899 (1025−5361)	2640 (426−7449)	2788 (783−6142)	2201 (544−5048)	2284 (1247−4400)	0.0366	0.0016
conduritol-β-epoxide	301 (108−647)	389 (128−637)	338 (55−649)	427 (120−683)	344 (167−890)	460 (131−897)	<0.0001	0.0233
erythritol	7684 (1722−13600)	9882 (4012−23901)	7756 (942−23788)	11018 (4177−22863)	6757 (2172−15878)	11240 (3068−32564)	<0.0001	0.5415
isogalactinol 1	3204 (515−11058)	2523 (445−8462)	3486 (593−17554)	2660 (352−7928)	4064 (574−14447)	2931 (631−13777)	0.0108	0.3052
galactinol	77828 (11287−187423)	48990 (10595−150281)	81832 (9997−197961)	47118 (5947−134747)	104681 (36985−251222)	52916 (7661−114773)	<0.0001	0.0171
isogalactinol 2	73455 (26863−129518)	59127 (10682−119502)	79504 (37044−139682)	56059 (26505−154362)	85991 (40498−147365)	63546 (26167−110309)	<0.0001	0.0356
glycerol	37645 (23700−93967)	40133 (22148−121781)	42087 (17740−127636)	41667 (22811−104928)	39576 (22817−65345)	40677 (20929−77498)	0.5482	0.3874
glycerol-3-galactoside	1147 (205−3390)	952 (304−1802)	1150 (58−2678)	1061 (105−2880)	1496 (110−3559)	938 (323−3464)	0.001	0.2555
glycerol-α-phosphate	11052 (4718−18538)	10590 (4907−17566)	12006 (5247−26062)	10334 (6311−17736)	13591 (6021−26094)	11158 (4606−19181)	0.0002	0.0048
inositol *myo*-	78463 (48128−109416)	79726 (51058−107163)	81679 (43991−133266)	79945 (53314−153115)	88054 (46861−123120)	80386 (48147−114589)	0.1559	0.0824
lyxitol	1245 (571−7776)	1332 (511−3131)	1467 (544−8149)	1739 (560−7879)	1288 (578−3428)	1449 (572−3995)	0.1448	0.0844
mannitol	67730 (5486−369433)	54765 (5938−197915)	38868 (2965−157530)	51491 (1307−211097)	60651 (1829−276425)	74520 (8391−369043)	0.5313	0.0322
ribitol	5228 (12−36839)	8486 (260−44866)	5477 (15−45409)	11560 (338−72706)	6453 (83−71786)	11167 (43−74812)	0.001	0.4832
sorbitol	224171 (40342−595540)	242126 (40358−534997)	184334 (25489−781634)	254969 (55823−504754)	171890 (10232−618488)	227435 (35462−690905)	0.0017	0.191
threitol	1401 (580−2898)	1752 (816−4991)	1459 (548−3754)	1897 (1053−3684)	1341 (750−3310)	1771 (970−3526)	<0.0001	0.3321
xylitol	712 (223−1168)	843 (421−2132)	755 (345−1830)	853 (359−1935)	758 (292−1326)	829 (150−1663)	0.0032	0.8196
isogalactinol 3	2134 (380−8137)	1710 (275−6035)	1880 (374−9368)	1752 (536−3835)	2189 (445−4004)	2219 (512−8383)	0.3772	0.2432
cellobiotol	1719 (224−3512)	1770 (378−2991)	1644 (132−3846)	1738 (751−3057)	1798 (764−4003)	1857 (165−3908)	0.4453	0.4359

[a] Units are ion counts based on data from GC-TOF-MS based metabolic profiling.

Table 7. Sugar Acid Composition[a] of Grain from Corn Grown in the United States in 2005

	Cambridge site		Huxley site		South Amana site		*p* value	
component	C103 mean (range)	Iodent mean (range)	C103 mean (range)	Iodent mean (range)	C103 mean (range)	Iodent mean (range)	genotype	location
galactonic acid	1190 (594−2948)	1250 (583−1995)	1254 (360−2265)	1154 (479−1996)	1090 (459−1604)	1133 (707−2339)	0.9823	0.0597
galacturonic acid	8421 (4966−15561)	8116 (3966−15340)	7758 (4307−13118)	7263 (2816−11921)	8405 (4994−13421)	7585 (4386−11763)	0.0293	0.043
gluconic acid	1573 (288−3207)	1483 (381−3055)	1417 (319−3489)	1518 (489−5550)	1054 (292−3088)	1233 (473−3943)	0.4829	0.0009
glucuronic acid	5140 (2715−9143)	6055 (3407−12061)	4912 (2032−8832)	5383 (2513−11518)	4379 (1713−7759)	5220 (2252−8173)	0.0002	0.005
glyceric acid	2103 (630−3447)	2169 (902−3748)	2284 (873−4478)	2327 (1074−4888)	2085 (748−2983)	2082 (535−3289)	0.6248	0.032
ribonic acid	1226 (467−2132)	1279 (399−2565)	1358 (487−2680)	1247 (389−2268)	1505 (457−2871)	1440 (503−2702)	0.511	0.011
saccharic acid	1863 (540−4395)	2010 (861−4138)	2030 (878−4318)	1816 (789−5223)	1707 (843−3849)	1693 (928−4306)	0.7466	0.031
threonic acid	1371 (672−1919)	1448 (830−2428)	1634 (669−2996)	1700 (1128−2637)	1627 (765−3275)	1623 (982−2810)	0.3538	<0.0001

[a] Units are ion counts based on data from GC-TOF-MS based metabolic profiling.

guanosine, and phosphoric acid in grain from the Iodent group at all sites and higher mean values were recorded for tocopherol in grain of the C103 group. Differences in growing location appeared to be a less significant contributor

Table 8. Sterols, Amines, and Other Compounds[a] in Grain from Corn Grown in the United States in 2005

	Cambridge site		Huxley site		South Amana site		*p* value	
component	C103 mean (range)	Iodent mean (range)	C103 mean (range)	Iodent mean (range)	C103 mean (range)	Iodent mean (range)	genotype	location
allantoin	4813 (1923−7826)	6329 (3158−10670)	4949 (1383−9472)	5833 (2062−9774)	5133 (1259−8270)	5645 (2473−10043)	<0.0001	0.6919
hydroxylamine	119659 (30430−220216)	110325 (21564−256468)	126878 (18381−289055)	104751 (15860−202495)	136717 (46507−228333)	111017 (29020−220403)	0.002	0.4169
putrescine	9924 (3091−36175)	16431 (4612−35019)	7928 (1586−25291)	12704 (2165−41808)	7929 (2038−40782)	13612 (5698−43264)	<0.0001	0.0087
2,5-dihydroxypyrazine	1661 (673−3114)	1492 (552−2664)	1738 (930−3085)	1554 (882−3257)	1765 (917−3790)	1772 (965−3548)	0.0548	0.0298
spermidine	7024 (693−18628)	8933 (1959−19284)	6209 (423−13252)	6933 (789−14168)	5684 (258−13277)	6720 (423−12504)	0.0016	0.0005
urea	42590 (15418−148631)	39304 (14297−191399)	51825 (14296−242441)	47088 (12275−200491)	50937 (13904−181103)	61898 (13514−177479)	0.8472	0.047
β-sitosterol	27616 (10512−47916)	26794 (5122−53232)	29518 (13529−52813)	26864 (12883−61892)	32039 (19192−51455)	31324 (15287−49377)	0.186	0.0014
stigmasterol	522 (242−864)	488 (225−857)	481 (191−873)	475 (175−1068)	590 (324−1043)	548 (160−921)	0.1662	0.0006
γ-tocopherol	4128 (1611−7782)	4764 (1020−13931)	4484 (1923−8118)	4717 (1916−8604)	4735 (1528−9914)	5408 (2633−9182)	0.0102	0.0277
α-tocopherol	718 (309−1420)	597 (285−1121)	698 (302−1364)	596 (245−1255)	774 (397−1428)	644 (255−1050)	<0.0001	0.1241
adenosine-5-monophosphate	558 (260−864)	589 (269−1308)	588 (244−915)	627 (287−1230)	665 (325−1043)	709 (334−1216)	0.0812	<0.0001
guanine	892 (186−2327)	1447 (389−2490)	976 (204−1979)	1642 (634−2893)	710 (263−1604)	1585 (764−2682)	<0.0001	0.0329
guanosine	229 (132−430)	298 (134−621)	235 (121−414)	280 (61−459)	244 (135−487)	294 (139−564)	<0.0001	0.507
uracil	2549 (439−6221)	2439 (335−5181)	2565 (1053−5002)	2441 (638−5119)	2614 (929−5267)	2509 (754−4491)	0.3289	0.8734
butane-2,3-diol	10108 (2423−38663)	10170 (3704−30950)	10549 (2360−30425)	11392 (2722−31803)	9849 (2303−41781)	10900 (3952−43869)	0.4341	0.7062
hydroxycarbamate	7316 (923−14523)	6720 (1221−16841)	7242 (1502−17279)	5988 (959−11284)	8172 (3167−19023)	6563 (1046−13562)	0.0046	0.307
phosphoric acid	326732 (130970−515442)	404597 (192029−617176)	322893 (114574−582804)	394735 (257349−549795)	314983 (173761−520931)	400100 (11658−583622)	<0.0001	0.7924

[a] Units are ion counts based on data from GC-TOF-MS based metabolic profiling.

to variation in these metabolites, and only adenosine-5-monophosphate showed a significant ($p < 0.0001$) location effect.

Overall Metabolic Phenotype. A pronounced tester effect was observed for the overall metabolite composition of the grain population. In hybrids of the Iodent group, most free amino acids and a range of carbohydrates were found at higher relative levels than recorded for the C103 population. Conversely, when a significant difference ($p < 0.0001$) was observed, several organic acids and free fatty acids were found at higher levels in the C103 population. In total, 37 of the identified metabolites showed a significant ($p < 0.0001$) tester effect. These differences in hybrid metabolic phenotypes were confirmed by a clear clustering of metabolite profiles using partial least squares multivariate statistics (**Figure 2**). Differences in growing location appeared to be a less significant contributor to variation in the metabolic phenotype.

Outlier Effects and Non-normal Distribution. Interestingly, the levels of a small number of metabolites when measured across all samples were not represented by a normal distribution but rather showed a tailed distribution (see **Figure 3** for illustrative examples). Some metabolites, such as putrescine, methionine, and trehalose, were characterized by maximum values 20-fold higher than the average value as measured across all samples. In some cases, non-normal distributions of metabolite levels could be attributed to large differences in metabolite values within specific inbred samples as discussed earlier for trehalose and ribitol. **Figure 4** presents a similar consideration for a different metabolite class, the amino acids lysine and methionine. When the different maize lines were summarized into the 48 main groups, some lines such as 46 were remarkably different from all other lines in level and distribution of both of these two amino acids. Interestingly, this was accompanied by similarly different values for the biosynthetically related pipecolic acid (**Figure 1**), yet when even higher values of pipecolic acid were found in another set, 15, there was no corresponding pattern in lysine and methionine.

As a further illustration of metabolites showing differential variation between inbred sets, 5-hydroxynorvaline represents another metabolite where there is considerable variability between metabolite levels within each inbred set whereas oleic acid represents a metabolite where variation appears to be more conserved (**Figures 1** and **4**). Oleic acid and other free fatty acids showed less genotypic variance; sugar alcohols and trehalose (**Figure 1**) were markedly different from all other compounds.

Concluding Remarks. Nontargeted metabolic profiling has been applied to grain collected from a maize population derived from 48 inbred lines crossed against two different tester lines and grown at three locations in Iowa. Results from metabolic profiling can be summarized as follows:

(1) Most differences were related to genotype, not location, although this may be a function of the limited geographical

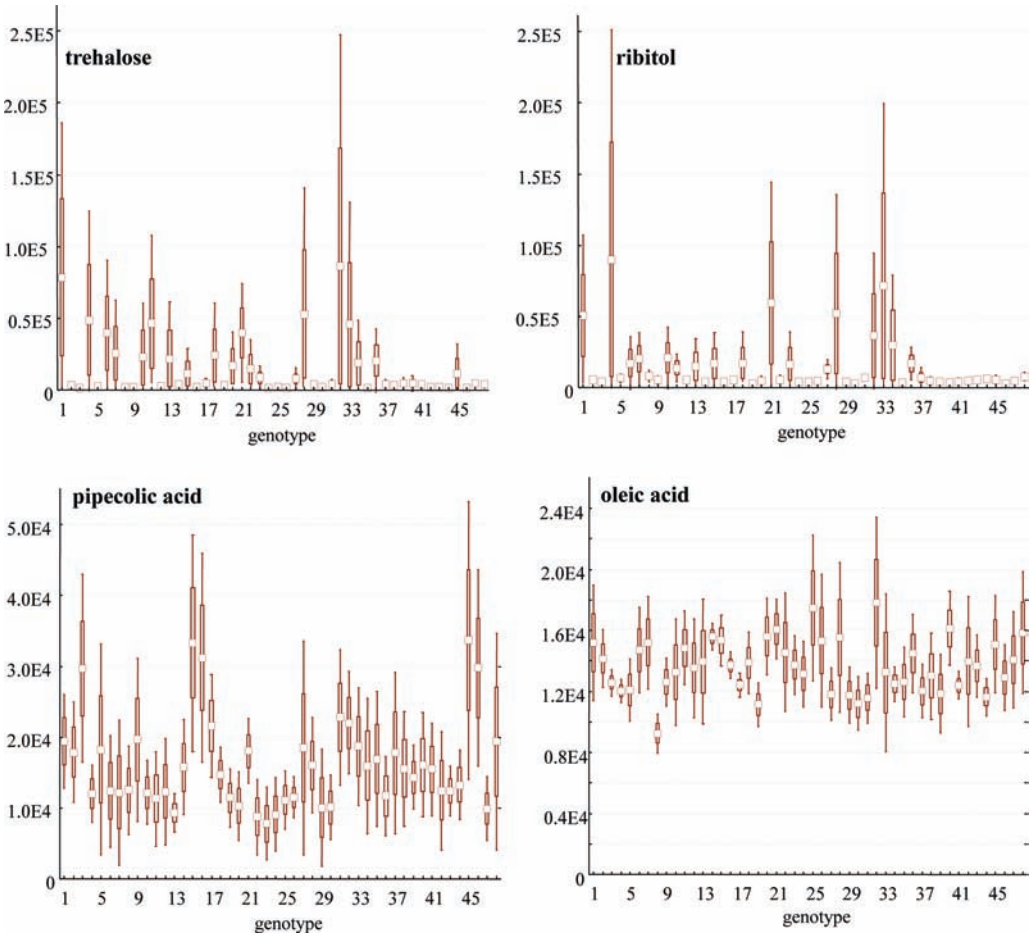

Figure 1. Means and 1.96 × SE ranges of selected metabolites within a given genotype [composed of six samples each (two tester crosses per inbred × three field sites)].

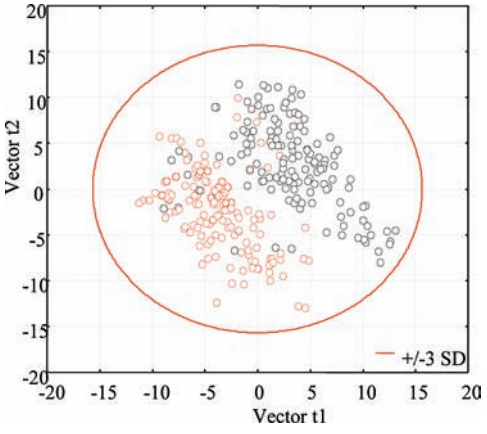

Figure 2. Partial least squares multivariate statistics demonstrating clustering of metabolite profiles according to tester cross (red = C103, black = Iodents). Vector t1 described 11.8% of the overall metabolite variability, and vector t2 accounted for 8.7% of the total variance.

variation represented in this study. Different metabolic phenotypes can be associated with the two different tester populations; C103 lines were generally associated with higher levels of free fatty acids and organic acids, and the Iodent group was associated with higher levels of amino acids and carbohydrates.

(2) Across all metabolites, the fold-range of mean values ranged from ~1.5- to 93-fold. Levels of organic acids, sugars, and sugar alchohols were, broadly speaking, more variable than other classes such as that for free amino acids and fatty acids. Some highly variable metabolites have been associated with regulatory functions such as trehalose and putrescine.

(3) Some metabolites showed a normal distribution, whereas others showed a non-normal distribution. Non-normal distributions could, at least in part, be attributed to large differences in metabolite values within specific inbred crosses relative to other inbred sets.

Thus, metabolic profiling has established that the small molecule metabolite pool is extremely variable across a diverse maize population. In addition to highlighting the degree of variability and the potential relationship of discrete metabolite classes to each other (such as the inverse relationship of free amino acids and fatty acids), it has provided some insights in understanding the nature and range of metabolite variability. The study population was constrained by limited geographical variability, and it is reasonable to suggest that environmental contributions to metabolite variability can be greater than indicated in this study. Although this preliminary investigation proved useful in defining biochemical changes associated with genotypic and phenotypic diversity in this study population, further targeted studies on a more geographically diverse population will be required to further understand natural variability in maize metabolomes and how best to implement the use of metabolomics as a tool in plant breeding.

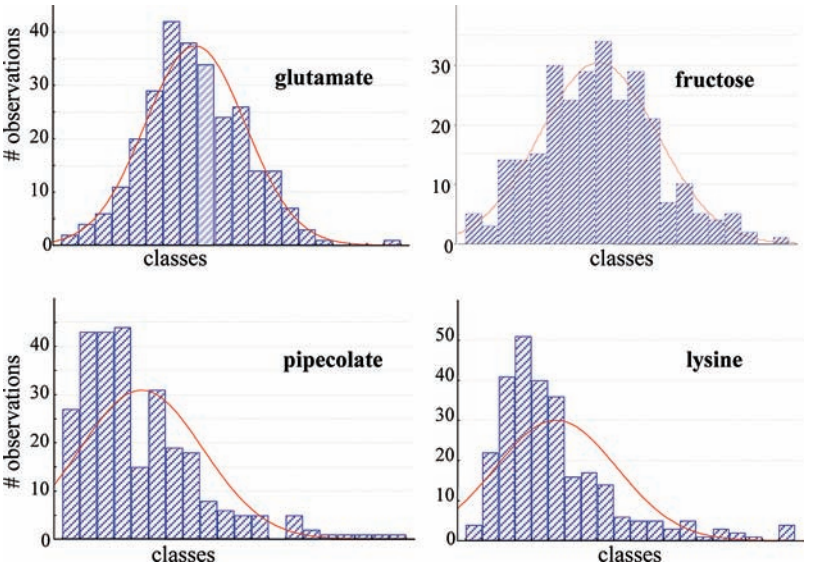

Figure 3. Frequency distribution histograms for four representative metabolites showing normal distributions (upper panel) or tailed distributions (lower panel) when sorted into 20 classes.

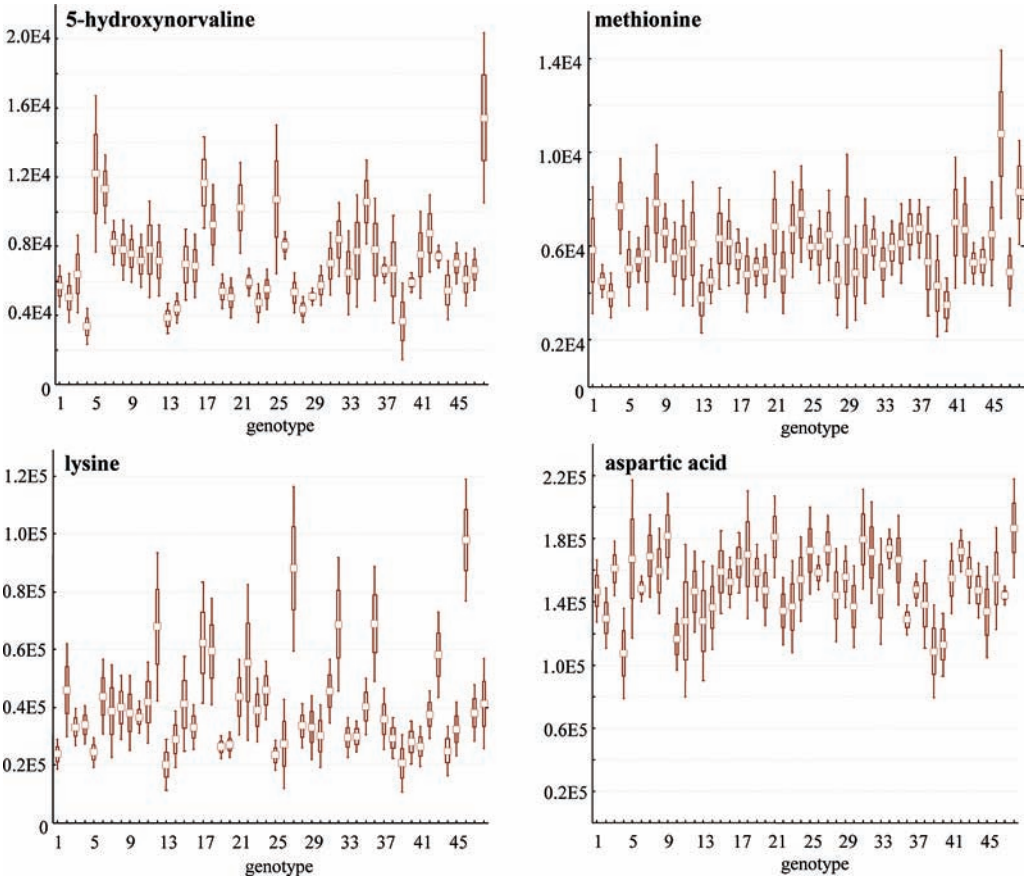

Figure 4. Means and 1.96 × SE ranges of selected metabolites within a given genotype [composed of six samples each (two tester crosses per inbred × three field sites)].

Supporting Information Available: PubChem CID numbers corresponding to **Tables 2−8**. This material is available free of charge via the Internet at http://pubs.acs.org.

LITERATURE CITED

(1) National Corn Growers Association, 2009; www.ncga.org.

(2) OECD. Consensus Document on Compositional Considerations for New Varieties of Maize (*Zea mays*): Key Food and Feed Nutrients, Anti-Nutrients and Secondary Plant Metabolites; Paris, France, 2002.

(3) Rocheford, T. R.; Wong, J. C.; Egesel, C. O.; Lambert, R. J. Enhancement of vitamin E levels in corn. *J. Am. Coll. Nutr.* **2002**, *21* (3 Suppl.), 191S–198S.

(4) Egesel, C. O.; Wong, J. C.; Lambert, R. J.; Rocheford, T. R. Combining ability of maize inbreds for carotenoids and tocopherols. *Crop Sci.* **2003**, *43* (3), 818–823.

(5) Raboy, V.; Gerbasi, P. F.; Young, K. A.; Stoneberg, S. D.; Pickett, S. G.; Bauman, A. T.; Murthy, P. P.; Sheridan, W. F.; Ertl, D. S. Origin and seed phenotype of maize low phytic

acid 1-1 and low phytic acid 2-1. *Plant Physiol*. **2000**, *124* (1), 355–368.
(6) Raboy, V. Seeds for a better future: 'low phytate' grains help to overcome malnutrition and reduce pollution. *Trends Plant Sci*. **2001**, *6* (10), 458–462.
(7) Shi, J.; Wang, H.; Schellin, K.; Li, B.; Faller, M.; Stoop, J. M.; Meeley, R. B.; Ertl, D. S.; Ranch, J. P.; Glassman, K. Embryo-specific silencing of a transporter reduces phytic acid content of maize and soybean seeds. *Nat. Biotechnol*. **2007**, *25* (8), 930–937.
(8) Watson, S. A. Description, development, structure and composition of the corn kernel. In *Corn Chemistry and Technology*; White, P., Johnson, L., Eds.; American Association of Cereal Chemists: St. Paul, MN, 2003; pp 1−38.
(9) Harrigan, G. G.; Stork, L. G.; Riordan, S. G.; Reynolds, T. L.; Ridley, W. P.; Masucci, J. D.; Macisaac, S.; Halls, S. C.; Orth, R.; Smith, R. G.; Wen, L.; Brown, W. E.; Welsch, M.; Riley, R.; McFarland, D.; Pandravada, 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.
(10) Fiehn, O. Extending the breadth of metabolite profiling by gas chromatography coupled to mass spectrometry. *TrAC−Trends Anal. Chem*. **2008**, *27* (3), 261–269.
(11) Fiehn, O.; Wohlgemuth, G.; Scholz, M.; Kind, T.; Lee do, Y.; Lu, Y.; Moon, S.; Nikolau, B. Quality control for plant metabolomics: reporting MSI-compliant studies. *Plant J*. **2008**, *53* (4), 691–704.
(12) Broadhurst, D. I.; Kell, D. B. Statistical strategies for avoiding false discoveries in metabolomics and related experiments. *Metabolomics* **2006**, *2* (4), 171–196.
(13) Lay, J. O.; Borgmann, S.; Liyanage, R.; Wilkins, C. L. Problems with the "omics". *TrAC−Trends Anal. Chem*. **2006**, *25* (11), 1046–1056.
(14) Moulin, M.; Deleu, C.; Larher, F.; Bouchereau, A. The lysine-ketoglutarate reductase−saccharopine dehydrogenase is involved in the osmo-induced synthesis of pipecolic acid in rapeseed leaf tissues. *Plant Physiol. Biochem*. **2006**, *44* (7−9), 474–482.
(15) Karim, S.; Aronsson, H.; Ericson, H.; Pirhonen, M.; Leyman, B.; Welin, B.; Mantyla, E.; Palva, E. T.; Van Dijck, P.; Holmstrom, K. O. Improved drought tolerance without undesired side effects in transgenic plants producing trehalose. *Plant Mol. Biol*. **2007**, *64* (4), 371–386.
(16) Ramon, M.; Rolland, F. Plant development: introducing trehalose metabolism. *Trends Plant Sci*. **2007**, *12* (5), 185–188.
(17) Avonce, N.; Leyman, B.; Thevelein, J.; Iturriaga, G. Trehalose metabolism and glucose sensing in plants. *Biochem. Soc. Trans*. **2005**, *33* (Part 1), 276–279.
(18) Wingler, A. The function of trehalose biosynthesis in plants. *Phytochemistry* **2002**, *60* (5), 437–440.
(19) Rolland, F.; Baena-Gonzalez, E.; Sheen, J. Sugar sensing and signaling in plants: conserved and novel mechanisms. *Annu. Rev. Plant Biol*. **2006**, *57*, 675–709.
(20) Weckwerth, W.; Loureiro, M. E.; Wenzel, K.; Fiehn, O. Differential metabolic networks unravel the effects of silent plant phenotypes. *Proc. Natl. Acad. Sci. U.S.A*. **2004**, *101* (20), 7809–7814.

Received for review October 22, 2009. Revised manuscript received January 20, 2010. Accepted January 25, 2010.