Original Paper **Open Access**

Characterization of maize germplasm: comparison of morphological datasets compiled using different approaches to data recording

John R Law[1], Steven R Anderson[2], Elizabeth S Jones[2], Barry K Nelson[2], Enver Mulaosmanovic[2], J Stephen Smith[2]*

[1]John Law Activities, 1 Willow Close, Little Paxton, St. Neots, Cambridgeshire, PE19-6JH, UK
[2]Pioneer Hi-Bred, 7300 NW62nd Ave., Johnston, Iowa, 50131, USA
*Corresponding author: E-mail: stephen.smith@pioneer.com

Abstract

Different methods are used to report morphological data to evaluate the eligibility criterion of distinctness for Plant Variety Protection (PVP). We compared morphological data for 152 maize (*Zea mays* L.) inbred lines generated according to both US (metric data) and EU data (summarized by reference to "note" scores from check inbreds). While modal notes were very similar for US and EU data, on average only 42.8% of the characteristics had identical scores. Associations among inbreds on the basis of comparisons of single or multiple characteristics were very similar regardless of data type. We conclude that maize morphological data collected according to either of these protocols cannot meaningfully be combined into a single dataset. However, each data type, provided an equivalent mechanism to test for distinctness. These results cause us to more fundamentally question the degree of reliance that should be placed upon morphological data for PVP and other applications, including for the management of accessions conserved in gene-banks.

Keywords: distinctness, maize, morphology, Plant Variety Protection, UPOV

Introduction

Comparative morphological studies of maize (*Zea mays* L.) play an important role in the management of crop diversity. Farmers working in traditional and subsistence agricultural communities use morphology to guide their use of germplasm (Perales et al, 2005; van Etten, 2006). Varietal management schemes practiced in subsistence communities are complex and sufficiently sophisticated that differences in heritabilities among morphological traits are recognized and used to direct management practices (Solieri and Cleveland, 2001). Farmer classifications of maize germplasm using ear characteristics can be equivalent to those obtained from sophisticated statistical analyses (Louette et al, 1997). Morphologies of US developed maize inbred lines and hybrids have changed since hybrids were introduced and crop management practices have changed (Duvick, 1997, 2005; Lauer, 2011). Morphological data provided the initial basis for taxonomic studies of maize (Sturtevant, 1884) and have remained a mainstay of maize racial taxonomy to current times (Ortiz et al, 2008). Morphological data also play an important role in the management of genetic resources that are conserved in ex situ gene-banks (Sanchez et al, 2000; Bioversity International, 2007).

Several guidelines for morphological descriptors of maize have been published (Bioversity International, 2007), including 1) by the International Board for Plant Genetic Resources (IBPGR) and the International Center for the Improvement of Maize and Wheat (IBPGR, 1991), 2) by Bioversity International/ CIMMYT (2009), the Council for Mutual Economic Aid (COMECON), the N.I. Vavilov All-Russian Scientific Research Institute of Plant Industry (VIR), and 3) by several national programs including the European Union Maize Landrace Database (EUMLDB, 2010) (http://www.ensam.inra.fr/gap/resgen88/results.htm) and by the USDA (2010) (http://www.ars-grin.gov/cgi-bin/npgs/html/desclist.pl?89) and 4) by the International Union for the Protection of Cultivated Varieties of Plants (UPOV) (UPOV, 1999, 2009)

Comparisons of morphological characteristics play a fundamental role in the determination of eligibility for the granting of intellectual property protection (IPP) under the auspices of UPOV in the form of Plant Variety Protection (PVP) or Plant Breeders' Rights (PBR). Issuance of a Plant Variety Protection (PVP) certificate depends on an ability to show that the candidate variety is morphologically distinct from all previously described varieties of that species. Detailed descriptions of characteristics are provided, for example in the "Guidelines for the conduct of tests for distinctness, uniformity and stability for maize" (UPOV, 1999, 2009). UPOV (2002a) designates subsets of characteristics to provide descriptions that can facilitate i) international harmonization of databases (asterisked characteristics) or ii) for grouping purposes so that similar varieties can be grouped together in the growing trial (UPOV, 1999, 2002a, 2009). UPOV notes the potential influences of environment

RECEIVED 02/04/2011

and examiner in a statement that "environmental effects, which, together with the effect of the examiner, may result in different descriptions for the same variety." (UPOV, C/37/10, 2003).

An important goal of UPOV is international standardization of field trial plot design, data recording, statistical analysis, reporting and comparison methodologies (van Wijk, 2003). van Hintum et al, (1995) conclude that the objective and interpretable scoring of traits in the UPOV system represents a "clear advantage" compared to the different schemes that have been published for the purposes of conducting taxonomy, comprehending phylogeny, or to assist in germplasm collection or gene-bank management.

However, in practice, the collection and analyses of morphological characteristics for the purpose of establishing the eligibility criterion of distinctness for the granting of PVP is not globally standardized. Prior to the development of the UPOV system, countries (including the US) had their own regulations. Technical criteria for the granting of rights differed among countries, the concept of what constituted a variety was not consistent, and technical standards and testing procedures were dependent on individual expertise (UPOV, 2000). Standardization was urgently required to better enable valid comparisons of varieties of common knowledge and to facilitate abilities to simultaneously obtain protection in several countries (UPOV, 2000).

Nonetheless, differences remain with regard to how descriptions are made, reported and compared. An examination of the protocols that are used in France and in the US is illustrative. In France, the Comite Technique Permanent de la Selection des Plantes Cultivees (CTPS, 2000) requests data for 16 morphological traits. States of characteristics for continuous quantitative morphological data are recorded or subsequently translated according (most usually) to a 1-9 scale composed of discrete classes or "notes". To facilitate scoring, "note" scores are illustrated by naming specific inbreds as exemplary control checks for each note score of each characteristic (UPOV, 1999). In some cases intermediate "notes" between defined notes are meaningful and allowable. In contrast, the translation of continuous data to discrete "notes" is not undertaken in the US system; instead, data are presented in terms of their mean and standard deviation. Also, the US PVP Office requests color characteristics to be recorded according to a Munsell color code (X-Rite Inc, Grand Rapids, Michigan) whereas UPOV records color using a discrete 1-9 scale. The US PVP Office requests breeders provide data on characteristics through their completion of a form titled "Exhibit C" (http://www.ams.usda.gov/science/pvpo/Forms/forms.htm). This form requests data for 49 morphological characteristics. And in further contrast to the situation in Europe, where the final round of data are generated by an official body (GEVES), all of the morphological data providing the objective description of the candidate variety are generated by individual applicants and presented to the US PVPO for examination.

The need to collect different sets of morphological data for many of the same inbreds in order to gain PVP protection in the US and in Europe provides a research opportunity to compare data generated under two quite different systems; one that translates scores to "notes" compared to another that utilizes the raw data from replicated field trials. We have access to data for morphological characteristics of maize inbred lines that were characterized according to the methodologies used in both the US and in Europe and that embrace a broad range of germplasm diversity. Comparisons of raw and transformed data and associations among inbred lines from each of the datasets provides a basis for determining 1) whether these data do, or could be sufficiently standardized to represent a single valid dataset, or 2) if they represent different datasets, then do they, at least, provide an equivalent basis for determining distinctness and for revealing associations among inbred lines. We wish to emphasize that we are not comparing or critiquing either system in regard to decisions that are made in respect of determination of varietal status. Rather, we seek to address a more fundamental question: Is it possible, with the data that are currently available (summary "note" scores from the EU and raw metric data from the US) to merge morphological data collected according to these different methodologies into a single meaningful database? While we would prefer to have access to raw metric data from the EU, rather than only for the summary type "note" scores, it is the note scores that indeed comprise the data describing each inbred line in the EU. If further research is warranted, then the collection of new data from replicated field trials carried out in both the EU and US would most likely be necessary.

Materials and Methods

Selection of characteristics and inbred lines

We obtained data describing the characteristics necessary for US PVP application for the inbreds included in this study during the period 1998-2005 from multi-location field trial plots. Between 2 and 3 locations were planted each year in the US located near Ankeny, Johnston, and Dallas Center, IA. Experiments were planted in late April or early May of each year using a randomized un-replicated experimental design nested by flowering date. Plots were planted at approximately 69,000 to 79,000 plants per ha. Most characteristics classified as discrete were collected at the plot level and assessed to give a single determination from the observation of 28 to 32 plants per inbred line. Quantitative traits were recorded from 5 plants per plot.

EU PVP data were provided to us by the EU PVP Office and by Pioneer staff in Europe. These data result from field trials that were conducted according to

Table 1 - List of characteristics used in the current analysis.

Trait	US units (for continuous traits) or scale (discrete traits)	EU scale	US and EU scale following transformation	US notes following transformation	EU notes following transformation	US Modal Note	EU Modal Note	US % range of expression	EU % range of expression	US % allele space used	EU % allele space used
Leaf Angle	degrees	1-9	1-9	1,2,3,4,5,6,7,8,9	1,2,3,4,5	4	3	100	56	100	56
Leaf Attitude†	1-5	1-9††	1-5	1,2,3	1,2,3	2	2	60	60	60	60
Brace Root Anthocyanin	1-5	1-9††	1-5	1,2,3,4,5	1,2,3,4,5	2	3	100	100	100	100
Growing Degree Units (GDU) to 50% Pollen Shed [G]*	GDUs	1-9	1-9	1,2,3,4,5,6,7,8,9	2,3,4,5,6,7,8,9	8	8	100	89	100	89
Tassel Glume Color	1-5	1-9††	1-5	1,3,4,5	1,2,3,4	1	1	100	80	80	80
Tassel Anther Color	1-5	1-9††	1-5	1,2,3,4,5	1,2,3,4	1	2	100	80	100	80
Tassel Axis Floret Density†	count	3-7	3-7	3,4,5,6,7	3,4,5,6	5	5	100	80	100	80
Tassel Branch Angle*	degrees	1-9	1-9	1,2,3,4,5,6,7,8,9	2,3,4,5,6,7	3	3	100	67	100	67
Tassel Attitude*†	1-5	1-9††	1-5	1,2,3,4	1,2,3,4	2	2	80	80	80	80
Tassel Primary Branch Number*	count	1-9	1-9	1,2,3,4,5,6,7,8,9	1,2,3,4,5	3	3	100	56	100	56
Growing Degree Units (GDU) to 50% Silk	GDUs	1-9	1-9	1,2,3,4,5,6,7,8,9	2,3,4,5,6,7,8,9	7	8	100	89	100	89
Silk Color [G]*	1-5	1-9††	1-5	1,2,3,4,5	1,2,3	1	3	100	60	100	60
Tassel Length	cm	1-9	1-9	1,2,3,4,5,6,7,8,9	3,4,5,6,7,8	5	5	100	67	100	67
Tassel Central Spike Length*†	cm	1-9	1-9	1,2,3,4,5,6,7,8,9	3,4,5,6,7,8,9	5	5	100	78	100	78
Bar Glume Color Absent-Present*	1-2	1-9††	1-2	1,2	1,2	1	1	100	100	100	100
Plant Height [G]*	cm	1-9	1-9	1,2,3,4,5,6,7,8,9	3,4,5,6,7,8,9	7	7	100	78	100	78
Ratio of Plant Height to Ear Height†	calculated	1-9	1-9	1,2,3,4,5,6,7,8,9	2,3,4,5,6,7,8,9	6	5	100	89	100	89
Leaf Width	cm	1-9	1-9	1,2,3,4,5,6,7,8,9	2,4,5,6,7,8,9	6	6	100	89	100	78
Shank Length	cm	1-9	1-9	1,2,3,4,5,6,7,9	1,2,3,4,5,6,7	3	3	100	78	89	78
Ear Length*	cm	1-9	1-9	1,2,3,4,5,6,7,8,9	3,4,5,6,7,8,9	6	5	100	78	100	78
Ear Diameter	mm	1-9	1-9	1,2,3,4,5,6,7,8,9	3,4,5,6,7,8	7	7	100	67	100	67
Ear Taper	1-3	1-3	1-3	1,2,3	1,2,3	2	2	100	100	100	100
Ear Row Number	count	1-9	1-9	1,2,3,4,5,6,7,8,9	3,4,5,6,7	5	4	100	56	100	56
Kernel Type (kernel texture score) [G]*	1-9	1-9	1-5	1,2,3,4,5	1,2,3,4,5	5	3	100	100	100	100
Cob Color [G]*	1-5	1-9††	1-5	1,5	1,5	5	5	100	100	40	40
Average								97.6	79.1	93.5	76.2

[G] Characteristics recommended by UPOV (1998) for grouping

* Characteristics that should be used on all varieties in every growing period and should always be included in the variety description (UPOV, 1998)

† Characteristics that are not required for US PVP applications but are collected due to occasional use in PVP applications for other countries

††The transformations of European data from 1-9 to 1-5 or 1-2 scales were employed temporarily to examine equivalency of trait modal values, distributions and compare trait data for each inbred. The non-transformed scales for EU data were utilized for all statistical analyses

UPOV protocols (UPOV, 1999, 2009). The vast majority of these data were obtained in France according to protocols promulgated by the (European) Community Plant Variety Office (CPVO) (http://www.cpvo.eu.int/main/) and as required by the European testing authorities for use in DUS examination as part of the European Plant Variety Protection process. Raw data are translated into "note" scores which are applied by comparison to specific check inbred lines that are also included in the field trials.

We identified 25 characteristics that describe the same botanical feature using comparable measuring or recording protocols documented by PVP officials in Europe and the US. Details of protocols can be found for Europe at http://www.upov.int/en/publications/tg-rom/tg002/tg_2_6.pdf and for the US at http://www.ams.usda.gov/AMSv1.0/ams.fetchTemplateData.do?template=TemplateC&navID=PlantVarietyProtectionOffice&rightNav1=PlantVarietyProtectionOffice&topNav=&leftNav=ScienceandLaboratories&page=PlantVarietyProtectionOffice&resultType=&acct=plntvarprtctn. We categorized these characteristics according to whether the data were classified as quantitative or discrete (qualitative) classes of expression. We included color in the discrete category. All 25 characteristics are listed in Table 1; the US continuous characteristics are listed alongside the units used to collect those data, while discrete characteristics are listed alongside the discrete scale used to collect data.

We then selected a set of 152 inbred lines on the basis that each inbred had been described for each of these characteristics both in Europe and also in the US. Each of the inbred lines is, or has been, used as a parent in at least one of 516 commercially sold US hybrids (first commercialized from 1991-2010) and 106 European hybrids (first commercialized 1992-2006). These inbred lines collectively represent a range of genetic backgrounds. According to their pedigrees the inbred lines can be assigned into:

a) predominantly Stiff-Stalk (SS) parentage (73 inbreds), b) predominantly non-Stiff-Stalk (NSS) parentage (73 inbreds, and c) predominantly Flint parentage (6 inbreds). The main pedigree backgrounds (mean percentage founder per inbred where contribution is > 5%) of the SS inbreds are: Argentinean Maiz Amargo (6%), Iodent (10%), Lancaster Low Breakage (7%), Leaming (14%), Osterland (5%), and Reid Yellow Dent (30%). The main pedigree backgrounds of the NSS inbreds are: Female Composite (FCOP) (7%), Iodent (28%), Lancaster Sure Crop (9%), Minnesota 13 (10%) and Reid Yellow Dent (9%). The main pedigree backgrounds of the flint inbreds are: Burr White (17%), European Flint (Eurflint) (20%), FCOP (7%), Flint open-pollinated (Flintop) (19%) and Hohenheim Synthetic (Hohensyn) (13%). Further details on these pedigree backgrounds are provided in Smith (2007) and in Smith et al (2004, 2006a, 2006b). The pedigree backgrounds of the SS and NSS inbreds were similar to those described for hybrids that were widely grown in the US that were developed by Pioneer Hi-Bred (Smith et al, 2006a, 2006b) and by other breeding organizations (Smith, 2007) during the 1990s and 2000s. The pedigree backgrounds of the predominantly flint inbred lines were similar to those of hybrids developed by Pioneer Hi-Bred that were widely used as parents of hybrids that were grown in northern France during the 1990s and 2000s (Smith et al, 2006a).

Continuous US characteristics and transformation to align with the EU format

In the US, data for continuous characteristics were obtained by direct measurement. These traits examined are listed in Table 1 adjacent to their units of measurement (discrete characteristics are listed with scales). Data were collected across multiple seasons and locations and reported to the US PVP Office as means and standard deviations. In order to facilitate comparisons of the data that were recorded in Europe and in the US, we transformed the US continuous data set to a format that would be more compatible with data obtained in Europe. The objective of this process of transformation was to translate the existing US scores to what they would be expected to have been if instead they had been recorded using the EU protocols in the US field trial locations. This approach was adopted because to make the reverse transformation from a specified range of EU "note" based final descriptions to plant based measurements, or plot-based scores, without an a priori range of minimum and maximum measurements would be highly speculative.

Transformation occurred as follows; for each inbred and for each continuous characteristic, the over-location and year mean value was calculated. The range for each characteristic over all inbreds was then divided by 8 to establish the "inter-note" difference. The mean value was then allocated into the appropriate discrete "note". For example, if the minimum value found for a trait was 123.5 and the maximum value was 276.5, the range of observed expression would be 153. This range was then divided by 8 to give an inter-note value of 19.125. Consequently, inbreds with trait values between the minimum (123.5) and minimum + 19.125 (142.625) would be assigned a transformed score of 1. To determine whether the transformation process had affected the distribution of the quantitative datasets, non-transformed and transformed US data were compared with EU data using Spearman's Rank Correlation analysis.

Discrete US characteristics

Characteristics collected in the US as discrete color characteristics were; Brace Root Anthocyanin, Tassel Glume Color, Tassel Anther Color, Silk Color, Bar Glume Color and Cob Color. The UPOV system generally prescribes a system of notes progressing from little or no anthocyanin expression to strong anthocyanin expression (1= no anthocyanin; 5 = very strong anthocyanin). In contrast, the US system requires that most color characteristics be recorded in the form of a Munsell code (http://www.xrite.com/top_munsell.aspx; Landa and Fairchild, 2005). Routinely we also convert each Munsell color to a 1-5 score reflecting the degree of anthocyanin present in order to tabulate data in the same way as other discrete traits. Here we utilized the US color data on a 1-5 scale to allow comparisons with data collected in Europe. One exception is Bar Glume Color which in the US is scored on a 1-2 scale, with 1 being absent and 2 being present. Other characteristics collected as discrete traits and examined here were Leaf Attitude, Tassel Attitude, Ear Taper and Kernel Type, with the associated scales being listed in Table 1.

The majority of the European traits were collected on a 1-9 scale. To align these data with the US scales, we transformed the European data to meet the US scale in order to assess equivalency of modal notes, % ranges of expression, % allele space used and individual note scores. Where European data scored on a 1-9 scale were transformed to a 1-5 scale, notes 1 and 2 = transformed note 1; 3 and 4 = 2; 5 and 6 = 3; 7 and 8 = 4; and 9 = 5. For the characteristic Bar Glume Color, the EU scale was 1-9 and the US scale was 1-2, but for both scales a note of 1 describes absence of color, with the remaining notes in the EU scale describing amount of anthrocyanin observed. Therefore, to transform the EU data, all notes not equal to 1 were transformed to a note of 2. For the dichotomous characteristic cob color, only 2 notes were scored in both the US and Europe; 1 and 5 in the US and 1 and 9 and in Europe. Therefore, for this trait, inbreds scored as a 9 in Europe were simply transformed to a 5. Ear Taper was scored on a 1-3 scale in both Europe and the US, but examination of individual protocols revealed that in Europe, a conical shape is scored as a note of 1 and a cylindrical shape is scored as a 3, whereas in the US the opposite is true. Therefore, for this trait, the 1 and 3 notes were reversed for the EU data. Kernel Type was scored on a 1-9 scale both in Europe and in the US, but the scales describe different phenotypes. In the US the PVP form designates a score of 1 = Sweet, 2 = Dent, 3 = Flint, 4 = Flour, 5 = Pop, 6 = Ornamental, 7 = Pipecorn and 8 = Other. Since Pioneer's germplasm is either Flint or Dent, instead we collect a Kernel Texture Score on a 1-9 scale with 1 describing a dent-type kernel and a note of 9 describing a flint type, and intermediate notes describing gradations between the two, and then transform this data to the US PVP scale for Kernel Type. In this study we utilized the 1-9 Kernel Texture Score. In Europe, a note of 1 = Flint, 2-4 = Flint/Dent to Dent/Flint, 5 = Dent, 6 = Sweet, 7 = Pop, 8 = Waxy and 9 = Flour. To align the US and European scales, both sets of data were transformed to a 1-5 scale (1 being flint and 5 being dent) with the notes for EU data remaining the same (there were no scored notes of 6-9). The US data were transformed so that note 1 = transformed note 5, notes 2 and 3 = 4, notes 4 and 5 = 3, notes 6 and 7 = 2, and notes 8 and 9 = 1.

Analysis of untransformed European and US data

The non-parametric Kolmogorov-Smirnov test was used to ascertain whether the underlying probability distributions for each trait differed between the non-transformed EU and US data. Second, non-transformed EU and US data were compared in a pair-wise fashion for each trait using Spearman Rank Correlation coefficient analyses.

Comparison of notes for transformed EU and US data

We refer to the US and EU datasets that includes the transformed data as trUS and trEU to differentiate these data from the raw PVP data. Several parameters were used to compare the notes for trEU and trUS data (Table 1). These were: 1) relative position of the modal (most frequently scored) note, which indicates whether the distribution of note scores resembles a "normal distribution" (mode centrally placed between minimum and maximum) or another type of distribution (e.g., clumped); 2) percent range of expression, which describes the range from minimum to maximum notes observed as a percentage of the full transformed scale; 3) percentage of allele space used which is the percent of observed notes compared to those that are possible to occur on the transformed scale and includes information on the contiguous nature of the notes used; (4) the percent agreement of individual note scores between the US and European data for each trait. This later parameter was assessed according to two levels of stringency; (a) exactly the same note scores required as qualification of agreement, and (b) one note difference allowable to qualify as agreement.

Overall congruence of EU and transformed US data

Further statistical approaches were used to investigate the overall level of congruence of the EU and US data. For these analyses, data for continuous US characteristics data were maintained in their transformed state so that both the EU and US datasets could be analyzed with scalar increments. All EU data were non-transformed to allow for maximum information on the scale utilized. The analyses utilized were; 1) Comparing the distributions of data among notes using the Shannon Entropy statistic, 2) correlating inbred characteristics determined in the EU versus the trUS using Spearman Rank Correlation Analyses, and 3) examining associations between all inbreds on the basis of their overall characteristics profile using multivariate statistical analyses.

The Shannon Entropy Index (Hill, 1973; Straathof, 2007) allows characteristics recorded in the EU and US datasets to be compared for "richness and uniformity of note distributions". We used the multi-state version of Shannon's entropy with a suitable logarithm base relevant to the expected number of expected “notes”. If the data are exactly uniformly distributed across the expected notes, then Shannon’s entropy would have the value of 1. Alternately, if all observations for a particular characteristic were reported as the same “note,” irrespective of the number of “notes”, then the computed value of Shannon’s entropy would be zero. Simulations showed that “centrally distributed data”, as are commonly found in biological experimentation, when scored over a nine “note” note-scale have a Shannon entropy value of about 0.7 (Straathof, 2007). Spearmans Rank Correlation Coefficients were calculated according to Kendall and Stuart (1973).

For multivariate analyses, a balanced dataset was required. We chose not to impute missing data, even by using commonly applied algorithms, lest the imputation process itself might become responsible for observed differences between the EU and trUS datasets. Instead we removed traits and inbreds with missing data resulting in the removal of data for 3 characteristics (Leaf Attitude, Tassel Attitude, Silk Color) and 23 inbreds from both EU and US matrices. This exercise resulted in two balanced data-set matrices of 22 characteristics (columns) and 129 inbreds (rows) representing data for each of the EU and trUS protocols. A further set of matrices (22 by 129 dimensions) was created to represent random data to serve as a benchmark of unstructured data. GenStat software (Payne et al, 2006) was utilized to compute similarity matrices based on both the Euclidean and City Block metrics, with subsequent analyses using Mantel Tests. For Mantel Tests we performed 100, 1,000, and 10,000 permutations of the matrices. Comparisons were made for EU to trUS data as well as EU to random data, trUS to random data, and random to random data. Principal coordinate (PCO) analysis was undertaken using GenStat and vector coordinates of the PCO’s which resulted in a graphical representation of the spatial relationships for inbreds for each source of data (EU or trUS).

Evaluation of the effect of maturity zone

We allocated inbred lines into one of four maturity zones (MZ) according to the number of heat units that are required for the inbred to reach flowering and maturity; MZ1 = inbreds with maturity 70-90 Comparative Relative Maturity (CRM) (Eckert et al, 1987; Olson and Sander,1988; Lauer, 1998), which corresponds to the maturity region of northern North America; MZ2 = 91-100 CRM, which corresponds to the maturity region of northern corn belt; MZ3 = 101-115 CRM, which corresponds to the maturity region occupying the central Corn Belt; and MZ4 = 116-126 CRM, which corresponds to the maturity region of southern United States, northern Mexico and more tropical longer season environments. We then investigated, for both EU and trUS data whether there was an effect of maturity zone in generating the morphological data for the plant characteristics. We compared the a priori MZ, allocated according to the known CRM data to 1) the most likely MZ that is predicted following Discriminant analysis and Canonical Variate Analysis (White and Law, 1991) and with 2) groupings obtained using inbred data with randomly applied MZ data. A high instance of agreement between actual MZ and those predicted by Discriminant Analysis indicates that there is a strong case that MZ is influencing the classification. Conversely, a low incidence of agreement would show that there is only weak support that the MZ classification influences the morphological data.

Results

Analysis of untransformed European and US data

Results from applying the Kolmogorov-Smirnov test (supplementary Table 1) showed that eight (out of 25; 32%) characteristics (Tassel Branch Angle, Tassel Central Spike Length, Bar Glume Color, Plant Height, Ratio of Plant Height to Ear Height, Ear Length, Ear Taper, and Cob Color) had no distributional differences between the EU and US datasets ($p > 0.05$). Spearman Rank Correlation Coefficients for the raw EU and US data were modest with an average of 0.63 nonetheless, all were positive and were highly significant ($p<0.001$; data not shown). The characteristics with the lowest correlations coefficients (and the only values below 0.4) were Ear Taper (0.14), Kernel Type (0.32) and Tassel Floret Density (0.29). Characteristics with the highest correlations were GDUs to 50% shed (0.92), GDUs to 50% Silk (0.90) and Cob Color (0.98).

Effect of the transformation process

Comparing US and trUS data for continuous characteristics, the average Spearman’s Ranked Correlation Coefficient was 0.98 (minimum 0.96; maximum 0.99) and each trait was significantly correlated at a probability <0.001, hence confirming that the “raw” and trUS data were not substantially affected by the transformation process. Rank correlation coefficients for inbreds on the basis of raw US compared to the EU data changed numerically compared to the correlation of the transformed US data with the EU data (mean change in Spearman’s Ranked Correlation Coefficient was 1.7%, maximum change of 10.4%; full data not shown) but there were no cases where the level of statistical significance changed as a result of the transformation procedure.

Comparison of note scores for transformed EU and US data

The initial inspection of the transformed data provided an overview of the range and distribution of note scores for each individual characteristic (Table 1). Transformed EU (trEU) and US (trUS) data concurred for the observed modal note for 16 (64%) characteristics, with a difference of only 1 note for a further 7 (28%) characteristics. Only Kernel Type and Silk Color had a larger modal note difference. Modal

Table 2 - Percent agreement of notes scores between transformed EU and US data (sorted by rank % agreement notes).

	% Agreement Notes	Rank % Agreement	% At Most One Note Difference	Rank % At Most One Note Difference
Cob Color ^{G*}	99.3	1	99.3	4
Ear Taper	89.3	2	100.0	1
Bar Glume Color*	83.6	3	100.0	2
Leaf Attitude*	59.4	4	100.0	3
Tassel Axis Floret Density	50.3	5	93.4	5
Silk Color	47.9	6	78.2	19
Ear Row Number	47.7	7	91.4	6
Tassel Glume Color	46.7	8	78.3	18
Tassel Attitude	45.7	9	85.4	10
Brace Root Anthocyanin	41.1	10	84.8	11
Plant Height	40.1	11	85.5	8
Ear Diameter	39.7	12	79.5	16
Growing Degree Units (GDU) to 50% Pollen Shed ^{G*}	39.5	13	85.5	9
Tassel Branch Angle*	39.1	14	79.5	17
Ratio of Plant Height to Ear Height ^{G*}	38.2	15	89.5	7
Tassel Anther Color	36.8	16	80.3	15
Ear Length*	32.5	17	84.8	12
Leaf Angle	29.7	18	76.6	20
Tassel Central Spike Length*	29.1	19	82.8	13
Tassel Length	28.5	20	68.2	22
Tassel Primary Branch Number*	25.3	21	50.7	25
Growing Degree Units (GDU) to 50% Silk ^{G*}	25.2	22	81.8	14
Shank Length	23.2	23	74.2	21
Kernel Type ^{G*}	16.9	24	61.5	23
Leaf Width	16.0	25	58.0	24
Average	42.8		81.9	

^{G} Characteristics recommended by UPOV (1998) for grouping

* Characteristics that should be used on all varieties in every growing period and should always be included in the variety description (UPOV, 1998)

notes that showed data were clumped or skewed on the transformed scales (modal notes of 1,2,8 or 9 for data assessed on a 1-9 scale, and 1 or 5 on a 1-5 scale) were; GDUs to 50% Pollen Shed (modal note of 8 on a 1-9 scale for both trUS and trEU data), GDUs to 50% silk (8 on a scale of 1-9 for trEU, modal note of 7 for trUS), Tassel Glume Color (1 = no color, on a 1-5 scale for trUS and trEU), Tassel Anther Color (1 = light pink, on a scale 1-5 for trUS, modal note of 2 for trEU), Kernel Type (5 = Dent, on a scale of 1-5 for trUS, modal note of 3 for trEU), Cob Color (5 = yellow cob, on a scale of 1-5 for trUS and trEU).

The trUS data exhibited higher metrics for “percent range of expression” and “percent allele space used” (averages of 97.6 and 94.0%, respectively) compared to the trEU data (79.1 and 76.4%, respectively). Higher metrics could be expected for the transformed continuous US characteristics as the data were aligned along the full 1-9 scale. However, these metrics were also higher (although not as pronounced) for the US data when just the non-transformed discrete characteristics were considered; for trUS data, averages were 94.0 and 86.0%, respectively, compared to 86 and 80% for Europe. These results demonstrate a general compression in the scale used for assessment of characteristics in Europe, at least for the germplasm examined here. For each characteristic, percent range of expression and percent allele space used generally agreed owing to the notes mostly all being utilized within the range of expression. For trUS, only three characteristics did not use all of the notes within the range of expression;

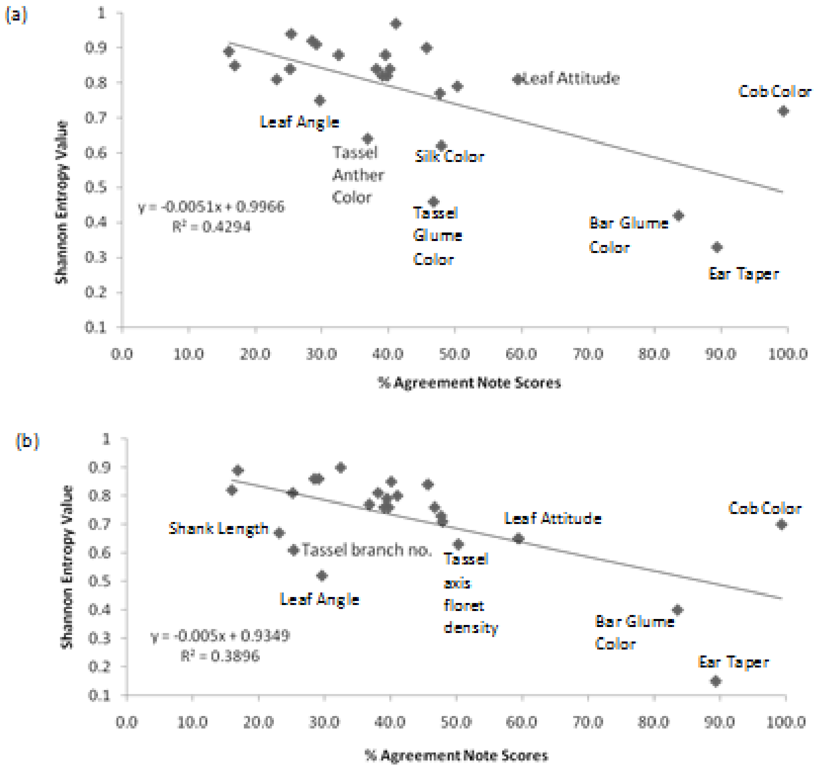

Figure 1 - Correlations between % agreement of notes scores between transformed US and EU data and Shannon Entropy values for characteristics assessed in the US (a) and Europe (b).

the discrete characteristics Tassel Glume Color and Cob Color, plus the continuous (transformed) characteristic Shank Length. For trEU data, only two characteristics, Leaf Width and Cob Color did not use all of the notes within the range of expression.

Table 2 presents the percentage agreement between the trEU and trUS datasets for note scores for each of the characteristics across the 152 inbreds. The range of agreement ranged from 99.3% of inbreds for the dichotomous characteristic of Cob Color, to 16% of inbreds for Leaf Width. Besides Cob Color, only 4 other characteristics exhibited >50% agreement; Ear Taper (89.3%), Bar Glume (83.6%), Leaf Attitude (59.4%) and Tassel Axis Floret Density (50.3%). Besides, Leaf Width, traits showing 25% or less agreement were Tassel Primary Branch Number (25.3%) , GDUs to 50% Silk (25.2%), Shank Length (23.2%), and Kernel Type (16.9%). Average agreement for note scores across all characteristics was 42.8%. When the criteria were relaxed to allow agreement of scores differing by 1 note then the mean level of agreement increased considerably to 82.0% (range 50.7-100%). The relative ordering of characteristics according to their % agreement of note scores did not markedly alter for most characteristics; the highest and lowest ranked characteristics were similar as the definition of agreement was relaxed. Traits that had considerably improved rankings when the criteria were relaxed were; Ratio of Plant Height to Ear Height, Ear Length, Tassel Central Spike Length and GDUs to 50% silk (Table 2).

Overall congruence of EU and transformed US data

Shannon's entropy H1 values are given in Table 3. Values are ranked to allow comparison between H1 values for the EU data and those based on the trUS PVP data. The Entropy values for the US characteristics were slightly shifted towards 1 (average 0.78, range 0.33 to 0.96) compared to the European data (average 0.72, range 0.15 to 0.90). The highest ranked EU characteristics were Ear Length (0.90), Kernel Type (0.89), Tassel Central Spike Length (0.86) and Tassel Length (0.86). The highest ranked trUS characteristics were Brace Root Anthocyanin (0.97), Tassel Primary Branch Number (0.94), Tassel Length (0.92) and Tassel Central Spike Length (0.91). Out of the top 10 ranked characteristics for Europe, 8 were also ranked in the top 10 for the trUS data, and correlations between EU and US Entropy values had a highly significant with a correlation coefficient of 0.72. The characteristics with the most different Shannon Entropy scores (difference >0.2) among the EU and US data were, (difference in parentheses), Leaf Angle (0.23), Tassel Glume Score (0.3), and Number of Primary Tassel Branches (0.33). Two characteristics (Bar Glume and Ear Taper) ranked the lowest in their Shannon Entropy values according to both the trUS and EU datasets.

Shannon Entropy values for the EU and trUS were also correlated against the % Agreement of Note Scores found. Correlations showed a negative trend and the slope was significant in both cases (correlation coefficients of 0.39 and 0.42, P <0.001 for EU and US data respectively; Figure 1).

Multivariate analyses

Comparisons of EU and trUS data can also be measured by comparing associations among inbred lines following multivariate statistical analyses. Mantel statistics for the degree of association between the matrices of inbred lines created from the multivariate analysis of EU compared to trUS characteristics data were 0.42 and 0.37 for the Pearson's and Spearman's correlation variants, respectively (Supplementary Table 2). With 10,000 permutations of the trUS matrix (keeping the EU fixed) and recalculating the Mantel statistics, there were no occasions where the permuted Mantel statistic exceeded values computed from the un-permuted data. Hence, we can determine that the association between the EU and trUS data matrices is very unlikely to have occurred by chance. Similarly, both Mantel and Pearson statistics showed a lack of association between the EU and a randomly generated dataset.

Maturity zone effects

The rates for allocation of inbred lines into the correct maturity zone (MZ) by using discriminant analysis were very similar for the EU and trUS data (80.6% and 79.1%, respectively). In contrast, the set of inbreds with randomly applied MZ data, (as a bench-mark assessment), showed a lower correct allocation rate of 48.1%. An analysis was also made of the extent to which the same inbreds were allocated correctly according to MZ in both the EU and trUS datasets by using discriminant analysis. Sixteen inbreds were misallocated in both EU and US datasets; 9 inbreds were misallocated with the EU data, and a further 11 inbreds were misallocated with the transformed US data. Results from Canonical Variate Analysis also showed a systematic effect of MZ, but all misallocations were, at most, by one MZ. In contrast, the analysis of random MZ data showed no discernible pattern of grouping according to MZ.

Discussion

We emphasize that the primary focus of this study is to determine whether morphological data collected for the same characteristics but according to different reporting protocols have the potential to be meaningfully merged, including via transformation, into a single database. We are not seeking to compare or to critique the specific ways by which these data are then utilized, either in the US or in the EU, as a basis to determine Distinctness for the purpose of determining eligibility for PVP. The process of determining Distinctness is a subsequent exercise to that of collecting the morphological data. For example, Distinctness testing in France is undertaken by the Groupe d'etude et de controle des varieties et des semences (GEVES) and relies, not only on the characteristics listed by the CTPS, but also upon the opinion of an expert committee. GEVES weights individual characteristics according to their presumed basis of complexity for their genetic control of expression thereby producing a phenotypic distance known as the logical de comparison de lignes de mais (LCLM), which also includes isozymic data. Committee experts also weight characteristics. These weightings are "probably different from the LCLM weight" and they "may also take into account other traits than those of the UPOV maize guidelines" (Dillmann and Guerin, 1998). We also wish to emphasize that this study can only represent an initial examination of the possibility that available data collected according to these different criteria could be merged into a single meaningful database. If further studies appear warranted then these would need to include new replicated field trials to generate raw data that form the basis of the "note score" methodology that is used in the EU.

In this study we aligned and compared maize inbred lines characterized using morphological data collected in the US and Europe and that are subsequently recorded according to contrasting protocols. Initial comparisons prior to any transformation of the data showed that 8 of 25 (32%) of the characteristics had no distributional differences between the US and EU datasets (Supplementary Table 1). Spearman Rank Correlation Coefficients for the raw EU and US data were modest with an average of 0.63 nonetheless, all were positive and were highly significant (p<0.001; data not shown).

Nevertheless, transformation of the data was necessary so as to better align data scores obtained from different methodologies employed in the EU and US. An immediate question then to be addressed was whether transformation of the data, particularly of the continuous US characteristics, injected bias. Consequently, we examined the effect of transformation by comparing the raw US with transformed US data, as well as the raw and transformed US data with EU data. Rank correlation analyses showed very little effect of the transformation process itself. Ranges of expression were shifted upwards for the trUS data compared with the equivalent characteristics collected in Europe due to the transformation process making use of the full scale of 1-9 notes. However, a shift upwards was also observed for the non-quantitative, discrete US characteristics, suggesting a general trend for a small portion of the scale being used in Europe compared with the US. This is possibly a consequence of the germplasm being studied here not representing the full range of germplasm that is studied in Europe. For example, flint germplasm is very widely used, most especially in northern Europe, whereas its usage in the US is minimal.

Comparison of note scores for transformed EU and US data

Modal values for note scores were very similar for the transformed US (trUS) and EU (trEU) characteristics. Modal notes agreed for 16 (64%) characteristics, with a difference of only 1 note for a further 7 (28%) characteristics. Skewing of the modal note towards one end of the scale could be attributed to the characteristics of the germplasm being examined. Modal values for GDUs to 50% Pollen and Silk Shed were 7 or 8 on a 1-9 scale; illustrative of the mid-late maturity material which predominated in this inbred set. Cob Color had a modal note of 5 as mostly yellow cob inbreds were studied here. Kernel Type had a modal note of 5 for trUS data, indicative of this set of inbreds being largely dent germplasm, although a modal note of 3 (flint-dent type) for US scored data may suggest some flint-dent admixture.

Despite the similarities in modal notes for the trUS and trEU data, the level of agreement for individual note scores was generally low, ranging from 16.0% to 99.3% with a mean of 42.8%. The strongest agreement was for the dichotomous character Cob Color at 99.3%. Such a result for this trait could be anticipated because it is recorded on the UPOV scale as either “note” 1 or “note” 9; an approach reserved for single gene characteristics for which the state of expression is usually more clear and robust over environments. Ear Taper and Bar Glume also had high values of > 80% agreement although these results were largely due to both of these traits having limited variation as evidenced by their low Shannon Entropy values. Out of the top 10 ranking characteristics for agreement of note scores, only 1, Cob Color, is a UPOV recommended characteristic for grouping and 3 (Cob Color, Bar Glume Color and Tassel Attitude) are considered characteristics that should always be included in the variety description. A larger number of UPOV grouping and required characteristics (3 grouping and 5 required) were found in the bottom 10 ranking traits. Most characteristics had levels of note score agreement below 50%. For Kernel Type, some of the disagreement may be attributed to the generally lower level of flint germplasm that is used

Table 3 - Shannon Entropy values and ranking (rank 1 nearest to uniform distribution [H1=1]) for EU and US data (sorted by Rank European Data Shannon’s Entropy).

Trait	European Data Shannon Entropy	Rank European Data Shannon’s Entropy	US Data Shannon Entropy	Rank US Data Shannon’s Entropy
Ear Length*	0.90	1	0.88	8
Kernel Type ^{G}*	0.89	2	0.85	9
Tassel Central Spike Length*	0.86	3	0.91	4
Tassel Length	0.86	4	0.92	3
Plant Height	0.85	5	0.84	11
Tassel Attitude*	0.84	6	0.90	5
Leaf Width	0.82	7	0.89	6
Ratio of Plant Height to Ear Height ^{G}*	0.81	8	0.84	12
Growing Degree Units (GDU) to 50% Silk ^{G}*	0.81	9	0.84	10
Brace Root Anthocyanin	0.80	10	0.97	1
Growing Degree Units (GDU) to 50% Pollen Shed ^{G}*	0.79	11	0.88	7
Tassel Anther Color	0.77	12	0.64	21
Tassel Glume Color	0.76	13	0.46	23
Ear Diameter	0.76	14	0.82	13
Tassel Branch Angle*	0.76	15	0.82	14
Ear Row Number	0.73	16	0.77	18
Silk Color	0.71	17	0.62	22
Cob Color ^{G}*	0.70	18	0.72	20
Shank Length	0.67	19	0.81	15
Leaf Attitude	0.65	20	0.81	16
Tassel Axis Floret Density	0.63	21	0.79	17
Tassel Primary Branch Number*	0.61	22	0.94	2
Leaf Angle	0.52	23	0.75	19
Bar Glume Color*	0.40	24	0.42	24
Ear Taper	0.15	25	0.33	25

^{G} Characteristics recommended by UPOV (1998) for grouping

* Characteristics that should be used on all varieties in every growing period and should always be included in the variety description (UPOV, 1998)

in the US compared to the EU. In addition, there has been a lack of consistency during the period in which these inbreds were developed with regard to how kernel type is scored in the US. Earlier data scores for this characteristic were made from visual inspection of the grain but some have later been changed to reflect the relative contributions by pedigree of flint or dent germplasm. Other examples of germplasm mixing in the pedigrees of varieties that have then lead to subsequent changes in scoring and classification procedures for grain have been reported for wheat (Williams, 2006; Canadian Grain Commission 2008a, 2008b).

When matching note criteria were relaxed to allow agreement where characteristic scores differed by, at most 1 "note", then the absolute numerical expression of agreement increased to a median of 81.8% (range 58-100%). The relative ranking of characteristics according to their percentage agreement (Table 2) did not markedly alter, although some characteristics appeared to rank highly due to being assessed on a small scale, eg Ear Taper (scale of 1-3), Bar Glume Color (scale 1-2) and Tassel Attitude (only notes 1-4 scored on a 1-5 scale).

The level of distribution of note scores across the scale were assessed using Shannon Entropy values. Entropy values for the US data were slightly shifted towards 1 (range of entropy 0.33 to 0.96) compared to the EU data (range 0.15 to 0.90). Again, the transformation process could be expected to have contributed some bias in this direction for continuous US data due to the linear transformation of the original US data forcing traits onto a 1-9 "note scale". Out of the top 10 ranked characteristics for Europe, 8 were also ranked in the top 10 for the trUS data, and correlations between EU and US Entropy values were highly significant. The weakest entropies for both the US and Europe were for the characteristics Bar Glume Color and Ear Taper. However, such general agreement was not universally the case. Characteristics with the most different Shannon Entropy scores (difference >0.25) among the EU and US data were Tassel Glume Color and Tassel Primary Branch Number. Indeed, Tassel Primary Branch Number had the 2nd highest Shannon's entropy value among the US data but ranked almost lowest (22^{nd}) among the EU data.

When considering whether a characteristic is reliable and useful (discriminative) for determining distinctness, several criteria should be considered. For example, a characteristic that scores consistently may merely be uninformative in the germplasm being examined (e.g. Ear Taper and Bar Glume Color). These characteristics, together with leaf angle also had low % note agreement and low Shannon Entropy values in both Europe and the US; they therefore showed a repeated lack of informativeness as well as inconsistency. In contrast, characteristics that performed well for both of these criteria and which would therefore appear to be both consistent and discriminative were Cob Color and Leaf Attitude. Cob Color is a UPOV recommended grouping characteristic, but Leaf Attitude is not and neither is it a required characteristic under the US PVP system. Many characteristics were incongruent among US and EU data, a finding which thus exemplifies the challenges of using morphological characteristics either to compare genotypes or to establish meaningful taxonomic groups.

The degree of equivalency between the EU and US datasets in respect of their ability to be used to show similar associations among inbred lines

Results from multivariate analysis showing associations of inbred lines using either the EU or trUS data support the hypothesis that there was a strong association between the EU and trUS data matrices and that this association was very unlikely to have occurred by chance. However, because only 42.8% of characteristics had the same note scores for trEU and trUS data for any inbred line, we conclude, at least initially and with data that are currently available, that the EU and US data are sufficiently different that they cannot be appropriately merged, even after transformation of the US data to best simulate data collection and reporting according to the EU format. Whether transformation of raw data that are obtained prior to assignation of "note" scores would provide a suitable basis for transformation and merging of the data sets will require additional field trials. The process of assigning note scores in reference to check inbred lines that have pre-designated note scores for each characteristic is designed to minimize the confounding effects of G x E. Nevertheless, to determine the extent to which the discrepancies we have found between EU and trUS data can be more precisely partitioned due to genotype by environmental effects in contrast to the different recording and data reporting methods will require additional experimentation including replicated field trials.

The characteristics selected by UPOV as a basis to determine DUS status are individually under simpler genetic control than yield, but nonetheless, the genetic control of many of characteristics used for DUS can be relatively complex so that expression is quantitative in nature (Sourdille et al, 1991; Austin et al, 2001; Bredemeijer et al, 2002; Mickelson et al, 2002; Enoki et al, 2006; Li et al, 2007). Consequently, the expression of morphological characteristics is subject to genotype x environmental interaction. For example, even when identical protocols have been used to collect morphological description data for the same genotypes in different locations, those data can be quite dissimilar (Jones et al, 2003; Hof and Reid, 2008). It should not be surprising; therefore, that abilities to make valid comparisons of morphological data that have been obtained using, not only contrasting systems for scoring and recording, but also in different geographies, will therefore be doubly challenging and practically impossible.

Nevertheless, it is also apparent that, despite the differences in environments, and methods of recording and reporting data, that analyses of data sourced from each of the EU and US system results in a very similar array of inbred lines when the relative ranking, or association of inbred lines are compared, either on an individual trait by trait basis, or on a multiple trait basis. Thus, while the majority of EU and US note scores cannot meaningfully be compiled into a single database, they do individually each provide a reliable basis upon which to determine distinctness in respect of the current requirements of UPOV. We also showed a systematic effect of maturity zone, implying that comparisons within maturity zones have greater validity than grouping all inbreds into a single maturity zone.

Standardization of any system to record and report data for morphological characteristics optimally requires use of exactly the same protocols to be applied and adhered to for the entire process. However, experience has shown that, even in those circumstances, obtaining data for morphological characteristics that can be meaningfully combined into a single database is still challenging and elusive (Jones et al, 2003; Hof and Reid, 2008). We recommend careful reexamination of the utility of collecting data for numerous morphological characteristics, including for the purposes of characterizing germplasm for conservation and subsequent management in genebanks. Our conclusions concur with other recommendations that a limited number of morphological characteristics: Relative maturity, plant height, and a photograph of the ear together with collection location, should form the basis for gene-bank management (Global Crop Diversity Trust, 2007), plus additional characteristics that have been shown to be demonstrably reliable and informative (such as Cob Color and Leaf Attitude), be used in contrast to the use of a more numerous list of morphological descriptors (CIMMYT/IBPGR, 1991).

References

Austin DF, Lee M, Veldboom LR, 2001. Genetic mapping in maize with hybrid progeny across testers and generations: plant height and flowering. Theor Appl Genet 102:163-176

Bioversity International, 2007. Guidelines for the development of crop descriptor lists. Bioversity Tech Bull Series. Bioversity international, Rome, Italy. XII

Bioversity International/CIMMYT, 2009. Key access and utilization descriptors for maize genetic resources

Bioversity International (Rome) and the International Wheat and Maize Improvement Center, Mexico City

Bredemeijer GMM, Cooke RJ, Ganal MW, Peeters R, Isaac P, Noordijk Y, Rendell S, Jackson J, Roder MS, Wendehake K, Dijks M, Amelaine M, Wickaert V, Bertrand L, Vosman B, 2002. Construction and testing of a microsatellite database containing more than 500 tomato varieties. Theor Appl Genet 105:1019-1026

Canadian Grain Commission, 2008a. About KVD – Changes to western Canadian wheat classes. http://www.grainscanada.gc.ca/wheat-ble/classes/classes-2008-eng.htm

Canadian Grain Commission, 2008b. Kernel visibility distinguishability in western Canadian wheat classes. http://www.grainscanada.gc.ca/wheat-ble/classes/kvd-dvg-eng.htm

CTPS, 2000. Questionnaire Technical Additional Zea mays L. Mais. Ministere de l'acriculture, Comite Technique Permanent de la Selection des Plantes Cultives, GEVES, Guyancourt, France

Dillmann C, Guerin D, 1998. Comparison between maize inbred lines: Genetic distances in the expert's été. Agronomie 18: 659-667

Duvick DN, 1997. What is yield? pp. 332–335. In: Developing Drought and Low N-tolerant Maize. Edmeades GO, Bänziger B, Mickelson HR, Pena-Valdivia CB (eds). CIMMYT, El Batan, Mexico

Duvick DN, 2005. The contribution of breeding to yield advances in maize (*Zea mays* L.). Adv Agron 86:83-145

Eckert DJ, Hunter RB, Keener HM, 1987. Hybrid maturity-energy relationships in corn drying. In National Corn Handbook NCH-51. Nielsen R (ed)

Enoki H, Miki K, Koinuma K, 2006. Mapping of quantitative trait loci associated with early flowering of a northern flint maize (*Zea mays* L.) inbred line. Maydica 51:515-523

Global Crop Diversity Trust, 2007. Global Strategy for the *ex situ* conservation and utilization of maize germplasm. Global Crop Diversity Trust, Rome, Italy

Hill MO, 1973. Diversity and Evenness: A unifying notation and its consequences. Ecology 54:427-432

Hof IL, Reid A, 2008. Construction of an integrated microsatellite and key morphological characteristic database of potato varieties on the EU common catalogue Part 1: Discussion of morphological and molecular data (revised). 11th session of the working group on biochemical and molecular techniques and DNA profiling in particular, Madrid, September 16-18, 2008. BMT/11/0 Rev, UPOV, Geneva, Switzerland

IBPGR,1991. Descriptors for Maize. International Wheat and Maize Improvement Center, Mexico City/International Board for Plant Genetic Resources, Rome

Jones H, Jarman RJ, Austin L, White J, Cooke RJ. 2003. The management of variety reference collections in distinctness, uniformity and stability testing of wheat. Euphytica 132:175-184

Kendall MG, Stuart A, 1973. The Advanced Theory of Statistics, Volume 2: Inference and Relationship, Griffin. ISBN 0852642156 (Sections 31.19, 31.21)

Landa ER, Fairchild MD, 2005. Charting Color from the Eye of the Beholder. American Scientist 93 (5): 436–443.

Lauer J, 1998. The Wisconsin comparative relative maturity (CRM) actions, selection response and heterosis. p. 81–92. In: JG Coors, system for corn. Field Crops 28: 31–21.

Lauer S, Hall B, Mulaosmanovic E, Anderson SR, Nelson B, Smith S, 2011. Morphological Changes in Parental Inbred Lines of Maize Hybrids in the Central Corn Belt of the United States, Crop Sci submitted.

Li Y, Dong Y, Niu S, Cui D, 2007. The genetic relationship among plant-height traits found using multiple-trait QTL mapping of a dent corn and popcorn cross. Genome 50: 357-364

Louette D, Charrier A, Berthaud, 1997. *In situ* conservation of maize in Mexico: Genetic diversity and maize seed management in a traditional community. Econ Bot 51: 20-38

Mickelson SM, Stuber CW, Senior L, Kaeppler SM, 2002. Quantitative trait loci controlling leaf and tassel traits in a B73 x Mo17 population of maize. Crop Sci 42: 1902-1909

Olson RA, Sander DH, 1988. Corn Production. In Corn and Corn improvement 3rd ed. Sprague GF, Dudley JW (eds) Agronomy 18: 639-686

Ortiz R, Crossa J, Franco J, Sevilla R, Burgueno J, 2008. Classification of Peruvian highland maize races using plant traits. Genet Resour Crop Evol 55: 151-162

Payne, RW, Murray DA, Harding SA, Baird DB, Soutar DM, 2006. GenStat for Windows (9th Edition) Introduction. VSN International, Hemel Hempstead

Perales HR, Benz BF, Brush SB, 2005. Maize diversity and ethnolinguistic diversity in Chiapas, Mexico. PNAS 102: 949-954

Sanchez JJ, Goodman MM, Stuber CW, 2000. Isozymatic and morphological diversity in the races of maize in Mexico. Econ Bot 54: 43-59

StatSoft inc, 2005. STATISTICA (data analysis software system), version 7.1. www.statsoft.com

Smith S, 2007. Pedigree Background Changes in U.S. Hybrid Maize between 1980 and 2004. Crop Sci 47: 1914-1926

Smith JSC, Desbons P, Gogerty J, Niebur WS, 2006a. Changes in parentage and genetic diversity of widely used maize hybrids grown in the northern United States and France from 1930 to the present. Maydica 51: 57-77

Smith JSC, Duvick DN, Smith OS, Cooper M, Feng L, 2004. Changes in pedigree backgrounds of Pioneer Brand maize hybrids widely grown from 1930 to 1999. Crop Sci 44: 1935-1946

Smith S, Loffler C. Cooper M, 2006b. Genetic diversity among maize hybrids widely grown in contrasting regional environments in the United States during the 1990s. Maydica 51: 233-242

Straathof SM, 2007. Shannon's entropy as an index of product variety. Economics Letters 94: 297-303

Sturtevant EL, 1884. Maize: An attempt at classification. Democrat and Chronicle Print Rochester, NY

Sourdille P, Baud S, Leroy P, 1996. Detection of linkage between RFLP markers and genes affecting anthocyanin pigmentation in maize (*Zea mays* L.). Euphytica 91: 21-30

UPOV 1991. Act of 1991. International Convention for the Protection of New Varieties of Plants. UPOV, Geneva, Switzerland

UPOV 1999. Guidelines for the conduct of tests for distinctness, uniformity and stability; maize. TG/2/6 + Corr. (Revision of TG/2/4). UPOV, Geneva, Switzerland

UPOV 2000. UPOV-WIPO roving seminar on the protection of plant varieties under the UPOV Convention, the patent system and the TRIPS agreement. UPOV-WIPO/RO/00/2 Geneva, Switzerland

UPOV 2002a. General introduction to the examination of distinctness, uniformity and stability and the development of harmonizes descriptions of new varieties of plants. TG/1/3 UPOV, Geneva, Switzerland

UPOV 2002b. Ad Hoc subgroup of technical and legal experts on Biochemical and Molecular Techniques. TC/38/14 – CAJ/45/5, UPOV, Geneva, Switzerland

UPOV 2003. Progress Report on the work of the technical committee, the technical working parties and the working group on Biochemical and Molecular Techniques, and DNA-profiling in particular. C/37/10, UPOV, Geneva, Switzerland

UPOV 2009. Guidelines for the conduct of tests for distinctness, uniformity and stability; maize. TG/2/7. UPOV, Geneva, Switzerland

van Etten J, 2006. Changes in farmers' knowledge of maize diversity in highland Guatemala, 1927/37-2004. Jour Ethnobiology and Ethnomedicine. 2: 12 (http://ethnobiomed.com/content/2/1/12).

van Hintum TJL, Jongen MWM, Hazekamp T, 1995. Standardization in Plant Genetic resources Documentation. Report of the second technical meeting of focal points for documentation in East European genebanks. Centre for Genetic resources, The Netherlands, (CGN), Wageningen, The Netherlands

van Wijk A, 2003. Implementation of Plant Variety Protection. WIPO-UPOV Symposium on Intellectual Property Rights in Plant Biotechnology. Geneva, October 24, 2003 WIPO-UPOV/SYM/03/10

Williams P, 2006. Marketing Wheat by Electronics I. Electronic classification of wheat. Grain Kernels 3: 1-3. PDK Projects Inc, Nanaimo, British Columbia, Canada.

White J, Law JR, 1991. Differentiation between varieties of oilseed rape (*Brassica napus* L.) on the basis of the fatty acid composition of the oil. Plant Varieties and Seeds (United Kingdom) 4(3): 125-132