Genetic diversity trends in twentieth century crop cultivars: a meta analysis

Mark van de Wouw · Theo van Hintum · Chris Kik · Rob van Treuren · Bert Visser

Abstract In recent years, an increasing number of papers has been published on the genetic diversity trends in crop cultivars released in the last century using a variety of molecular techniques. No clear general trends in diversity have emerged from these studies. Meta analytical techniques, using a study weight adapted for use with diversity indices, were applied to analyze these studies. In the meta analysis, 44 published papers were used, addressing diversity trends in released crop varieties in the twentieth century for eight different field crops, wheat being the most represented. The meta analysis demonstrated that overall in the long run no substantial reduction in the regional diversity of crop varieties released by plant breeders has taken place. A significant reduction of 6% in diversity in the 1960s as compared with the diversity in the 1950s was observed. Indications are that after the 1960s and 1970s breeders have been able to again increase the diversity in released varieties. Thus, a gradual narrowing of the genetic base of the varieties released by breeders could not be observed. Separate analyses for wheat and the group of other field crops and separate analyses on the basis of regions all showed similar trends in diversity.

Introduction

In the last century, scientific plant breeding has made an enormous impact on the agricultural landscape. In the early twentieth century knowledge about hybrids and mutations, and the application of Mendel’s work on inheritance was instrumental in a dramatic leap forward in plant breeding (Murphy 2007). Breeders have continuously introduced new varieties with higher yields and which are adapted to changing farming systems and user demands. This has contributed to a large extent to the major increases in agricultural productivity which have been observed in the twentieth century (Dudley 1994). Especially the Green Revolution of the 1960s and 1970s was a very important achievement of plant breeders, contributing to global food security (Evenson and Gollin 2003).

However, concern has been raised that the major breeding efforts in the twentieth century have been a strong force in the reduction of crop genetic diversity (Gepts 2006). It is generally thought that continuous selection among crosses of genetically related cultivars has led to a narrowing of the genetic base of the crops on which modern agriculture is based, contributing to the genetic erosion of the crop gene pools on which breeding is based (Plucknett et al. 1987).

In the past, genetic uniformity of crops has led to several devastating attacks of pests and diseases. Well-known examples are the potato blight epidemic in Ireland in the 1840s, and the corn leaf blight which devastated maize production in the USA in the 1970s (Lopez 1994). New strains of old diseases might threaten future agriculture productivity. A new strain of stem rust is now a cause of concern to wheat growers (Singh et al. 2006b) and it is feared that the global banana production will face severe losses in the near future due to a new strain of Panama...
disease which is attacking the very widely planted ‘Cavendish’ banana clone (Ploetz 2006). Many breeders are very much aware of these painful reminders of the importance of diversity, but it is not clear to what extent scientific plant breeding has been instrumental in reducing crop diversity any further since the switch from traditional landraces and farmers’ varieties to modern cultivars.

In the last decade, the availability of new molecular techniques and the interest in crop diversity have led to many published scientific papers addressing trends in diversity in released crop varieties. Diversity levels as found in these studies often fluctuate strongly from one time period to the next (Christiansen et al. 2002; Tian et al. 2005; White et al. 2008). Papers from the same crop and same regions sometimes show conflicting results (Huang et al. 2007; Roussel et al. 2004) and in general no clear pattern emerges from these studies, as both increases and decreases in diversity are observed (e.g. Fu et al. 2006; Hazen et al. 2002; Hysing et al. 2008).

A traditional, qualitative and narrative, literature review of publications on the impact of breeding on diversity trends in crops demonstrated the difficulty in drawing general conclusions, although the tentative conclusion was that the possible diversity reduction over time was most likely small (Fu 2006). In a traditional narrative review or by using “vote counting” methods (e.g. Peters et al. 2009; Stewart et al. 2004), the direction and magnitude of the effect which is being studied is difficult to estimate as the number of significant outcomes in the studies has little direct relation with the size of the effect (Rosenberg et al. 2000). A meta analysis can overcome these limitations by using formal statistical techniques for combining the results of independent experiments in order to reach general conclusions.

In the current study, meta analysis techniques were used to distil a general trend in the genetic diversity as output by breeders in the form of released varieties. The aim of this study was to quantify the impact plant breeding has had in the last century on the genetic diversity of crops using the large body of literature which has appeared on the subject in recent years. As this is the first meta analysis using diversity indices, the meta analytical methods were adapted for use with these indices.

Materials and methods

Selection of studies

A meta analysis starts with the selection of appropriate studies for inclusion in the analysis. Using Scopus, Web of Sciences and CAB abstracts a literature search was made for publications on diversity trends in modern breeding in the twentieth century using molecular techniques. Publications based on morphological diversity or pedigree analysis were not included. Keywords used in these searches were, among others, genetic erosion, impact of breeding and diversity trends. The search was limited to publications in the following languages: English, French, Spanish, Portuguese, German and Dutch. When a relevant study was found, papers which were cited by the study, as well as papers which cited the study were checked, to obtain an as complete set of papers as possible.

Extracting data

Diversity measures such as Nei’s gene diversity index (Nei 1973) or average genetic distances were extracted from the text, tables or figures of the papers selected. Variants of these measures were found in most papers. All these indices have in common that they look at evenness in the distribution of alleles, whereby diversity decreases when the distribution of alleles over the varieties gets more skewed. In some papers more than one diversity measure was used, in these cases preference was given to either Nei’s gene diversity index or average distances measured with the simple matching coefficient, as these two measures are identical (Kosman 2003). Many of the diversity indices which were used are essentially the same and only needed a simple transformation to obtain Nei’s gene diversity index. In addition to the diversity indices, other information extracted from the papers are the number of varieties studied for each period, the number of loci studied, the marker system used, the average PIC value (Anderson et al. 1993) of the markers, the region of study and the crop.

For the meta analysis a temporal division in decades was chosen. Although some studies also included varieties from before 1900 and after 2000, it was decided to limit the meta analysis to the twentieth century as both the varieties from the first decade of the twenty-first century and varieties released before the 1900s were not well represented. Since for the older decades only very few varieties were included in the studies, it was decided to combine the first three decades into a single group, 1900–1929. The meta analysis was limited to the diversity in modern cultivars and diversity data on historical landraces were excluded.

In many cases, the definitions of the time periods in the publications did not follow a decadal division or the diversity data were not reported in a way which allowed them to be directly used in the meta analysis. In one case a simple recalculation could transform the data into a decadal division (Priolli et al. 2004). In all other cases where a decadal division was not followed or the relevant data were difficult to extract from the paper, authors were contacted to obtain missing information or some clarification about...
the figures or the analysis. Many authors responded by recalculating their data to follow the decadal divisions and the use of adjusted Nei’s diversity index or by providing the original data (Table 1).

When raw data were available either from the authors or directly extracted from some of the papers (Metakovsky et al. 1991, 1994, 2000; Novoselskaya-Dragovich et al. 2003), they were analysed using spreadsheet software or Arlequin 3.1 (Excoffier et al. 2005) to obtain Nei’s adjusted gene diversity index on a decade by decade basis.

Not all authors responded or were able to provide the requested assistance, and these papers were reviewed to find other ways to fit them in the meta analysis. In studies where diversity was not exactly split into decades, the diversity values found were assigned to the most closely related decade (Bauer et al. 2007; Novoselskaya-Dragovich et al. 2007; Reif et al. 2005a, b). In cases where the reported diversity measure was more than one decade, they were assumed to be constant for these decades (Figuilo et al. 2007; Mantegazza et al. 2008; Nersting et al. 2006; Novoselskaya-Dragovich et al. 2007; Reif et al. 2005a). In the subsequent meta analysis these values are used as relative to the other values of pooled decades in the same study. In these cases the number of cultivars was divided over the decades to avoid an additional weighting of these diversity indices in the subsequent analyses. Obviously, diversity indices derived from pooled decades were not used to compare the decades within the pool. The other way round, when the time period in the study was shorter than needed for the meta analysis, the average diversity of the time periods belonging to a decade was calculated for use in the meta analysis (Figuilo et al. 2007). One paper did not indicate the number of cultivars analyzed for each time period, and in this case the number of cultivars analyzed for each decade was assumed to be proportional to the total number of cultivars used in the study (Novoselskaya-Dragovich et al. 2007). From Kim et al. (2005) the average dissimilarities were used by transforming the reported average similarities, as in this paper the reported PIC values were calculated by averaging over primers and not over loci.

Several papers which were originally selected for the meta analysis had to be rejected at closer inspection as the data in the papers were not presented in a way that made them usable in the meta analysis (Khlestkina et al. 2004; Koebner et al. 2003; Smýkal et al. 2008). One other paper (Martos et al. 2005) was rejected on closer inspection as the sample size used was so small that a further division in decades was not sensible.

It appeared that several times the same sets of varieties were used in different studies using different molecular marker systems (Donini et al. 2000; Cooke and Law 1998; Fu et al. 2003b, 2004, 2005, 2006; Kolodinska Brantestam et al. 2003, 2004, 2007). To avoid an implicit additional weighting of these sets of varieties, the results of the studies were averaged based on the number of loci in each study. Several publications have used more than one marker system. In cases where separate analyses were reported for each marker system (Donini et al. 2000; Manifesto et al. 2001; Zheng et al. 2003), the analyses were combined by taking weighted averages of the diversity indices based on the number of loci studied in each marker system.

A single publication might contribute more than once to the meta analysis if a separate analysis was done for different regions (Huang et al. 2007; White et al. 2008) or if more than one crop type was studied (Kim et al. 2005; Le Clerc et al. 2006). Also in a paper on wheat (Novoselskaya-Dragovich et al. 2007) two different regions were studied, but as the Italian data originated from another paper (Metakovsky et al. 1994) which had already been included in the meta analysis, these were disregarded. For papers in which in addition to an overall analysis, also an analysis split into regions was done, only the results from the overall analysis were used in the meta analysis (e.g. Kolodinska Brantestam et al. 2004).

Effect size and study weight

Using the diversity values extracted from the publications, the effect size for use in the meta analysis was calculated as the log response ratio:

\[ E_{ij,k} = \ln \left( \frac{H_{i,k}}{H_{j,k}} \right) \]

where \( E_{ij,k} \) is the effect size of study \( k \) and decades \( i \) and \( j \) and \( H_{i,k} \) and \( H_{j,k} \) are the diversity indices from study \( k \) in respectively decade \( i \) and \( j \). The natural log of the response ratio was used as this has preferable statistical properties, since the log ratio is equally affected by changes in either numerator or denominator (Hedges et al. 1999). The diversity effect was calculated for all possible combinations of decades.

In a meta analysis weighting of effect sizes from the individual studies occurs either on the basis of the reciprocal of the sampling variances (Gurevitch et al. 2001) or, when sampling variances are not available, a weight might be based on sample size (Sánchez-Meca and Marín-Martínez 1998) or some other assessment of reliability (Manly 2001). In the current meta analysis a weight on the basis of the sampling variance was not an option as sampling variances are often not mentioned.

Diversity indices, as used in the current study, show two dimensions to the sampling process: the number of varieties used in each study and the number of loci studied (Nei 1987). A variance curve will, with increasing sample sizes...
Table 1  Studies on diversity trends in crops using molecular techniques

| Reference                  | Crop   | Region                  | Time period     | No of Cultivars | Notes                                           |
|----------------------------|--------|-------------------------|-----------------|-----------------|------------------------------------------------|
| Bauer et al. (2007)        | Maize  | Serbia                  | 1966–2001       | 24              | Temporal adjustments made                       |
| Christiansen et al. (2002) | Wheat  | Nordic Countries        | 1901–1993       | 75              | Original data received from authors             |
| Condón et al. (2008)       | Barley | Mid West, USA           | 1958–1998       | 61              | Original data received from authors             |
| Cooke and Law (1998)       | Wheat  | UK                      | 1930–1999       | 55              | Data from paper                                 |
| Donini et al. (2000)       | Wheat  | UK                      | 1930–1999       | 55              | Diversity data from Law et al. (1997)           |
| Feng et al. (2006)         | Maize  | USA                     | 1930–2004       | 53              | Recalculated by authors                         |
| Figliuolo et al. (2007)    | Wheat  | Italy                   | 1900–1999       | 131             | Temporal adjustments made                       |
| Fu et al. (2003a)          | Oat    | Canada                  | 1886–2001       | 96              | Original data received from authors             |
| Fu et al. (2003b)          | Flax   | North America           | 1908–1998       | 54              | Original data received from authors             |
| Fu et al. (2004)           | Oat    | Canada                  | 1886–2001       | 96              | Original data received from authors             |
| Fu et al. (2005)           | Wheat  | Canada                  | 1845–2004       | 75              | Original data received from authors             |
| Fu et al. (2006)           | Wheat  | Canada                  | 1845–2004       | 75              | Original data received from authors             |
| Fu et al. (2007)           | Soybean| Canada                  | 1934–2001       | 45              | Original data received from authors             |
| Hao et al. (2006)          | Wheat  | China                   | 1950–1999       | 968             | Data from paper                                 |
| Hazen et al. (2002)        | Wheat  | Shaanxi, China          | 1950–1999       | 23              | Data from paper                                 |
| Huang et al. (2007)        | Wheat  | UK & Europe             | 1940–1999       | 282             | Data from paper                                 |
| Hysing et al. (2008)       | Wheat  | Nordic Countries        | 1900–2003       | 166             | Recalculated by authors                         |
| Khlestkina et al. (2004)   | Wheat  | Siberia                 | 1926–1999       | 54              | Excluded from the analysis                      |
| Kim et al. (2005)          | Barley | Korea                   | 1932–1999       | 44              | Data from paper                                 |
| Koebner et al. (2003)      | Barley | UK                      | 1920–1999       | 134             | Excluded from the analysis                      |
| Kolodinska Brantestam et al. (2003) | Barley | Nordic & Baltic countries | 1900–2000       | 240             | Original data received from authors             |
| Kolodinska Brantestam et al. (2004) | Barley | Nordic & Baltic countries | 1900–2000       | 227             | Original data received from authors             |
| Kolodinska Brantestam et al. (2007) | Barley | Nordic & Baltic countries | 1900–2000       | 197             | Original data received from authors             |
| Landjeva et al. (2006)     | Wheat  | Bulgaria                | 1925–2003       | 91              | Recalculated by authors                         |
| Le Clerc et al. (2005)     | Maize  | France                  | 1950–1999       | 114             | Recalculated by authors                         |
| Le Clerc et al. (2006)     | Maize  | France                  | 1970–2004       | 2137            | Recalculated by authors                         |
| Le Clerc et al. (2006)     | Pea    | France                  | 1950–2004       | 268             | Recalculated by authors                         |
| Malysheva-Otto et al. (2007) | Barley | Europe                  | 1900–1999       | 510             | Recalculated by authors                         |
| Manifesto et al. (2001)    | Wheat  | Argentina               | 1938–1995       | 105             | Original data received from authors             |
| Mantegazza et al. (2008)   | Rice   | Italy                   | 1850–2001       | 135             | Temporal adjustments made                       |
| Martos et al. (2005)       | Wheat  | Italy & Spain           | 1900–2000       | 24              | Excluded from the analysis                      |
| Metakovsky et al. (1991)   | Wheat  | Yugoslavia              | 1970–1988       | 38              | Recalculated from data in paper                 |
| Metakovsky et al. (1994)   | Wheat  | Italy                   | 1930–1990       | 121             | Recalculated from data in paper                 |
| Metakovsky et al. (2000)   | Wheat  | Spain                   | 1950–2000       | 97              | Recalculated from data in paper                 |
| Metakovsky and Branlard (1998) | Wheat | France                  | 1945–1992       | 187             | Original data received from authors             |
| Nersting et al. (2006)     | Oat    | Nordic                  | 1898–2000       | 64              | Temporal adjustments made                       |
| Novoselskaya-Dragovich et al. (2003) | Wheat | SE Russia               | 1924–2000       | 39              | Recalculated from data in paper                 |
| Novoselskaya-Dragovich et al. (2007) | Wheat | Serbia                  | 1960–1999       | 94              | Temporal adjustments made                       |
| Prioli et al. (2004)       | Soybean| Brazil                  | 1970–1999       | 184             | Data from paper                                 |
| Qi et al. (2006)           | Rice   | China                   | 1950–1999       | 257             | Data from paper                                 |
| Reif et al. (2005a)        | Maize  | Central Europe          | 1951–2001       | 85              | Temporal adjustments made                       |
| Reif et al. (2005b)        | Wheat  | Developing countries    | 1950–1997       | 123             | Temporal adjustments made                       |
| Roussel et al. (2004)      | Wheat  | France                  | 1840–2000       | 497             | Original data received from authors             |
| Roussel et al. (2005)      | Wheat  | Europe                  | 1840–1999       | 480             | Original data received from authors             |
| Russell et al. (2000)      | Barley | UK                      | 1884–1999       | 95              | Original data received from authors             |
| Smykal et al. (2008)       | Pea    | Czech & Slovak Rep.     | 1950–2000       | 164             | Excluded from the analysis                      |
| Tian et al. (2005)         | Wheat  | China                   | 1940–1999       | 242             | Data from paper                                 |
and/or number of loci, reach a point whereby it will be very
difficult to decrease the variance any further, showing
diminishing marginal returns of an extra investment in
increasing sample size or number of loci. A study weight
needs to reflect all these aspects and we have therefore
developed the following formula, which is a modification
of a weight based on sample size:

\[
W_{i,j,k} = \frac{1}{\left(\frac{N^i_k + N^j_k}{N^i_k \times N^j_k} + \frac{C_1}{\text{PIC}_k \times N^i_k} + C_2\right)}
\]

where \(W_{i,j,k}\) is the weight of the effect size of decades \(i\) and \(j\)
of study \(k\); \(N^i_k\) and \(N^j_k\) the number of varieties in decade \(i\) and
decade \(j\) of study \(k\); \(\text{PIC}_k\) the average PIC value of study \(k\);
\(N^2_k\) the number of loci in study \(k\); and \(C_1\) and \(C_2\) are the
constants.

The first part of the denominator is a reflection of a weight
based on sample size developed for use in a meta analysis
(Hedges and Olkin 1985), which has been found to be an
improvement on using the total sample size only (Sánchez-
Meca and Marín-Martínez 1998). The second part of the
denominator reflects the variance reduction obtained by
increasing the number of loci in the study. Also the molec-
ular marker system that has been used plays an important
role, as the potential number of alleles per locus depends on
the type of marker system used. The average Polymorphic
Information Content (PIC) value of all loci used in the study
was therefore added as an additional weighting of the
number of loci. The first constant reflects the relative con-
tribution of the number of loci and the number of varieties in
the reduction of the sampling variance. The second constant
sets a limit to the maximum weight possible, and avoids in
this way that very large studies would dominate the final
results. This latter constant can be viewed as a reflection of
any other components in the variance not covered in the first
two parts of the denominator, such as the effect of the crop or
the region. The final weight of each study is obtained by
adding the various components, reflecting the way the
variance of the diversity indices is constructed (Nei 1987),
and then taking the inverse.

For this study the first constant was set to a value of 2
based on the different recommendations for sample size
(10–20 samples, Singh et al. 2006a; Zhao et al. 2006), and
number of loci (30 for SSRs and 70 for AFLP loci). The
second constant was set to a value of 0.4, resulting in a
maximum weight of 2.5 for a study in which the other
components are approaching 0. In the current meta analysis
this resulted in a fivefold difference in weight between the
smallest and largest study. The influence of small varia-
tions of the constants in the weight formula on the results
of the meta analysis was tested.

### Meta analysis

Using the Metawin 2.0 program (Rosenberg et al. 2000)
95% bias corrected bootstrap confidence intervals were calculated for all mean effect sizes obtained from the pair
wise comparisons of decades, and a significant effect was
concluded where the confidence interval did not intersect 0.

Both unweighted and weighted analyses were carried
out, and the results compared. The data were explored for a
possible publication bias using a modified funnel plot
(Light and Pillemer 1984), whereby the effect sizes are
plotted against the study weight. The data were further
checked for possible underlying structures by classifying
studies into groups according to crop group or region and
carrying out separate analyses for each group.

The relative diversity values for the eight decades that
collectively best explained the total set of decadal com-
parisons were estimated by an optimisation procedure
using all effect sizes and their weights. In this procedure at
first the estimated effect sizes between the decades were all
set to zero. Then the weighted sum of squared differences
between estimated and observed effect sizes for the total
set of pair wise comparisons was minimized by an iterative
process in which the estimated values were varied in small
steps. The estimated effect sizes belonging to the minimal
sum of squared differences was considered to best fit the
observed effect sizes and from these values the diversity
values for each decade were calculated, setting the diver-
sity in the decade with the lowest diversity to 100. The
optimisation procedure was carried out by a tailor made
program written in Turbo Pascal (available upon request)
and the results were plotted in graphs.

### Results

A total of 48 publications were found which compare
varieties from different eras using molecular marker tech-
niques (Table 1). The earliest paper, based on seed pro-
teins, dated from 1991, but only four papers were published
before 2000 and the large majority of papers dated from the last few years. Although a variety of different crops have been studied, the studies have in common that they address field crops (mostly cereals) which are planted from seeds. Furthermore most of the studies focussed on North American or European varieties, although also studies from South America, Asia and Australia were found. A large array of molecular techniques has been used in these studies. Large fluctuations in diversity were often observed in studies based on only few loci or few cultivars (e.g. Christiansen et al. 2002; Hazen et al. 2002; Novoselskaya-Dragovich et al. 2003).

A total of 732 pair-wise comparisons of diversity levels between sets of decades were extracted from the publications. The more recent decades showed the highest number of comparisons (up to 42), while the older decades were represented in only few studies and as a result the number of comparisons remained as low as 13 (Tables 2, 3).

The unweighted meta analysis showed a maximum difference of almost 9% between the decade with the lowest and the decade with the highest genetic diversity. A clear and significant reduction in diversity could be observed in the 1960s (Fig. 1a), the level of diversity being significantly lower then in the three previous decades (Table 2). The diversity appeared to recover in the periods thereafter, and in the 1990s it was significantly higher then it was in the 1960s. The weighted meta analysis showed similar trends (Fig. 1b), but in general smaller changes in diversity between decades were observed. Also in the weighted meta analysis a low diversity was observed in the 1960s, with a significant reduction in diversity from the 1950s to the 1960s and 1970s (Table 3). The oldest period showed comparatively low diversity in the released varieties. However, relatively few studies included varieties from this time period, and the values obtained did not show significant differences with any of the other decades.

Varying constant $C_1$ in the weight formula between 1.5 and 3.5 and constant $C_2$ between 0.2 and 0.6 resulted in only small differences in the final results, demonstrating that the weight is reasonably robust under small fluctuations in these constants.

The funnel plot (Fig. 2) followed the expected pattern of more variation in the effect size observed in the smaller sized studies. The larger studies showed small diversity differences of less then 10%, while large diversity effects are mostly limited to decadal comparisons with few loci and few varieties. The difference in the diversity effect

Table 2 Number of comparisons of diversity for pairs of decades (above diagonal) and significance of the effect size using 95% bootstrap confidence intervals in the unweighted meta analysis (below diagonal, *s* significant, *ns* not significant)

| Decade | <1929 | 1930s | 1940s | 1950s | 1960s | 1970s | 1980s | 1990s |
|--------|-------|-------|-------|-------|-------|-------|-------|-------|
| <1929  | 14    | 13    | 15    | 16    | 16    | 16    | 16    | 16    |
| 1930s  | *ns*  | 16    | 19    | 23    | 25    | 25    | 25    | 25    |
| 1940s  | *ns*  | *ns*  | 19    | 24    | 26    | 26    | 26    | 26    |
| 1950s  | *ns*  | *ns*  | *ns*  | 30    | 36    | 35    | 36    | 36    |
| 1960s  | *s*   | *s*   | *s*   | 35    | 37    | 37    | 38    | 38    |
| 1970s  | *ns*  | *ns*  | *ns*  | *ns*  | *ns*  | *ns*  | *ns*  | 41    |
| 1980s  | *ns*  | *ns*  | *ns*  | *ns*  | *ns*  | *ns*  | *ns*  | 42    |
| 1990s  | *ns*  | *ns*  | *ns*  | *ns*  | *ns*  | *ns*  | *ns*  | *ns*  |

Table 3 Number of comparisons of diversity for pairs of decades (above diagonal) and significance of the effect size using 95% bootstrap confidence intervals in the weighted meta analysis (below diagonal, *s* significant, *ns* not significant)

| Decade | <1929 | 1930s | 1940s | 1950s | 1960s | 1970s | 1980s | 1990s |
|--------|-------|-------|-------|-------|-------|-------|-------|-------|
| <1929  | 14    | 13    | 15    | 16    | 16    | 16    | 16    | 16    |
| 1930s  | *ns*  | 16    | 19    | 23    | 25    | 25    | 25    | 25    |
| 1940s  | *ns*  | *ns*  | 19    | 24    | 26    | 26    | 26    | 26    |
| 1950s  | *ns*  | *ns*  | *ns*  | 30    | 36    | 35    | 36    | 36    |
| 1960s  | *ns*  | *ns*  | *ns*  | *s*   | 35    | 37    | 37    | 38    |
| 1970s  | *ns*  | *ns*  | *ns*  | *s*   | *ns*  | 41    | 42    | 42    |
| 1980s  | *ns*  | *ns*  | *ns*  | *ns*  | *ns*  | *ns*  | *ns*  | *ns*  |
| 1990s  | *ns*  | *ns*  | *ns*  | *ns*  | *ns*  | *ns*  | *ns*  | *ns*  |

Fig. 1 Crop genetic diversity in the twentieth century based on an unweighted (a) and a weighted (b) meta analysis of 44 publications. The diversity in the decade with the lowest diversity was set to 100.
found in the weighted (Fig. 1b) and unweighted (Fig. 1a) meta analysis form another indication that the smaller studies showed larger diversity differences between the decades. The shape of the funnel plot does not indicate a publication bias, as also small effects are reported for the smaller sized studies.

Wheat has been the most popular crop for studies on genetic diversity trends as 421 out of a total of 732 decadal comparisons in this meta analysis involved wheat. To check whether the wheat studies dominated the final result, two additional analyses were carried out with only the wheat studies and excluding the wheat studies respectively (Fig. 3a, b). Although in these smaller analyses no significant effects could be observed, the general trends were similar, with a low diversity around the 1960s still visible in both cases. The analyses of the wheat studies showed a relatively later, in the 1990s, recovery of the genetic diversity, while the other crops showed this recovery already in the 1970s.

Two regions with a larger number of studies (North America and Europe) were also analysed separately to check for regional differences (results not shown). Also the regional analyses showed similar general trends as were observed in the overall analysis. In the North American group a dip in diversity was observed in the 1970s, albeit without significant effects. The European group, which included the largest number of studies, had its lowest diversity in the 1960s, which significantly differed from the diversity as observed in the 1950s.

Discussion

Meta analysis

Studies with small sample sizes and modest effects are vulnerable to a negative bias (Gurevitch et al. 2001), that is a bias against detecting true effects. Many of the studies used in this meta analysis had small sample sizes, and the overall diversity effect was not very large. Under such circumstances, “vote counting” techniques or narrative reviews of results would easily result in the true effects remaining unrecognized. Indeed, a narrative review of some of the publications used in this meta analysis (Fu 2006) did not recognize the significant decrease in diversity in the 1960s and 1970s as we found in this analysis.

Contacting authors for clarifications and missing information for use in a meta analysis has been reported to give in general a very low level of return (Gurevitch et al. 2001). In the current meta analysis this has not been the case, as an unexpectedly high level of cooperation with the authors of the original studies has been experienced, whereby authors handed over original data or recalculated their data for use in the meta analysis. The level of cooperation might vary according to the professional field to which the meta analysis applies, but it appears worthwhile, when conducting a meta analysis, to at least try contacting authors. Apparently in the field of plant breeding and agriculture the level of cooperation can be extremely high, possibly as the result of a remaining spirit of open access to research results.
A weighted meta analysis on the basis of sample size where large studies contribute more than small studies, has shown to be an improvement compared to an unweighted meta analysis (Hunter and Schmidt 2004). However, the approach could be troublesome if one study would have an extremely large sample size as compared to all the other studies, as this study would then completely dominate the final result of the meta analysis. If at the same time this large study is deviant in a certain way, or uses a very specific sample group which is not representative for the group as a whole, the weighted meta analysis might result in wrong conclusions. In the current meta-analysis we introduced a constant \( C_2 \) in the weight formula, which effectively sets a limit to the maximum weight possible, thereby avoiding that a very large study will entirely define the result of the meta analysis, while still maintaining much of the added accuracy of weighting in the meta analysis.

No evidence for a publication bias in the form of under reporting of non significant results was found. This type of publication bias does not seem likely to have influenced the current meta analysis, given that many studies on which this meta analysis is based do not report clear significant differences or do not mention confidence levels of the results at all.

Trends in diversity

Although individual studies might report large fluctuations in diversity, this meta analysis shows that overall no major reduction in diversity as released by plant breeders has taken place. We did find a significant reduction in the diversity of released varieties in the 1960s, but even then the diversity reduction as compared to the diversity levels in the 1950s is only 6%. If this would have been the start of a continuing trend in reduction in diversity thereafter, very severe losses in diversity would have occurred by now. However, indications are that after the 1960s and 1970s breeders have been able to again increase the diversity of released varieties. An overall gradual narrowing of the genetic base of the varieties released by breeders has not been observed.

In the 1960s and 1970s the introduction of the new Green Revolution-type cultivars for the major staple crops led to concerns on the disappearance of the world’s varietal wealth of crop plants (Harlan 1970; Pistorius 1997). The widely shared concerns ultimately resulted in the establishment of a worldwide network of international genebanks hosted by the CGIAR research centres. The seed samples stored in these genebanks facilitated access of the world’s crop diversity to plant breeders world wide. It seems likely that the easy access to crop diversity provided by the genebanks, improved communication among breeders and easier exchange of seeds were factors contributing to the reversal of the initial trend in diversity reduction as observed in this meta analysis. Also the increased use of crop wild relatives for breeding and in recent years the use of synthetic wheats will have contributed to the observed diversity increase.

The low diversity observed in the meta analysis among the cultivars released before 1930 might have been caused by an initially low output of breeding programmes, as this is the time when scientific breeding started to gain momentum, but it could as well be the result of sampling problems. The diversity as observed in the oldest group is based on few studies, with often few varieties, and therefore yields an unreliable estimate of the diversity, which resulted in no significant effects and large fluctuations in diversity between subgroups of publications in this meta analysis.

The availability of historic cultivars is essential for temporal studies. The longer the period covered in the study, the more likely that part of the historic cultivars is no longer available and cannot be included in the study. A bias might thus be introduced, as those cultivars which can still be found in collections, might be in particular those cultivars which have been used for subsequent breeding and are therefore more closely related to the newer cultivars. Old cultivars which were not used in breeding programmes might have been lost, which could cause an underestimate of the diversity in the early decades. An additional problem is that the historic cultivars will have gone through many regeneration cycles in order to maintain viable seeds. As a result, genetic integrity might have been compromised and some diversity might have been lost due to genetic drift or selective pressures during regeneration (van Hintum et al. 2007).

Cultivars released in the 1950s were probably still widely used at the time the large genebanks were set up and are probably well represented in the germplasm collections. This might explain the high difference in numbers of studies that include cultivars of the 1940s and cultivars of the 1950s found in this meta analysis. One might also expect that cultivars which were in use when the large gene banks were set up might more closely resemble the original material, based on growing awareness of proper seed storage and regeneration protocols.

The diversity measures used in this meta analysis are an indication of evenness based on the frequencies of alleles in the group of cultivars studied. Maintenance of diversity levels does not mean that no alleles get lost at all, since already rare alleles might still have disappeared and a shift in alleles might have occurred. Whether the alleles which have been lost are in fact useful is not known, but they might be of importance for future plant breeding efforts and therefore it is useful to maintain the varieties which contain these alleles in ex situ collections.
Scope of the meta analysis

Two stages can be recognized in the genetic erosion of crop diversity due to a modernization bottleneck: the initial replacement of landraces by modern cultivars and further trends in diversity as a consequence of modern breeding (van de Wouw et al. 2009). Whereas this meta analysis focused on the second stage of the modernization bottleneck and showed that little or no loss of diversity occurred in varieties released during the twentieth century by the formal sector, this leaves the question unanswered how much crop diversity was lost at the time of shift from landraces and farmers’ varieties to modern varieties. Evidence exists that such loss of diversity did occur (e.g. Reif et al. 2005b; Roussel et al. 2004) and this loss might be the major component of genetic erosion, threatening both the sustainability of agricultural production and the raw material for future plant breeding.

The publications used for this meta analysis all focused on the diversity at a regional level, and the meta analysis has shown that in general diversity has been maintained. Whether this conclusion can be extrapolated to trends in global crop diversity remains to be determined. A reduction of the diversity in the global set of released varieties might still occur if unique regional alleles are replaced by globally common alleles, even though at a regional level diversity levels have been maintained. The increased dominance of large multinational seed companies in breeding (Pistorius and van Wijk 1999) might be instrumental in a possible diversity reduction at a global scale if the increased role of these companies means that increasingly similar cultivars are released in different regions.

The studies used as input in this meta analysis have focused on the time of release of the cultivars. Although this reflects the amount of diversity released by the breeders in a specific decade, it is also a reflection of the level of breeding activity in that decade for the crop under study. Less diversity will be released during periods with little breeding activity, even though the total number of cultivars available to farmers and the diversity at farm level or at seed producers’ level might not have changed as a result of the persistence of older varieties in the farming systems.

Although breeders have an important role in supplying diverse material, farmers make the ultimate varietal choices. The choice of the farmers can still strongly limit the diversity grown in the fields if most farmers select the same varieties. This danger is especially high in a period in which large advances in breeding are made, and in which a few of the newly released cultivars outcompete other varieties because of a significant higher yield or an unique resistance to an important pest or disease. This has happened in Australia at the time of introduction of new midge resistant sorghum hybrids, which were planted by more than 80% of the farmers. The move to these hybrids was associated with a narrowing of genetic diversity (Jordan et al. 1998). The dip in diversity observed in this meta analysis in the 1960s, which coincides with the introduction of the short straw cereal cultivars, might have been aggravated by the varietal choices made by farmers.

Often it is not clear how well the cultivars from the studies used in this analysis represent the cultivars actually released. Information on the total number of cultivars which have been released in the studied time periods is often lacking. Therefore it remains often unknown what part or proportion of the released diversity has been studied although in some studies it is stated that the analysed cultivars are representative of the cultivars which were used the most (e.g. Reif et al. 2005a; Russell et al. 2000). Only one study on maize and pea (Le Clerc et al. 2006) looked at the complete set of cultivars listed in the French catalogue. This study showed a very large increase in the number of varieties listed in the last 50 years.

Although a higher number of released cultivars does not necessarily translate in a higher diversity, allelic richness will most probably be higher when the group of cultivars is larger. The basis of plant variety protection is the DUS test, which stipulates that a new variety has to be distinct, uniform and stable. The distinctness principle ensures that a new variety contributes new diversity to the existing pool of varieties. This can be just a new combination of alleles already present in the existing varieties, but there is a strong likelihood that also new alleles are introduced, therefore increasing the richness of the group of released varieties.

In most of the studies selected for the meta analysis anonymous markers are used. These markers target randomly both non-coding and coding regions of the genome. In two studies on Canadian wheat cultivars both anonymous SSR (Fu et al. 2005) and EST-derived SSR markers (Fu et al. 2006) were used to study the same group of cultivars. Very similar results were found using either marker set, which seems to indicate that anonymous markers are also a reflection of the diversity found in functional regions of the genome. More research would be needed to confirm this finding.

The studies used in this meta analysis were based on eight different crop species. These species have in common that they are all major field crops grown from seeds with an emphasis on cereals. No studies on temporal trends in diversity as a result of breeding were found for horticultural crops. Whether horticultural crops will exhibit the same general diversity trends as field crops remains unclear. In horticultural crops breeding objectives are more consumer directed, while in field crops breeding objectives are mostly grower directed (Janick 2005). These different breeding objectives may have caused different trends in diversity. Diversity trends for horticultural crops might
follow socio-economic trends, by which more affluent consumers pose higher quality demands and will expect a wider choice in the horticultural crops they consume.

Concluding remarks

No general trends pointing at a loss in regional genetic diversity as released by breeders in the last century have been observed in this meta analysis. However, this does not rule out the possibility of diversity loss for specific crops and regions. Also, as the technical possibilities in plant breeding are advancing rapidly, it is unclear what will happen with future crop diversity levels. New techniques leading to the incorporation of genes from organisms outside the primary and secondary gene pool of the crop and the use of synthetic wheats in wheat breeding could in theory lead to an increased diversity among the released varieties. However, the strong technical demands and large investments needed to apply these new techniques in developing new cultivars might also mean that breeding continues to become more centralized and run by an ever reducing numbers of breeding companies, a development which could then be counter effective with regards to diversity levels. It remains important that trends in diversity are monitored, and that a back-up system for the conservation of important crop diversity is in place.

Acknowledgments Drs Agnese Kolodinska, Federico Condón, Merethe Christiansen, Jorge Dubcovsky, Marcela Manifesto, Yong-Bi Fu, Gérard Branlard and François Balfourier very kindly provided the original data from their respective studies. Drs David Zhang, Lyudmilla Malysheva-Otto, Svetlana Landjeva, Mark Cooper, Jon White (on behalf of Pioneer Hi-Bred International Inc., National Institute of Agricultural Botany, Molecular Plant Breeding CRC and the Cooperative Research Centre for Value Added Wheat), Shu-Chin Hysing and Joëlle Lallemend recalculated diversity indices from their studies for use in the meta analysis. We would also like to thank all other researchers who were very helpful, even though they were not successful in recovering the data as people had moved into other jobs or retirement. Lastly I would like to thank Prof Dr Rolf Hoekstra for suggesting a meta analysis. This project was financed by the Dutch Ministry of Agriculture, Nature and Food Quality.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

Anderson JA, Churchill GA, Autrique JE, Tanksley SD, Sorrells ME (1993) Optimizing parental selection for genetic-linkage maps. Genome 36:181–186

Bauer I, Drinic SM, Drinic G, Micic DJ (2007) Assessing temporal changes in genetic diversity of maize hybrids using RAPD markers. Cereal Res Commun 35:1563–1571

Christiansen MJ, Andersen SB, Ortiz R (2002) Diversity changes in an intensively bred wheat germplasm during the 20th century. Mol Breed 9:1–11

Condón F, Gustus C, Rasmussen DC, Smith KP (2008) Effect of advanced cycle breeding on genetic diversity in barley breeding germplasm. Crop Sci 48:1027–1036

Cooke RJ, Law JR (1998) Seed storage protein diversity in wheat varieties. Plant Var Seeds 11:159–167

Donini P, Law JR, Koehn RMD, Reeves JC, Cooke RJ (2000) Temporal trends in the diversity of UK wheat. Theor Appl Genet 100:912–917

Dudley JW (1994) Plant breeding—a vital part of improvement in crop yields, quality and production efficiency. In: Frey KJ (ed) Historical perspectives in plant sciences. Iowa State University Press, Ames, pp 163–177

Evenson RE, Gollin D (2003) Assessing the impact of the Green Revolution, 1960 to 2000. Science 300:758–762

Excoffier L, Laval G, Schneider S (2005) Arlequin ver 3.0: an integrated software package for population genetics data analysis. Evol Bioinform Online 1:47–50

Feng L, Sebastian S, Smith S, Cooper M (2006) Temporal trends in SSR allele frequencies associated with long-term selection for yield of maize. Maydica 51:293–300

Figliuolo G, Mazzeo M, Greco I (2007) Temporal variation of diversity in Italian durum wheat germplasm. Genet Resour Crop Eval 54:615–626

Fu YB (2006) Impact of plant breeding on genetic diversity of agricultural crops: searching for molecular evidence. Plant Genet Resour Charact Util 4:71–78

Fu YB, Peterson GW, Scoles G, Rossnagel B, Schoen DJ, Richards KW (2003a) Allelic diversity changes in 96 Canadian oat cultivars released from 1886 to 2001. Crop Sci 43:1989–1995

Fu YB, Rowland GG, Duguid SD, Richards KW (2003b) RAPD analysis of 54 North American flax cultivars. Crop Sci 43:1510–1515

Fu YB, Kibite S, Richards KW (2004) Amplified fragment length polymorphism analysis of 96 Canadian oat cultivars released between 1886 and 2001. Can J Plant Sci 84:23–30

Fu YB, Peterson GW, Richards KW, Somers D, DePauw RM, Clarke JM (2005) Allelic reduction and genetic shift in the Canadian hard red spring wheat germplasm released from 1845 to 2004. Theor Appl Genet 110:1505–1516

Fu YB, Peterson GW, Yu JK, Gao LF, Jia JZ, Richards KW (2006) Impact of plant breeding on genetic diversity of the Canadian hard red spring wheat germplasm as revealed by EST-derived SSR markers. Theor Appl Genet 112:1293–1277

Fu YB, Peterson GW, Morrison MJ (2007) Genetic diversity of Canadian soybean cultivars and exotic germplasm revealed by simple sequence repeat markers. Crop Sci 47:1947–1954

Gepts P (2006) Plant genetic resources conservation and utilization: the accomplishments and future of a societal insurance policy. Crop Sci 46:2278–2292

Gurevitch J, Curtis PS, Jones MH (2001) Meta-analysis in ecology. Adv Ecol Res 32:199–247

Hao C, Wang L, Zhang X, You G, Dong Y, Jia J, Liu X, Shang X, Liu S, Cao Y (2006) Genetic diversity in Chinese modern wheat varieties revealed by microsatellite markers. Sci China Ser C Life Sci 49:218–226

Harlan JR (1970) Evolution of cultivated plants. In: Frankel OH, Bennett E (eds) Genetic resources in plants—IBP Hand- book No 11. International Biological Programme, London, pp 19–32

Hazen SP, Zhu L, Kim HS, Tang G, Ward RW (2002) Genetic diversity of winter wheat in Shaanxi Province, China, and other common wheat germplasm pools. Genet Resour Crop Evol 49:437–445
Koebner RMD, Donini P, Reeves JC, Cooke RJ, Law JR (2003)

Khlestkina EK, Roder MS, Efremova TT, Borner A, Shumny VK

Janick J (2005) Horticultural plant breeding: past accomplishments,
Hysing S-C, Sal T, Nybom H, Liljeroth E, Merker A, Orford S,
Hunter JE, Schmidt FL (2004) Methods of meta-analysis: correcting
Huang XQ, Wolf M, Ganal MW, Orford S, Koebner RMD, Roder MS
Hedges LV, Olkin I (1985) Statistical methods of meta-analysis.
Theor Appl Genet (2010) 120:1241–1252 1251

Light RJ, Pillemer DB (1984) Summing up: the science of reviewing
research. Harvard University Press, Cambridge
Qi Y, Zhang D, Zhang H, Wang M, Sun J, Wei X, Qiu Z, Tang S, Cao Y, Wang X, Li Z (2006) Genetic diversity of rice cultivars (Oryza sativa L.) in China and the temporal trends in recent fifty years. Chin Sci Bull 51:681–688

Reif JC, Hamrit S, Heckenberger M, Schipprack W, Maurer HP, Bohn M, Melchinger AE (2005a) Trends in genetic diversity among European maize cultivars and their parental components during the past 50 years. Theor Appl Genet 111:838–845

Reif JC, Zhang P, Dreisigacker S, Warburton ML, van Ginkel M, Hoisington D, Bohn M, Melchinger AE (2005b) Wheat genetic diversity trends during domestication and breeding. Theor Appl Genet 110:859–864

Roussel V, Koenig J, Beckert M, Balfourier F (2004) Molecular diversity in French bread wheat accessions related to temporal trends and breeding programmes. Theor Appl Genet 108:920–930

Roussel V, Leisova L, Exbrayat F, Stehno Z, Balfourier F (2005) SSR allelic diversity changes in 480 European bread wheat varieties released from 1840 to 2000. Theor Appl Genet 111:162–170

Russell JR, Ellis RP, Thomas WTB, Waugh R, Provan J, Booth A, Fuller J, Lawrence P, Young G, Powell W (2000) A retrospective analysis of spring barley germplasm development from ‘foundation genotypes’ to currently successful cultivars. Mol Breed 6:553–568

Sánchez-Meca J, Marín-Martínez F (1998) Weighting by inverse variance or by sample size in meta-analysis: a simulation study. Educ Psychol Meas 58:211–220

Singh M, Chabane K, Valkoun J, Blake T (2006a) Optimum sample size for estimating gene diversity in wild wheat using AFLP markers. Genet Resour Crop Evol 53:23–33

Singh RP, Hodson DP, Jin Y, Huerta-Espino J, Kinyua MG, Wanyera R, Njuu P, Ward RW (2006b) Current status, likely migration and strategies to mitigate the threat to wheat production from race Ug99 (TTKS) of stem rust pathogen. CAB Rev Perspect Agric Vet Sci Nutr Nat Resour 1:1–13

Smykal P, Hybl M, Corander J, Jarkovský J, Flavell A, Griga M (2008) Genetic diversity and population structure of pea (Pisum sativum L.) varieties derived from combined retrotransposon, microsatellite and morphological marker analysis. Theor Appl Genet 117:413–424

Stewart GB, Coles CF, Pullin AS (2004) Does burning degrade blanket bog? Systematic Review No. 1. Collaboration for environmental evidence. Birmingham, UK, p 30

Tian QZ, Zhou RH, Jia JZ (2005) Genetic diversity trend of common wheat (Triticum aestivum L.) in China revealed with AFLP markers. Genet Resour Crop Evol 52:325–331

van de Wouw M, van Hintum T, Kik C, van Treuren R, Visser B (2010) Genetic erosion in crops: concept, research results and challenges. Plant Genet Resour Charact Util. doi:10.1017/S1479262109990062

van Hintum TJL, van De Wiel CCM, Visser DL, van Treuren R, Vosman B (2007) The distribution of genetic diversity in a Brassica oleracea gene bank collection related to the effects on diversity of regeneration, as measured with AFLPs. Theor Appl Genet 114:777–786

White J, Law JR, MacKay I, Chalmers KJ, Smith JSC, Kilian A, Powell W (2008) The genetic diversity of UK, US and Australian cultivars of Triticum aestivum measured by DArT markers and considered by genome. Theor Appl Genet 116:439–453

Zhao R, Cheng Z, Lu W, Lu B (2006) Estimating genetic diversity and sampling strategy for a wild soybean (Glycine soja) population based on different molecular markers. Chin Sci Bull 51:1219–1227

Zheng YL, Zhang QZ, Wei YM, Wu W, Yan ZH (2003) Genetic diversity of Sichuan elite wheat cultivars based on microsatellites and STS-PCR markers. J Genet Breed 57:47–57