1. Introduction

Unpredictable and increasingly variable climatic conditions have highlighted the importance of developing new and innovative strategies for crop breeding and management [1]. Breeding for diversity has been gaining much attention in the recent past [2,3] due to the fact that diversity within varieties or crop populations should increase adaptability to changing and increasingly variable growing environments. However, this genetic diversity may also allow for the adaptation of crop populations to specific environments, provided selection pressure is strong and consistent enough [4–6]. One approach to enhancing within-crop diversity is evolutionary breeding [7], which can be achieved through the use
of heterogeneous crop populations or evolutionary populations (EPs), of which composite cross populations (CCPs) are an example.

Heterogeneous populations result from the crossing of a minimum of at least five parental genotypes [8,9], after which the seeds produced by the crossings are multiplied (if needed) and the harvested seeds are equally mixed together [2]. This population can be sown and harvested over a number of generations, either under natural or supervised selection. Depending on their genetic make-up and environmental conditions, CCPs may respond to climatic uncertainties, adapt to biotic and abiotic stresses, and maintain genetic diversity [10–13]. CCPs and other genotype mixtures result in greater intra- and inter-varietal diversity within the agricultural landscape and ensure a wider adaptation capacity for crop varieties [2]. As discussed by Ceccarelli and Grando [5], the use of EPs selected within and adapted to target environments has the advantage of buffering strong genotype-year interactions within location interactions, which can be hugely variable, particularly due to climate change. However, despite the advantages of decentralised selection for genotype-location interactions and the ability of EPs to adapt to these target environments over time, many EPs show great adaptation capacity to non-targeted environments and performance stability across environments [4,6,14,15].

Heterogeneous populations of barley and wheat are often comparable to or superior to modern pure line varieties for both yield and yield stability, mainly under organic and low-input conditions [10,15–21]. Exposing genetically similar wheat and barley populations to differing environments may result in population divergence through adaptation to specific climatic conditions, biotic stresses, or farming systems [22–28]. Overall, genetic diversity can improve crop resilience and performance and provide greater buffering capacity under challenging conditions [1,2,5,13,29,30].

According to Patel et al. [31], large environmental variances, such as those experienced by unpredictable weather conditions, will result in slow population evolutions as genotype diversity is preserved through the differing genotypes that thrive and reproduce under particular, but changing, environments. This, as mentioned by Döring et al. [2], provides “a buffer against the environmental fluctuations through compensatory effects”. In more stable environments, selection pressures may be stronger and more consistent, enabling quicker divergent population evolutions according to local selection pressures such as pathogen presence, interspecific competition, or nutrient availability. In this way, specifically suited genotypes will be preferred in a specialised environment, thereby reducing the genetic diversity of the population but favouring local adaptation [22–26]. The success of environmental adaptation, however, is also largely dependent on the suitability of parental pedigree [6]. A slow adaptation to an ever-changing environment is thus to be expected if populations only have one season to adapt to a given location before being moved to a different and contrasting environment.

As the effects of climate change on crop agronomic performance are difficult to quantify, particularly in short-term trials, a unique opportunity arose within a European network working together on a winter wheat (Triticum aestivum L.) CCP that was created in 2001 in the United Kingdom (UK). Sub-populations originating from this CCP were exposed to changing pedoclimatic conditions across Europe for five consecutive years (2008/09: F8–2012/13: F12) in eight sites, resulting in “cycling CCPs”. In contrast, “non-cycling CCPs” were maintained in parallel in the UK, Hungary, and Germany.

In summer 2013, the seeds of all eleven CCPs were harvested and evaluated (the F13) at one site (Germany) together in 2013/14 with three reference line varieties. As the seeds of the eleven CCPs had been grown at eight different sites in 2012/13, seed size was inevitably variable among the CCPs. Therefore, in the second experimental year (2014/15), not only stored seed of the F13 (harvested by partners in 2013) but also F14 seed that had been produced during the first experimental year were included in the experiment. There was a smaller seed size variation in the F14 seed, which allowed us to address the question of (i) how seed size affects the performance of the populations under ambient environmental conditions. Additionally, the performance (agronomic and baking
quality) and phenotypic diversity of the populations were assessed to (ii) test if there are agronomic and morphological changes among non-cycling and cycling CCPs, pointing to specific adaptations depending on management history. Finally, (iii) we asked if there were differences in yield performance and baking quality of the CCPs in comparison to pure lines.

2. Materials and Methods

2.1. Genetic Material

The CCPYQ is the result of crossing a total of 20 wheat varieties (release dates of CCP parental varieties from 1934 to 2000) conducted in 2001 in the UK. Eight high-yielding parents (Y) were crossed with eleven high-baking-quality parents (Q) (88 crosses). The variety “Bezostaya”, which is considered both high yielding and of high baking quality, was intercrossed with all other 19 parental varieties (107 crosses in total). Crosses of the 19 Y and Q parents with one of four naturally male sterile lines were also bulked into the CCP to enhance further outcrossing in later generations [29,32]. Since 2005, bulked progenies of the F₅ generation of this population have been maintained in the UK, Hungary (HU), and Germany (D) under natural selection only.

Since the F₅ generation (2008/09), a circulation scheme was developed by eight partners to submit the CCP (CCPYQ) that had been maintained in the UK, HU, and D since 2005 to changes in environment every year, creating cycling populations in order to test the effects of differing environments on populations and possible robustness through compensatory effects to environmental variability through exposure to changing conditions over time. The CCPs were grown under organic or low-input conditions in a plot > 100 m² at one site, harvested without selection, and sent to the next cycling partner the following year (Figure 1). The original partners (UK, HU, and D) also maintained their original non-cycling populations at the same location for comparison. Each partner involved in the cycling experiment multiplied and forwarded the received population on to the following partner from 2008/09 (F₅) until the F₁₃ generation (2012/13) (Figure 1).

Figure 1. Cycling scheme developed between the eight partners for the composite cross populations derived from the F₅ generation from Germany, the UK, and Hungary (see Table 1 for more details). (Google Maps, 2023). Numbers indicate partner sites: (1) The Organic Research Centre, United Kingdom; (2) University of Copenhagen and Agrologica, Denmark; (3) Technical University of Munich, Germany; (4) Centre for Agricultural Research, Hungary; (5) Wageningen University and Research, the Netherlands; (6) University of Kassel, Germany; (7) FiBL, Switzerland; (8) INRAE, Le Moulon, France.
Table 1. CCP code, origin, and cycling history of the winter wheat CCPs compared in 2014 and 2015 under organic growing conditions (Neu-Eichenberg, Germany).

<table>
<thead>
<tr>
<th>Origin</th>
<th>2008/09</th>
<th>2009/10</th>
<th>2010/11</th>
<th>2011/12</th>
<th>2012/13</th>
<th>CCP Code Used in This Study</th>
<th>Cycling (C)/Non-Cycling (NC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D(KS)</td>
<td>CH</td>
<td>F</td>
<td>DK</td>
<td>D(TUM)</td>
<td>HU</td>
<td>HU13NC</td>
<td>C</td>
</tr>
<tr>
<td>CH</td>
<td>F</td>
<td>UK</td>
<td>DK</td>
<td>D(TUM)</td>
<td>HU</td>
<td>HU13NC</td>
<td>C</td>
</tr>
<tr>
<td>NL</td>
<td>D(KS)</td>
<td>CH</td>
<td>F</td>
<td>UK</td>
<td>UK13NC</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>D(TUM)</td>
<td>HK</td>
<td>UK</td>
<td>UK</td>
<td>UK</td>
<td>UK13NC</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>UK</td>
<td>UK</td>
<td>UK</td>
<td>UK</td>
<td>UK</td>
<td>UK13NC</td>
<td>C</td>
<td></td>
</tr>
</tbody>
</table>

* Partner codes and institutes: CH = FiBL, Switzerland; DK = University of Copenhagen and Agrologica, Denmark; D(KS) = University of Kassel, Germany; F = INRAE, Le Moulon, France; HU = Centre for Agricultural Research, Hungary; NL = Wageningen University and Research, the Netherlands; D(TUM) = Technical University of Munich, Germany; UK = The Organic Research Centre, United Kingdom. ** F13 CCPs (harvested in 2013 by all partners) and F14 CCPs (harvested in 2014 from the first experimental trial year) are indicated in the text by F13 and F14, respectively.

This resulted in three non-cycling (i.e., home) populations, which all originated from the same batch of seed in 2005, as well as eight cycling populations, which originated from these three non-cycling populations in either 2008 or 2009 (Table 1). These 11 CCPs were compared at the University of Kassel, Germany, in a two-year trial in 2013/14 and 2014/15.

2.2. Field Site and Experimental Design

The trials were carried out in the two experimental seasons of 2013/14 and 2014/15 at the research fields of the Faculty of Organic Agricultural Sciences (University of Kassel, Witzenhausen) in Neu-Eichenberg (51°22′24.7″ N and 9°54′12.5″ E), mean annual precipitation (September to August 2005–2015): 674 mm, mean annual temperature (September to August 2005–2015): 9.5 °C, altitude: 247 m above sea level (see [21] for detailed weather data for the experimental site). The soil is classified as fine loamy loess soil (Haplic Luvisol), scoring 76 points according to the German soil point system (0–100). The fields have been managed organically since 1984 without the use of synthetic fertilisers or pesticides (see [15,18,21] for more details).

In 2013/14, seeds from the eight cycling CCPs, as well as seeds from the three non-cycling CCPs (F13) (Table 1), were received from all partners and sown in the same field at one site in Neu-Eichenberg, Germany, in a randomised complete block design with four replications. In addition, three pure line varieties, Achat, Akteur, and Capo, were included as references (14 experimental entries). All three references are classified as baking-quality varieties commonly grown by organic farmers in Germany. In 2014/15, stored seed of all CCPs from 2013 (F13) and harvested seed (F14) of all CCPs from 2014 in Neu-Eichenberg, Germany, were included with the same three reference varieties in a randomised complete block design with four replicates (25 experimental entries) at the same site in Neu-Eichenberg, Germany (Table 1). The pre-crop for both experiments was a two-year grass-clover mixture. Soil samples were taken at two depths (0–30 cm and 30–60 cm) per replicate block shortly after sowing and analysed according to the VDLUFA standards [33]. Total mineral nitrogen availability (kg N/ha) in the two soil levels was 49 kg N/ha in 2013 and 54 kg N/ha in 2014. No organic fertiliser was applied before or during the experimental seasons. The wheat was sown on 31 October 2013 and 29 October 2014 in plot sizes of 18 m² (11 × 1.5 m) at 350 germinable kernels/m². Each plot contained...
five rows, which were spaced 30 cm apart to allow for weed control through harrowing and/or hoeing at the tillering stage.

Both experimental seasons were characterised by dry winters and springs, with the exception that in May of 2013/14, enough rain fell, providing sufficient water during the critical period of flowering to prevent serious yield loss. In 2013/14, the dry months over winter and into early spring were coupled with higher-than-average temperatures, particularly between January and April, resulting in a mean temperature of 10.3 °C (September to August) and a total precipitation of 604 mm for the time relevant for wheat growth (September to July) (Figure 2). Monthly temperatures in 2014/15 were generally similar to the long-term average (1971–2000), between February and mid-June, with higher-than-average temperatures during the autumn and winter and from mid-June to August (Figure 2). However, these were coupled with uninterrupted dry conditions from November to June, leading to extremely early maturation. Thus, the 95 mm that fell in July had no effect on yield, and effectively, only 334 mm of precipitation (from September to June) were relevant for the wheat season (Figure 2). In comparison, the long-term total seasonal precipitation (September to July) was 573 mm (1971–2000).

Figure 2. Mean monthly temperatures (°C) and total monthly precipitation (mm) for the experimental years of 2013/14 and 2014/15. The long-term monthly means for Göttingen, Germany, (1971–2000) for both temperature and precipitation are also given.

2.3. Assessments

Foliar diseases caused by fungal pathogens were assessed on 10 plants per plot at BBCH stages 45–55 (26 May 2014 and 29 May 2015) and BBCH stages 60–70 (16 June 2014 and 17 June 2015). Non-green leaf area (NGLA) was recorded as a percentage of the flag leaf (F), the leaf below the flag leaf (F-1), and, if possible, the F-2 leaf (first assessment only). The two main foliar leaf diseases observed were also recorded at each assessment date. Due to the strong similarity of symptoms between Stagonospora nodorum and Zymoseptoria tritici, these two pathogens were not distinguished separately from one another in the assessments. The occurrence of foot diseases was low in 2013/14 and close to zero in 2014/15, and as such, they were not considered relevant to the agronomic results.
Morphological assessments were performed on three 0.5-m rows of mature plants cut as close to the ground as possible from each plot (0.450 m² per plot). The number of ears for ear-bearing tillers per m² and the percentage of counted ears with awns (awned ears) were recorded. From these samples, total biomass and kernel weight were also recorded to calculate the harvest index (HI) and kernel number per ear. Combine harvested grain yields (t/ha) and thousand kernel weights (TKWs) (g) were adjusted to 14% moisture content. The TKW (g) was measured three times per plot (the weight of 1000 kernels), and the mean of these three measurements was used for analysis. Stem length (measured from base of stem to start of ear) and ear length (awn length not included) were recorded for 45 randomly chosen mature plants per plot (n = 180 plants in total per entry).

Protein content (%) was determined for all CCP entries and reference varieties in both generations and in both years. Protein content (%) was calculated according to seed nitrogen content (% N × 5.7), which was analysed using an Elementar Analysator vario MAX (Elementar Analysetechnik GmbH, Hanau, DE). In addition, a number of baking quality parameters were assessed for all F₁₄ CCPs and reference varieties in 2014/15. Hagberg Falling Number (HFN), sedimentation value (Zeleny), wet gluten content (%), and water absorption capacity (%) (ICC Methods no. 107, 116, 106/2, and 179, respectively) were analysed by Aberham Laboratories, Großaitingen, DE. The analysis of baking volume (mL/100 g flour) was based on an internal test developed by Aberham Laboratories using wholemeal flour and the addition of malt flour in order to stabilise the loaf crust only when HFN values were too high (personal communication, Dr. Aberham).

2.4. Data Processing and Statistics

The mean non-green leaf area (NGLA%) over all three leaf levels was calculated by weighting the F leaf by the factor of 4, and the F-1 and F-2 leaves by the factor of 3 (as described by [18]). The higher weighting of the flag leaf takes into account its greater contribution to total grain yield in comparison to lower leaves [34]. The Area under the Disease Progress Curve (AUDPC) [35], as well as the Relative Area under the Disease Progress Curve (RAUDPC) in accordance with Fry [36], were calculated to allow for comparison of the populations and reference varieties within an experimental year or over the two experimental years, respectively.

Data analysis was conducted using the statistical software R (Version 3.4.4) [37], and graphics were produced using Excel (Microsoft Office, Version 18.2305.1222.0) and the R package “ggplot2” [38]. Linear models were built using the R package “stats” [37] and were performed on a number of datasets described as follows: Dataset 1a: all experimental entries (F₁₃ CCPs and reference varieties) over both experimental years (2013/14 and 2014/15); dataset 1b: all experimental entries (F₁₄ CCPs and reference varieties) in 2014/15 only; dataset 2: the F₁₃ generation CCPs (over both experimental years, not including reference varieties); and dataset 3: the F₁₃ and F₁₄ CCPs in 2014/15 only (not including reference varieties). Linear models applicable for a randomised complete block design (RCBD) were used, and model residuals were tested to ensure all assumptions were met. Estimated marginal means were calculated using the “emmeans” package [39], followed by a Tukey post hoc test with pairwise comparison and Holm correction.

For comparison of the agronomic performance of the F₁₃ CCPs and the reference varieties in 2013/14 and 2014/15 (dataset 1a), as well as the F₁₄ CCPs and reference varieties in 2014/15 (dataset 1b), the following model was used (ANOVA type III):

\[
\text{Response}_{ijr} = \mu + E_i + Y_j + (E_i \times Y_j) + B_r(Y_j) + e_{ijr}
\]

where Response is the trait of interest, \(\mu\) is the effect of the mean, \(E\) is the effect of the experimental entry (CCPs with differing management histories and reference varieties), \(Y\) is the year effect (2013/14 and 2014/15), \(B\) is the replicate block effect, and \(e\) is the residual errors. The same model without the factor year was used to analyse differences between the F₁₄ CCPs and reference varieties for agronomic and morphological traits, as well as baking quality, in 2014/15 only. To test differences between the CCP entries only, the
The abovementioned model was used for datasets 2 and 3, with the factor year replaced by generation for the comparison of F$_{13}$ and F$_{14}$ CCP entries in 2014/15.

The model used to quantify differences in the CCPs of the F$_{13}$ generation based on their origin and cycling history (dataset 2) was the following ANOVA type III:

$$\text{Response}_{ijk\ell} = \mu + O_i + C_j + Y_k + (O_i \times C_j) + (O_i \times Y_k) + (C_j \times Y_k) + (O_i \times C_j \times Y_k) + B_k(Y_k) + e_{ijk\ell}$$

where Response is the trait of interest, $\mu$ is the effect of the mean, $O$ is the effect of the origin (UK, D, and HU), $C$ is the effect of cycling (cycling/non-cycling), $Y$ is the year effect (2013/14 and 2014/15), $B$ is the replicate block effect, and $e$ is the residual errors.

To detect differences between the CCPs based on the effects of origin, cycling history, and generation, the following model (ANOVA Type III) was used on dataset 3:

$$\text{Response}_{ijk\ell} = \mu + O_i + C_j + G_k + (O_i \times C_j) + (O_i \times G_k) + (C_j \times G_k) + (O_i \times C_j \times G_k) + B_k + e_{ijk\ell}$$

where Response is the trait of interest, $\mu$ is the effect of the mean, $O$ is the effect of the origin (UK, D, and HU), $C$ is the effect of cycling (cycling/non-cycling), $G$ is the effect of generation (F$_{13}$/F$_{14}$), $B$ is the replicate block effect, and $e$ is the residual errors.

In order to further explore the differences between the F$_{13}$ CCPs in both experimental years (dataset 2), GGE (genotype plus genotype-by-environment) biplots, namely the “which-won-where view” and “mean vs. stability” biplots, were implemented using the “metan” package in R [40]. In the “which-won-where view” biplot, the blue dotted lines indicate differential sectors based on the measured genotype traits. If an environment falls within a sector, then the genotype on the perimeter of the polygon indicates the genotype best suited to a particular environment for the selected trait [41]. The “mean vs. stability” biplot is used as an indication as to which genotypes rank where in terms of the measured trait. The AEC (average environment coordination) axis is the second axis that runs through the biplot origin and is perpendicular to the AEA. The greater the distance away from the origin of the biplot along this line, the greater the instability of the genotype in terms of genotype-environment interaction (GEI) [41].

Separate genotype by trait (GT) biplots were also created for the F$_{13}$ CCPs in 2013/14, the F$_{13}$ CCPs in 2014/15, and the F$_{14}$ CCPs in 2014/15. The analysis of GT biplots has two major objectives: the first is to understand the relations among traits, and the second is to understand the trait profiles of the genotypes. In GT biplots, the cosine of the angle between two traits approximates the correlation between them. Additionally, the length of the vectors indicates the closeness of the association with other traits. Traits with shorter vectors tend to have weaker associations with other traits. The same is true for genotypes and traits in terms of the cosine of the angle between the genotypes and the traits in the interpretation of trait profiles. Genotypes found to be closely located (<90°) to specific traits indicate a high propensity for the trait [42].

Correlation coefficients were calculated between agronomic parameters within each generation using the “GGally” package in R [43], as well as between TKW at sowing and agronomic parameters for each generation. For the calculation of the correlation coefficients, the data was tested for normality, and either a Pearson product-moment correlation or a Spearman’s rank correlation was performed. Grain protein deviation (GPD), as a measurement to identify genotypes with greater grain protein than would be expected from their grain yield [44], was calculated as the standardised residuals derived from the regression of grain protein content on grain yield. GPD was calculated by averaging GPD values over replicate blocks for each experimental entry (references and CCPs) and each year, as described by Bogard et al. [45].
3. Results

The main foliar disease observed in both experimental years was stripe rust caused by *Puccinia striiformis* f. sp. *tritici*. The diseases *Stagonospora nodorum/Zymoseptoria tritici* and *Drechslera tritici-repentis* were also present, although less frequently than stripe rust. Foliar disease pressure was lower in 2013/14, with NGLA values in 2013/14 ranging from 2 to 13% (1st assessment) and from 7 to 14% (2nd assessment). In 2014/15, NGLA values at the 1st assessment ranged from 4 to 11% and from 10 to 33% at the 2nd assessment.

3.1. Seed Size Effects

As seeds of the F$_{13}$ CCPs were sent from various partners, the TKW values of the seeds at sowing differed for each population. The TKW values ranged from 42 g (NL13 CCP) to 47 g (UK13 CCP) (standard deviation (SD) of TKW at sowing F$_{13}$: 1.42). In contrast, the range of TKW for the F$_{14}$ (i.e., derived from the harvested seed of the F$_{13}$ in 2014) was smaller, ranging from 45 g to 47 g (SD of TKW at sowing F$_{14}$: 1.38). In 2013/14, significant positive correlations were present between TKW at sowing and AUDPC and TKW values at harvest (both $p < 0.05$, respectively). In the drought year of 2014/15, however, TKW at sowing of the F$_{13}$ correlated positively with stem length and negatively with HI (both $p < 0.01$, respectively), but no correlations were observed for the F$_{14}$ (Table 2).

<table>
<thead>
<tr>
<th>Gen.</th>
<th>Yield</th>
<th>HI</th>
<th>AUDPC</th>
<th>TKW</th>
<th>Ears.m$^{-2}$</th>
<th>Kernel no.</th>
<th>Protein</th>
<th>Stem length</th>
</tr>
</thead>
<tbody>
<tr>
<td>F$_{13}$</td>
<td>0.08</td>
<td>-0.19</td>
<td><strong>0.32</strong></td>
<td><em>0.36</em></td>
<td>-0.01</td>
<td>-0.27</td>
<td>0.12</td>
<td>0.22</td>
</tr>
<tr>
<td>F$_{14}$</td>
<td>-0.01</td>
<td>-0.43 **</td>
<td>0.02</td>
<td>0.14</td>
<td>0.01</td>
<td>-0.25</td>
<td>-0.14</td>
<td><strong>0.41</strong></td>
</tr>
</tbody>
</table>

Table 2. Correlation coefficients (r) for various agronomic parameters and TKW at sowing for winter wheat CCP entries of the F$_{13}$ in 2013/14 and F$_{13}$ and F$_{14}$ in 2014/15 grown in Neu-Eichenberg, Germany. Significant correlations are marked in bold.

The interrelations among agronomic parameters of the CCPs were affected both by the strong seasonal differences as well as by seed size variation at sowing due to their different origins. Despite similar mean AUDPC values in both the F$_{13}$ and F$_{14}$ CCPs in 2014/15, negative correlations of AUDPC with TKW, kernel number per ear, and HI tended to be stronger in the F$_{13}$, most likely due to the slightly higher AUDPC values found for certain CCPs in the F$_{13}$ in comparison to the F$_{14}$ of 2014/15 (Figure S1, Table S2). The F$_{13}$ CCPs grown in the wetter season of 2013/14 displayed a few differential interrelations in comparison to the 2014/15 season (Figure 3A,B). In 2014/15, when both generations were grown in the same experimental year, differences in the direction and/or significance of the correlation coefficients were found depending on the generation (Figure S1). These apparent differential yield component profiles between generations are compounded by the legacy effect of seed size variation at sowing, resulting in a significant CCP generation effect on TKW (Table S1), in addition to the significant negative correlation with HI and positive correlation to stem length with seed size at sowing for the F$_{13}$ but not the F$_{14}$ in 2014/15 (Table 2).
Figure 3. The genotype by trait (GT) biplot for the parameters yield (Yield), protein yield (Pyield), TKW, straw yield (Strawyield), ear-bearing tillers per m² (Ears_m2), percentage of ears with awns (Awned_ears), protein content (Protein), RAUDPC, stem and ear length (Stem_length, Ear_length), kernel number per ear (Kernel_no), and harvest index (HI) for winter wheat CCPS for (A) the F13 in 2013/14, (B) the F13 in 2014/15, and (C) the F14 in 2014/15 (Neu-Eichenberg, Germany).

3.2. Comparison of the Cycling and Non-Cycling CCPs and Their Origins

Total yields were significantly greater in 2013/14 (mean yield of 4.7 t/ha) in comparison to both the F13 and F14 of the drought year 2014/15 (mean yield of 4.2 t/ha) (Table 3,
Tables S2 and S4). With few exceptions, significant differences between cycling and non-cycling populations were not consistent in each generation and experimental year and were intermingled with legacy effects due to the seed production environment. Thus, the strongest effect of cycling history on CCPs of different origins was found in the percentage of ears with awns. These differences were consistent for each generation and year, with the HU non-cycling CCP indicating a significantly greater percentage of ears with awns in comparison to the HU cycling CCPs and the UK non-cycling CCP tending towards significantly fewer awned ears in comparison to the UK cycling CCPs (Table 3). Significant differences between non-cycling CCPs for percentage of awned ears were consistent for the F13 generation in both years, with a higher degree of awned ears found in the D and HU CCPs, and this same pattern was found in the F14 non-cycling CCPs, although it was no longer significant.

Table 3. Yield (t/ha), HI, RAUDPC, TKW (g), ear-bearing tillers/m² (Ears_m²), kernel number per ear (Kernel no.), protein content (%), percentage of ears with awns (Awned ears (%)), stem length (cm), and ear length (cm) of the winter wheat CCP entries grouped by origin and cycling history in the experimental years 2013/14 (F13) and 2014/15 (F13 and F14). Differing small letters indicate significant differences between CCP origin groups based on their cycling history and within each generation and year. Differing capital letters indicate significant differences between cycling and non-cycling CCPs for each generation and year. Asterisks indicate significant differences between cycling and non-cycling CCP groups of the same origin within a given generation and year.

<table>
<thead>
<tr>
<th>Origin</th>
<th>Yield (t/ha)</th>
<th>HI</th>
<th>RAUDPC</th>
<th>TKW (g)</th>
<th>Ears_m²</th>
<th>Kernel no.</th>
<th>Protein (%)</th>
<th>Awned ears (%)</th>
<th>Stem length (cm)</th>
<th>Ear length (cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F13</td>
<td>F14</td>
<td>F13</td>
<td>F14</td>
<td>F13</td>
<td>F14</td>
<td>F13</td>
<td>F14</td>
<td>F13</td>
<td>F14</td>
</tr>
<tr>
<td></td>
<td>NC</td>
<td>C</td>
<td>NC</td>
<td>C</td>
<td>NC</td>
<td>C</td>
<td>NC</td>
<td>C</td>
<td>NC</td>
<td>C</td>
</tr>
<tr>
<td>D</td>
<td>4.92 ab</td>
<td>4.82</td>
<td>4.09</td>
<td>4.31 b</td>
<td>4.09</td>
<td>4.37 b</td>
<td>0.39</td>
<td>0.40</td>
<td>0.38</td>
<td>0.39</td>
</tr>
<tr>
<td>HU</td>
<td>4.34 a b</td>
<td>4.76</td>
<td>4.24</td>
<td>4.33 b</td>
<td>4.19</td>
<td>4.11 b</td>
<td>0.40</td>
<td>0.40</td>
<td>0.40</td>
<td>0.40</td>
</tr>
<tr>
<td>UK</td>
<td>5.04 b*</td>
<td>4.56</td>
<td>4.34</td>
<td>4.00 a</td>
<td>4.22</td>
<td>4.27 b</td>
<td>0.40</td>
<td>0.40</td>
<td>0.41</td>
<td>0.41</td>
</tr>
<tr>
<td>Mean</td>
<td>4.77</td>
<td>4.71</td>
<td>4.22</td>
<td>4.22 b</td>
<td>4.17</td>
<td>4.25 b</td>
<td>0.40</td>
<td>0.40</td>
<td>0.39</td>
<td>0.39</td>
</tr>
<tr>
<td>RAUDPC</td>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.06</td>
<td>0.05</td>
<td>0.14 b*</td>
<td>0.11 a</td>
<td>0.10</td>
<td>0.11 a</td>
<td>45.9 b*</td>
<td>44.1</td>
<td>45.7</td>
<td>45.2 b</td>
</tr>
<tr>
<td>HU</td>
<td>0.07</td>
<td>0.07</td>
<td>0.17 b*</td>
<td>0.12 a</td>
<td>0.15 b</td>
<td>0.13 a</td>
<td>41.7 a*</td>
<td>43.3</td>
<td>43.3 a*</td>
<td>44.8 b</td>
</tr>
<tr>
<td>UK</td>
<td>0.04</td>
<td>0.05</td>
<td>0.10 a</td>
<td>0.12 b</td>
<td>0.12 ab</td>
<td>0.12 ab</td>
<td>42.7 a</td>
<td>43.5</td>
<td>44.1 ab</td>
<td>43.4 a</td>
</tr>
<tr>
<td>Mean</td>
<td>0.06</td>
<td>0.06</td>
<td>0.13 B</td>
<td>0.12 A</td>
<td>0.12</td>
<td>0.12 A</td>
<td>43.4</td>
<td>43.9</td>
<td>44.3</td>
<td>44.4</td>
</tr>
<tr>
<td>Protein (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>9.7 ab</td>
<td>29</td>
<td>9.7</td>
<td>30</td>
<td>9.7</td>
<td>27</td>
<td>11 b</td>
<td>7</td>
<td>10</td>
<td>7 ab</td>
</tr>
<tr>
<td>HU</td>
<td>10.1 b</td>
<td>10</td>
<td>9.8</td>
<td>14</td>
<td>9.6</td>
<td>14 b*</td>
<td>7</td>
<td>11</td>
<td>10</td>
<td>7 a</td>
</tr>
<tr>
<td>UK</td>
<td>9.4 a</td>
<td>9.7</td>
<td>27</td>
<td>3 a*</td>
<td>10</td>
<td>3 a*</td>
<td>5</td>
<td>10</td>
<td>5</td>
<td>10 b</td>
</tr>
<tr>
<td>Mean</td>
<td>9.7</td>
<td>9.7</td>
<td>9.8</td>
<td>9.7</td>
<td>9.7</td>
<td>9</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>Stems (cm)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>92.2</td>
<td>89.9</td>
<td>88.5</td>
<td>90.6</td>
<td>8.56</td>
<td>8.39</td>
<td>7.68</td>
<td>7.42</td>
<td>7.08</td>
<td>7.67</td>
</tr>
<tr>
<td>HU</td>
<td>94.5</td>
<td>92.7</td>
<td>87.9</td>
<td>87.5</td>
<td>8.44</td>
<td>8.36</td>
<td>7.51</td>
<td>7.53</td>
<td>7.61</td>
<td>7.61</td>
</tr>
<tr>
<td>UK</td>
<td>94.1</td>
<td>95.3</td>
<td>88.8</td>
<td>88.3</td>
<td>8.30</td>
<td>8.47</td>
<td>7.59</td>
<td>7.62</td>
<td>7.47</td>
<td>7.56</td>
</tr>
<tr>
<td>Mean</td>
<td>93.6</td>
<td>94.0</td>
<td>87.8</td>
<td>88.0</td>
<td>8.43</td>
<td>8.41</td>
<td>7.59</td>
<td>7.52</td>
<td>7.39</td>
<td>7.61</td>
</tr>
</tbody>
</table>

An overall significant difference was found between the cycling and non-cycling F13 CCPs in 2014/15 for RAUDPC, mainly due to the significant differences between cycling and non-cycling CCPs of both D and HU origin in the experimental year with higher
foliar disease pressure, namely stripe rust (Table 3). A significant overall effect of CCP origin was also found for TKW for both the F13 generation over both years (Table S3) and for the F13 and F14 generation in 2014/15 (Table S1). However, despite this overall effect of origin, significant differences between CCPs of differing origin were only consistent for the non-cycling CCPs of the F13 generation and for the cycling CCPs in the 2014/15 experimental year (Table 3).

No general yield advantage was observed in the D13NC CCP (multiplied since the F5 generation in Germany); however, TKW was found to be higher in this CCP in both F13 generations (not in the F14) in comparison to the cycling CCPs, most likely due to seed size variation. This legacy effect was also evident for differences in TKW between cycling and non-cycling CCPs of HU origin, which were only consistent in the F13 generation, despite the strong overall effect of the experimental year for the majority of the measured traits in the comparison of the F13 generation (Tables S3 and S4). The overall effect of generation in 2014/15 was also only significant for TKW, with the F14 CCPs indicating a greater TKW (44.9 g) in comparison to the F13 CCPs (44.4 g) (Tables S1 and S2). Significant differences between cycling and non-cycling CCPs were found for both ear length and ear number per m² in the F14 CCPs grown from seeds with a similar TKW (removal of potential legacy effect). In contrast, this cycling history effect was not found in the F13 generation in either experimental year, again an indication of seed size legacy (Table 3).

3.3. Comparison of the CCP Entries

In terms of yield performance of the F13 CCPs, the DK13 and UK13NC CCPs tended towards the highest yields in both experimental years, followed by the UK13, TUM13, and NL13 CCPs (Figure 4A, Table 4 and Table S4). In terms of yield stability based on GEI, the CCPs DK13, UK13, and TUM13 indicated the greatest stability at a high yield level, with D13 indicating yield stability albeit at the lowest yield level (Figure 4B) (note that in this panel the lowest yields are to the right, see Table 4). Both the D13 and HU13NC CCPs indicated the highest grain protein content in both experimental years, followed by the TUM13 CCP (Figure 4C). The CCPs D13NC and NL13, despite their moderate grain protein content, indicated greater grain protein stability in comparison to the other CCP entries, while the CCP TUM13 indicated stability at a higher grain protein level (Figure 4D).

Significant year effects for the F13 CCPs were found for yield, RAUDPC, TKW, kernel number per ear, and stem and ear length (Tables S3 and S4), and significant CCP entry effects were found for yield, HI, RAUDPC, TKW, kernel no., protein, percentage of awned ears, and stem and ear length (Table S4). However, a significant CCP and year interaction effect was only found for kernel number per ear (F10 = 2.92, p ≤ 0.01), indicating that apart from this trait, the CCP entries maintained a similar ranking for the measured traits irrespective of the experimental year (Table S4). A significant generation effect between the F13 and F14 CCPs was only found for TKW (Tables S1 and S2). Significant differences between CCP entries occurred for both the F13 and F14 in the 2014/15 experimental year for HI, kernel no., percentage of awned ears, and stem length. In contrast, entry effects on TKW and RAUDPC were only significant for the F13 in that year (Table S4). The interaction effect of CCP entry and generation was only significant for TKW (F10 = 2.66, p ≤ 0.01) and ears per m² (F10 = 2.21, p ≤ 0.05), indicating that for all other traits, the ranking of the CCP entries within each generation was similar (Table S2).

In contrast to the yield components, where CCP entry ranking was variable depending on the season and generation, CCP origin affected stripe rust severity. In 2015, when disease pressure due to stripe rust was higher, the RAUDPC of the F13 HU13NC (0.17) was significantly higher in comparison to the DK13 and UK13NC CCPs (0.09). This pattern was similar for the F14; however, significant differences between the CCPs were no longer found (Table S2, Figure 3A–C).
Figure 4. The “which-won-where view” biplot for grain yield (A) and protein content (C), as well as the “mean vs. stability” biplot for grain yield (B) and protein content (D) of the F13 winter wheat CCPs in experimental years 2013/14 and 2014/15 (Neu-Eichenberg, Germany). Note that in Panel B, the lowest yields are to the right, and in Panel D, the highest protein content is to the right (see Table 4 for the values for each CCP).
**Table 4.** Yield (t/ha), HI, AUDPC, TKW (g), ear-bearing tillers/m² (Ears_m²), and kernel number per ear (Kernel no.) of the reference varieties and winter wheat CCP entries of the F₁₃ in 2013/14 and F₁₃ and F₁₄ in 2014/15 (Neu-Eichenberg, Germany). Differing small letters indicate significant differences between reference varieties and CCP entries within each year and generation. Differing capital letters indicate significant differences between the means of the reference varieties and CCP entries within each year and generation.

<table>
<thead>
<tr>
<th>Origin</th>
<th>Entry</th>
<th>Yield (t/ha) F₁₃</th>
<th>F₁₄</th>
<th>HI F₁₃</th>
<th>F₁₄</th>
<th>AUDPC F₁₃</th>
<th>F₁₄</th>
<th>TKW (g) F₁₃</th>
<th>F₁₄</th>
<th>Ears_m² F₁₃</th>
<th>F₁₄</th>
<th>Kernel no. per ear F₁₃</th>
<th>F₁₄</th>
</tr>
</thead>
<tbody>
<tr>
<td>References</td>
<td>Achat</td>
<td>4.9 bc</td>
<td>4.3</td>
<td>0.43 c</td>
<td>0.41</td>
<td>0.41 bc</td>
<td>0.41</td>
<td>48.1 e</td>
<td>48.0 d</td>
<td>389 ab</td>
<td>409 ab</td>
<td>401 ab</td>
<td>31 ab</td>
</tr>
<tr>
<td></td>
<td>Akteur</td>
<td>3.7 a</td>
<td>3.8</td>
<td>0.40 abc</td>
<td>0.38</td>
<td>abc</td>
<td>0.38</td>
<td>46.6 de</td>
<td>46.0 a</td>
<td>324 a</td>
<td>336 a</td>
<td>336 a</td>
<td>27 a</td>
</tr>
<tr>
<td></td>
<td>Capo</td>
<td>5.6 c</td>
<td>4.2</td>
<td>0.41 abc</td>
<td>0.37 a</td>
<td>abc</td>
<td>0.37 ab</td>
<td>46.2 de</td>
<td>46.0 cd</td>
<td>393 ab</td>
<td>378 ab</td>
<td>379 ab</td>
<td>30 ab</td>
</tr>
<tr>
<td>D</td>
<td>D13NC</td>
<td>4.9 bc</td>
<td>4.1</td>
<td>0.39 a</td>
<td>0.38 abc</td>
<td>0.37 a</td>
<td>0.40 abc</td>
<td>0.40</td>
<td>45.9 cde</td>
<td>45.7 bcd</td>
<td>44.2 ab</td>
<td>367 ab</td>
<td>462 b</td>
</tr>
<tr>
<td></td>
<td>DK13</td>
<td>5.0 bc</td>
<td>4.5</td>
<td>0.41 abc</td>
<td>0.40 abc</td>
<td>0.39 abc</td>
<td>0.38</td>
<td>44.9 bcd</td>
<td>45.0 abc</td>
<td>45.9 bc</td>
<td>442 b</td>
<td>428 b</td>
<td>388 ab</td>
</tr>
<tr>
<td></td>
<td>HU13</td>
<td>4.5 ab</td>
<td>4.4</td>
<td>0.40 abc</td>
<td>0.38 abc</td>
<td>0.39 abc</td>
<td>0.38</td>
<td>43.3 abc</td>
<td>44.4 abc</td>
<td>45.6 b</td>
<td>412 ab</td>
<td>427 ab</td>
<td>29 ab</td>
</tr>
<tr>
<td></td>
<td>TUM13</td>
<td>4.9 bc</td>
<td>4.5</td>
<td>0.39 a</td>
<td>0.38 abc</td>
<td>0.39 abc</td>
<td>0.38</td>
<td>44.1 abcd</td>
<td>46.1 cd</td>
<td>45.4 b</td>
<td>439 b</td>
<td>441 b</td>
<td>29 ab</td>
</tr>
<tr>
<td>HU</td>
<td>HU13NC</td>
<td>4.3 ab</td>
<td>4.2</td>
<td>0.40 abc</td>
<td>0.40 abc</td>
<td>0.40 abc</td>
<td>0.39</td>
<td>41.7 a</td>
<td>43.3 ab</td>
<td>44.8 ab</td>
<td>389 ab</td>
<td>417 ab</td>
<td>418 ab</td>
</tr>
<tr>
<td></td>
<td>F13</td>
<td>4.6 b</td>
<td>4.3</td>
<td>0.40 abc</td>
<td>0.41 bc</td>
<td>0.42</td>
<td>0.40</td>
<td>43.3 ab</td>
<td>45.0 abc</td>
<td>44.7 ab</td>
<td>379 ab</td>
<td>434 b</td>
<td>374 ab</td>
</tr>
<tr>
<td></td>
<td>UK13</td>
<td>4.9 bc</td>
<td>4.4</td>
<td>0.39 ab</td>
<td>0.39 abc</td>
<td>0.39 abc</td>
<td>0.39</td>
<td>44.1 abcd</td>
<td>46.1 cd</td>
<td>45.4 b</td>
<td>439 b</td>
<td>441 b</td>
<td>29 ab</td>
</tr>
<tr>
<td>UK</td>
<td>UK13NC</td>
<td>5.0 bc</td>
<td>4.3</td>
<td>0.40 abc</td>
<td>0.41 cd</td>
<td>0.41 bc</td>
<td>0.39</td>
<td>45.0 bcd</td>
<td>45.7 abc</td>
<td>45.9 bc</td>
<td>442 b</td>
<td>438 b</td>
<td>388 ab</td>
</tr>
<tr>
<td></td>
<td>CH13</td>
<td>4.8 bc</td>
<td>3.9</td>
<td>0.39 ab</td>
<td>0.39 abc</td>
<td>0.39 abc</td>
<td>0.39</td>
<td>42.7 ab</td>
<td>44.1 abc</td>
<td>44.2 ab</td>
<td>408 b</td>
<td>423 b</td>
<td>441 b</td>
</tr>
<tr>
<td></td>
<td>D13</td>
<td>4.2 ab</td>
<td>3.7</td>
<td>0.38 a</td>
<td>0.39 abc</td>
<td>0.39 abc</td>
<td>0.38</td>
<td>43.3 abc</td>
<td>44.4 abc</td>
<td>45.6 b</td>
<td>412 ab</td>
<td>427 ab</td>
<td>29 ab</td>
</tr>
<tr>
<td></td>
<td>NL13</td>
<td>4.7 bc</td>
<td>4.4</td>
<td>0.42 bc</td>
<td>0.42 d</td>
<td>0.40</td>
<td>0.42</td>
<td>44.1 abcd</td>
<td>46.1 cd</td>
<td>45.4 b</td>
<td>439 b</td>
<td>441 b</td>
<td>29 ab</td>
</tr>
<tr>
<td>Reference mean</td>
<td></td>
<td>4.7</td>
<td>4.1</td>
<td>0.41 B</td>
<td>0.39</td>
<td>0.39</td>
<td>105</td>
<td>241</td>
<td>241</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCP mean</td>
<td></td>
<td>4.7</td>
<td>4.2</td>
<td>0.39 A</td>
<td>0.39</td>
<td>0.39</td>
<td>71</td>
<td>230</td>
<td>224</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.4. Comparison of the CCP Entries with the Reference Varieties

In 2013/14, the highest and lowest yields were achieved by reference varieties Capo (5.6 t/ha) and Akteur (3.7 t/ha), respectively (Table 4). The CCP entries in 2013/14 yielded similarly to one another, ranging from 4.2 to 5.0 t/ha. In 2014/15, the yields of the references ranged between 3.8 and 4.3 t/ha. The F₁₄ CCPs (4.1–4.4 t/ha) had a narrower yield range than the F₁₃ CCPs (3.7–4.4 t/ha) (Table 4). In the F₁₃ of 2014/15, all CCPs except the HU13NC CCP were comparable in their AUDPC values to the two reference varieties, Achat and Capo. Overall, the mean AUDPC values of the CCPs were generally lower than the mean AUDPC values of the reference varieties (Table 4).

Across seasons, the TKWs of the reference varieties and CCPs were similar. In the F₁₃ of 2013/14, TKWs ranged from 48.1 g (Achat) to 41.7 g (HU13NC). The overall mean...
TKW of the reference varieties was significantly higher than that of the F13 CCPs in both years but not compared to the F14 CCPs (Table 4). While the numbers of ear-bearing tillers per m² of the CCP entries in any generation or experimental year were similar in both years and generations (414 to 418 per m²), a significant CCP and generation interaction effect in 2014/15 (F₁₀ = 2.21, p ≤ 0.05) indicated a change in ranking of the CCP entries depending on generation (Table 4). However, all CCPs had a significantly higher number of ear-bearing tillers per m² than the reference variety with the lowest number of ears, Akteur, independent of generation and experimental year, and the overall CCP mean for this parameter was always significantly higher than the reference variety mean (Table 4).

Grain protein deviation (GPD) was calculated for each experimental entry in each environment (year and generation). GPD values above 0 indicate experimental entries with greater grain protein than would be expected from their grain yield, with a number of experimental entries achieving positive GPD values in both experimental years (Figure S2, Table S5). The reference variety Akteur indicated the highest GPD value consistently in both experimental years, followed by the D13 and TUM13 CCPs. In the F13 generation, the HU13NC indicated GPD values above 0, however, this was not the case in the F14 of 2014/15 (Figure S2), with the reference variety Achat also tending towards variable GPD values depending on the experimental year. The TUM13 CCP, in addition to constant positive GPD values and grain protein stability (Figure 4C,D), indicated a high yield potential and low GEI (Figure 4A,B, Table 4, Figure S2). The CCPs DK13, HU13, UK13NC, and F13 were found to have the lowest GPD values in both experimental years, as did the reference variety Capo (Figure S2, Table S5).

The mean protein content of the F14 CCPs and reference varieties was low (9.8%), reflecting the very dry spring conditions and low soil nitrogen availability (Table 5). The reference variety Akteur achieved the highest protein content (10.7%), followed by the D13 CCP and the reference variety Achat (10.0%).

Table 5. Baking quality parameters for the winter wheat F14 entries (2014/15) grown in Neu-Eichenberg, Germany, including water absorption (%), wet gluten content (%), sedimentation value (Zeleny, mL), Hagberg Falling Number (seconds), and baking volume (mL/100 g flour). Differing letters indicate significant differences between the entries (p < 0.05). Colour codes: red = poor/low rating, orange = moderate, and green = good/high rating *.

<table>
<thead>
<tr>
<th>Origin</th>
<th>Entry</th>
<th>Protein (%)</th>
<th>Water (%)</th>
<th>Gluten (%)</th>
<th>Sedi. (Zeleny)</th>
<th>HFN (sec.)</th>
<th>Baking vol. (mL/100g)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ref</td>
<td>Achat</td>
<td>10.0 ab</td>
<td>60.5</td>
<td>21.9 abc</td>
<td>31.8 cd</td>
<td>370 d</td>
<td>306.0</td>
</tr>
<tr>
<td></td>
<td>Akteur</td>
<td>10.7 b</td>
<td>62.3</td>
<td>24.3 c</td>
<td>35.8 d</td>
<td>373 d</td>
<td>318.4</td>
</tr>
<tr>
<td></td>
<td>Capo</td>
<td>9.7 a</td>
<td>62.4</td>
<td>21.4 abc</td>
<td>31.0 cd</td>
<td>313 c</td>
<td>278.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>DT13NC</td>
<td>9.8 ab</td>
<td>60.5</td>
<td>21.7 abc</td>
<td>25.8 ab</td>
<td>277 abc</td>
<td>303.0</td>
</tr>
<tr>
<td></td>
<td>DK13</td>
<td>9.5 a</td>
<td>59.7</td>
<td>19.8 ab</td>
<td>24.3 ab</td>
<td>304 bc</td>
<td>297.1</td>
</tr>
<tr>
<td></td>
<td>HU13</td>
<td>9.6 a</td>
<td>60.0</td>
<td>20.2 ab</td>
<td>22.8 a</td>
<td>260 ab</td>
<td>300.0</td>
</tr>
<tr>
<td></td>
<td>TUM13</td>
<td>9.9 ab</td>
<td>59.8</td>
<td>21.3 abc</td>
<td>22.0 a</td>
<td>281 ab</td>
<td>281.0</td>
</tr>
<tr>
<td>HU</td>
<td>HU13NC</td>
<td>9.7 a</td>
<td>59.9</td>
<td>20.5 ab</td>
<td>25.8 ab</td>
<td>261 ab</td>
<td>295.0</td>
</tr>
<tr>
<td></td>
<td>F13</td>
<td>9.3 a</td>
<td>59.3</td>
<td>18.6 a</td>
<td>21.5 a</td>
<td>249 a</td>
<td>279.0</td>
</tr>
<tr>
<td></td>
<td>UK13</td>
<td>9.8 ab</td>
<td>60.5</td>
<td>20.9 abc</td>
<td>25.0 ab</td>
<td>271 abc</td>
<td>301.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UK</td>
<td>UK13NC</td>
<td>9.4 a</td>
<td>60.0</td>
<td>20.0 ab</td>
<td>21.5 a</td>
<td>283 abc</td>
<td>294.4</td>
</tr>
<tr>
<td></td>
<td>CH13</td>
<td>9.8 ab</td>
<td>62.0</td>
<td>20.8 abc</td>
<td>25.3 ab</td>
<td>287 abc</td>
<td>307.0</td>
</tr>
<tr>
<td></td>
<td>D13</td>
<td>10.1 ab</td>
<td>60.6</td>
<td>22.3 bc</td>
<td>28.0 bc</td>
<td>288 abc</td>
<td>310.0</td>
</tr>
<tr>
<td></td>
<td>NL13</td>
<td>9.7 a</td>
<td>60.1</td>
<td>20.1 ab</td>
<td>22.8 a</td>
<td>290 abc</td>
<td>283.3</td>
</tr>
</tbody>
</table>

* Water absorption capacity is considered good above 58%, while anything below 54% is considered a poor rating, affecting shelf life, reducing dough volume, and negatively affecting dough proof times. In terms of wet gluten content (%), any value less than 23% is considered a poor rating, and any value over 27% can be classified as good or acceptable. Poor ratings for sedimentation are values under 22 mL, while good values are indicated over 28 mL. The Hagberg Falling Number (HFN) is measured in seconds, whereby a poor rating is considered anything less than 180 s and a good rating is from 240–320 s. However, any HFN values over 320 s are also poorly rated. The baking volume was measured using wholemeal test loaves with loaf volumes below 330 mL for poor and above 350 mL for good (Dr. Aberham, personal communication,[18]).

The water absorption capacity of all experimental entries was good, with the best values achieved by the reference varieties. In contrast, none of the entries achieved satis-
factory wet gluten content, with the majority of the CCPs, except D13, achieving lower values than the reference varieties (Table 5). Similarly, the sedimentation values of the three reference varieties, particularly Akteur (35.8 mL), were higher than those of the CCP entries. Hagberg Falling Numbers (HFNs) for all 11 CCPs were good (values ranged from 249–304 s), while both reference varieties, Achat and Akteur, achieved HFN values above 320 s and were considered poor (low enzymatic activity). Baking volume was poor for all entries, including the reference varieties, with no statistically significant differences. The greatest baking volume was achieved by the reference variety Akteur (319 mL), followed by the CCPs D13 (310 mL) and CH13 (307 mL). The reference variety Capo achieved the lowest baking volume of 278 mL, followed by the F13 CCP (279 mL) (Table 5).

4. Discussion

The initial seed size variation among CCPs with different histories affected their performance depending on growing conditions. In 2013/14, characterised by higher precipitation and low disease pressure, high TKW at sowing carried through to greater TKW at harvest. Potential effects on stem length at early developmental stages did not significantly affect final stem length. In contrast, in the very dry season of 2014/15, seed weight effects on stem length carried through to maturity. As seed quality, which includes seed size, affects not only yield potential but also production and resource-use efficiency [46], the legacy effect of seed size variation due to differing production environments resulted in strong differences in trait profiles between the CCP entries depending on the experimental season. Stanton [47] reported on the positive effects of seed size on growth rate and flower number in wild radish plants; however, these effects were only present in competitive environments. In wheat, both Iqbal et al. [48] and Shoaib et al. [49] reported increased plant height and ear length with increasing seed size under semi-arid conditions. Under the challenging climatic conditions of 2014/15, the differences in seed size at sowing in the F13 CCPs may thus have affected stem length. Consequently, this effect was no longer present in the F14 CCPs when seed size was more similar. The marked water deficit in 2014/15 restricted later crop growth and grain filling. As seed size at sowing led to greater investment in stem length, this resulted in negative effects on HI. Removing most of the variation in TKW at sowing for the F14 was effective in removing the effects of seed size at sowing. Thus, the results of the F14 are likely only influenced by genotype-environment effects, while those of the F13 were also affected by the previous growing conditions of the seeds.

Under the drought conditions of 2014/15, the CCPs achieved the same yields as the reference varieties. In 2013/14, four out of eleven CCPs were outyielded by the variety Capo, while seven CCPs, Capo, and Achat achieved higher yields than Akteur. Thus, only under less stressful conditions in the first year could at least one of the reference varieties (Capo) take advantage of its genetic background compared to the CCPs. In contrast, the excellent buffering capacity of genetic diversity in the face of stress enabled the CCPs to perform comparably to the included reference varieties under drought conditions. The performance of the CCPs under these conditions is impressive considering their older parental genetics (see [50]) and the development of these CCPs under natural selection in comparison to the reference varieties, which were strongly selected over several generations in order to pass the official DUS (distinctiveness, uniformity, and stability) and VCU (value for cultivation and use) tests for variety release.

Superior capacity of evolutionary populations (EPs) to maintain yields under drought conditions has been reported for barley EPs [51], as well as for winter wheat CCPs under organic management [21]. A number of researchers have reported on the advantages of EPs in terms of yield and yield stability in comparison to a number of line selections and reference varieties [4,6,14,15,19,20]. Raggi et al. [19] attribute the yield stability and good agronomic performance of a barley EP to the unpredictability and changing environmental conditions of the breeding station during population development over nine years, which would have selected for genotypes better suited to “contrasting climatic conditions”. While natural selection of a population within a single location will result in more specific or
higher “temporal” adaptation [28], “spatial” adaptation may be lower [4]. Despite natural selection driving greater temporal adaptation of genotypes grown in a single environment, temporal adaptation is confounded not only by genotype-location interactions (GEI) but also by the complexity that year effects are becoming increasingly variable, and as such, GEI are replaced by genotype-year interactions within location interactions [5,52]. For this reason, the individual and population buffering potential of EPs is greatly advantageous under variable environments [52], particularly due to the inherent genetic diversity and the response of natural selection under highly variable environments favouring those genotypes with higher yielding capacity, which is strongly associated with individual fitness as those genotypes most successful in a number of given environments also have the highest potential to maximise yield [53]. The HU13NC CCP, which in the first experimental year was one of the lowest-yielding CCPs, yielded only 0.1 t/ha less during the dry second experimental year, while the other CCPs yielded between 0.3 and 0.8 t/ha less, suggesting specific adaptation of this non-cycling population to drier continental environmental conditions. However, the overall similarity of the CCPs based on their origins and cycling histories for agronomic traits in this study may indicate that these CCPs have undergone little directional selection and therefore still have a high buffering potential, as discussed by Patel et al. [31].

Similar results were reported by Mežaka et al. [54], who reported on small differences in yield but not in grain quality traits in barley EPs cultivated under both organic and conventional management, in this case indicating little adaptation to specific growing conditions under natural selection. A study by Rhoné et al. [25] found a high level of adaptive diversity in some wheat EPs even after 12 generations, despite local genetic drift and selection for certain earliness traits. However, depending on the scale and consistency of the selection pressure, divergent evolutions of EPs for agronomic traits due to specific adaptation have been reported by Bocci et al. [4] for grain yield in a wheat EP grown in Sicily, a location with sufficiently constant climatic conditions to ensure directional selection. Merrick et al. [6] discuss the relevance of a sufficient number of generations needed for directional selection of EPs as well as the importance of sufficient genetic diversity and potential based on parental selection.

Kingsolver et al. [55] report that natural selection for morphology is greater than selection for phenology. Significant EP origin and cycling history effects on degree of awnedness were also reported in a number of dynamically managed French wheat EPs. Three out of four French wheat EPs had a significantly greater number of awned ears in comparison to their original EP after eight years of multiplication at two sites [10,56]. No comparison to the original CCPs in 2008 is available for the degree of awnedness in this study. However, the DK13, F13, and UK13NC CCPs had significantly reduced frequencies of awned ears in comparison to the HU13NC consistently, irrespective of generation and experimental year. This suggests divergent evolutions for the morphological trait of awned ears per m$^2$ in the relatively short period of five generations. Awns are associated with better drought tolerance as awns can contribute significantly to photosynthesis, which is essential in drought conditions where leaves may no longer be able to fulfil this capability due to early senescence [57]. The HU13NC CCP was grown in a continental climate characterised by warmer and drier summers and, as pointed out above, did not differ in yield due to the drought conditions of 2014/15. Thus, the different environmental conditions under which the CCP has been grown may have selected for a higher degree of ear awnedness, particularly under drier conditions, contributing to the population’s ability to cope better with drought stress.

The speed with which heterogeneous populations may diverge from one another is dependent not only on the potential of the inherent genetic diversity for adaptive traits and their heritability but also on the degree and consistency of the selective pressure in a given environment [55]. Bertholdsson et al. [27] and Vijaya Bhaskar et al. [28] reported on the divergent evolutions of wheat CCPs for seminal root length and root weight under organic versus conventional management conditions after only four generations. Additionally, even
within the same management system, two CCPs with the same genetic background differed from one another for all recorded seedling traits already in the F₆. The divergent trajectory of these two conventionally managed CCPs was also confirmed by yield stability [20]. Yield differentiation in a wheat EP multiplied in Sicily over four generations diverged significantly from those EPs that had been grown in Tuscany [4].

The genetic background of the CCP entries in this study is based on both high baking quality and high yielding varieties, and while yield may be subject to natural selection [11, 53], baking quality parameters are not associated with plant fitness and, as such, are not considered to be shaped by natural selection [53]. In 2013, the German non-cycling CCPs based on the 20 parents had similar baking volume, protein content, and HFN values as Capo, Akteur, and Achat. However, the baking volume of the CCPs based only on the 12 high-quality parents was significantly higher in comparison to the CCPs based on all 20 parents and those based on the high-yielding parents, confirming the importance of parental variety choice [18]. Although the baking volume values were poor in 2015 for all experimental entries, no significant differences were found between the CCPs and the reference varieties. Interestingly, CCP D13, which has a higher frequency of darker-coloured ears and kernels, tended towards the best protein content and GPD values of all CCP entries. Darker-coloured kernels have been associated with higher protein values [58]. Using a colour sorter to separate darker and lighter kernels in a wheat CCP, Döring et al. [59] found that the darker grain fraction had a protein content 1.6% greater than the lighter kernel fraction. In the F₁₄ of 2014/15, the CCP D13 had achieved a moderate protein content, which was comparable to the reference varieties Achat and Akteur. Not only was the protein content in D13 greater in comparison to most CCP entries, but sedimentation value, wet gluten content, and baking volume tended to be as good as the baking quality of the reference varieties. The darker kernels and better baking quality parameter values in the CCP D13 may indicate divergent evolution for a morphological trait that apparently occurred within five generations.

Despite the generally low disease pressure (measured as non-green leaf area, NGLA) in both years, the extended drought conditions experienced from November to June 2014/15 led to physiological stress and, as such, may have caused a higher predisposition to disease in the plants [60]. This was combined with early senescence, leading to an overall higher NGLA in all entries. As stripe rust does not usually play a role in Hungary, it is possible that changes in resistance frequencies were different in the cycling Hungarian CCPs in comparison to the HU13NC CCP. Changes in resistance frequencies due to diseases have been reported for wheat [23,61,62] and barley [63–66] EPs. The non-cycling German CCP (D13NC) did not indicate a lower disease incidence value and, as such, adaptation to the experimental site. However, over the past 13 generations, disease severity in the CCPs at the experimental site has been low, even in the presence of high disease pressure, in comparison to various susceptible reference varieties used in other years and some of the parental varieties grown in 2008/09 and 2014/15 [21], demonstrating the capacity of genetic diversity to buffer disease pressure in populations.

The variable interrelations of the agronomic traits as affected by year and generation demonstrate the plasticity and therefore general adaptability of the CCPs, based on the premise that plant plasticity is thought to be adaptive, especially as the result of environmental variation [67,68]. Changes in seed size and the number of kernels per ear have been found between generations in a number of dynamically managed French wheat populations [69]. Within this study, significant CCP and generation interaction effects were found for seed size as well as the number of ears per m², with a significant CCP and year interaction effect also found for kernel number per ear. Seed size heritability is considered low [70], and variation in seed size and number tends to be a result of phenotypic plasticity, as confirmed by the lower variation in seed sizes among the CCPs grown for one year in the same field. Kernel number per ear and number of ears per m² reacted dynamically depending on either experimental year or generation effects.
The number of ears per m² of the CCPs was generally greater than the reference varieties, but was only significantly greater than the variety Akteur. Increased diversity has been shown to increase plant biomass [71], and a study by Baresel et al. [15] reported on the greater number of ears per m² of a number of CCPs with the same genetic base as those used in this study in comparison to reference varieties under organic conditions. Generally, the combination of a high level of heterogeneity and a high level of heterozygosity, especially in early generations, allows for a wider response range to various environments [6]. Additionally, the differential trait profiles most noticeable between the two generations in 2014/15 (Figures 4 and S1) may be influenced by the legacy effect of seed size at sowing interacting with the dry weather, as well as significant differences between the generations for TKW and their subsequent effect on stem length and HI. The correlation between HI values and the number of ear-bearing tillers per m² of the F₁₃ was negative in 2013/14 and positive in 2014/15. As HI represents the ratio of grain yield to total plant biomass, the negative correlation between HI and number of ears/m² in 2013/14 indicates that greater plant biomass in the form of ears per m² reduced HI values as resources were invested in plant biomass rather than grain yield (kernel number and TKW). However, in the drought conditions of 2014/15, a higher number of ears per m² was correlated with higher HI values, reflecting the more challenging growing conditions of the season and leading to significantly shorter plants with less overall biomass. The increased number of ears per m² thus contributed to grain yield, thereby increasing HI. The fact that a minimal correlation between HI values and the number of ear-bearing tillers per m² was found in the F₁₄ of 2014/15 suggests that seed weight at sowing may have played a determinant role in the correlation between HI and the number of ear-bearing tillers per m² that interacted with the climatic conditions.

Protein content was significantly and negatively correlated with kernel number per ear and TKW in the F₁₃ of 2013/14, with no such correlations found in the F₁₄ of 2014/15. This indicates that differing growth stages were important for final protein content, which would have been affected by the differing climatic conditions. In addition, the positive correlation between disease values and protein content in the F₁₃ of 2013/14 may indicate that disease pressure restricted yield early on, resulting in higher grain protein due to the inverse relationship found between yield and protein content, despite the very low disease pressure. In both the F₁₃ of 2013/14 and the F₁₄ of 2014/15, there was a significant negative correlation between protein content and HI values. This highlights the complexity of dissecting the environmental effects of differing seed weights at sowing. Considering that protein and yield values did not vary much overall, this demonstrates the plasticity of the interrelationships of agronomic parameters in the CCPs in general.

5. Conclusions
1. While the CCPs were capable of levelling out the effects of differing seed sizes on final yield and protein content through yield component parameters when environmental conditions were less stressful in 2013/14, this was not the case during the drought in 2014/15. Thus, for stringent comparisons of genotype-environment interactions, differences between the CCP entries and generations due to differing seed weight at sowing should be avoided, e.g., by a round of seed reproduction in a common environment before future experiments. This will ensure that only genetic differences between entries are reported.
2. The similarity in performance of the cycling and non-cycling populations and the high trait profile plasticity in their response to environmental conditions demonstrate the value of well-designed heterogeneous populations in the face of unpredictable environmental conditions. The CCPs were more climate resilient and better able to buffer environmental stress than the pure-line reference varieties, resulting in a lower yield reduction under drought stress.
3. Constantly changing environmental conditions did not confer obvious agronomic advantages or disadvantages to the cycling CCPs, highlighting the buffering capacity...
of intraspecific diversity and confirming the research by Patel et al. [31] that large environmental variance results in slow population evolution. This is of interest to farmers who want to cultivate EPs but do not want to produce their own seed, relying rather on regionally decentralised certified seed production, which should not affect EP performance.

4. In contrast, continuously applying differential selection environments for only five to eight years can lead to significant changes. Thus, the increased number of awned ears in the Hungarian home population, as well as the slightly higher susceptibility to stripe rust, fit the overall more continental environmental conditions with colder winters and lower precipitation in Hungary.

5. Overall, EPs show great potential in terms of agronomic performance, particularly under conditions where higher biotic and abiotic pressures exist. Decentralised development of such populations allows for the dynamic maintenance of intra-specific diversity with high plasticity. In addition to providing an interesting alternative to genetically homogeneous varieties, it can contribute to the development of future genetic resources.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/agronomy13061662/s1, Table S1: ANOVA results of the origin (D, HU, UK) and cycling history effects on the F13 and F14 winter wheat CCPs in 2014/15 (Neu-Eichenberg, Germany), including mean squares (MS) and significance levels, Table S2: Yield (t/ha), HI, RAUDPC, TKW (g), ear-bearing tillers/m² (Ears_m²), kernel number per ear (Kernel no.), protein content (%), percentage of ears with awns (Awned ears (%)), stem length (cm) and ear length (cm) of the F13 and F14 winter wheat CCP entries in the experimental years 2014/15 (Neu-Eichenberg, Germany), Table S3: ANOVA results of the origin (D, HU, UK) and cycling history effects on the F13 winter wheat CCPs of 2013/14 and 2014/15 (Neu-Eichenberg), including mean squares (MS) and significance levels, Table S4: Yield (t/ha), HI, RAUDPC, TKW (g), ear-bearing tillers/m² (Ears_m²), kernel number per ear (Kernel no.), protein content (%), percentage of ears with awns (Awned ears (%)), stem length (cm) and ear length (cm) of the F13 winter wheat CCP entries in the experimental years 2013/14 and 2014/15 (Neu-Eichenberg, Germany), Table S5: Grain protein deviation (GPD) of the reference varieties and winter wheat CCP entries of the F13 in 2013/14 and F13 and F14 in 2014/15 (Neu-Eichenberg, Germany) and Figure S1: Correlation coefficients (r) for various agronomic parameters for the winter wheat CCP entries over all generations and experimental years (black), of the F13 in 2013/14 (red), of the F13 in 2014/15 (green) and of the F14 in 2014/15 (blue) (Neu-Eichenberg, Germany) and Figure S2: Grain protein deviation (GPD) and protein yield (kg/ha) of the experimental entries (winter wheat CCPs and reference varieties) in the (A) F13 of 2013/14 and 2014/15 and (B) F14 of 2014/15 (Neu-Eichenberg, Göttingen).


Funding: The authors O.D.W., S.B., A.H. and M.R.F. at the University of Kassel (BÖLN-Grant agreement 2812OE021), J.P.B. at TUM School of Life Sciences (BLE-Grant Agreement 12OE022) and A.B. in Denmark (Naturerhverv Grant agreement 34009-12-0589) were supported through the project CORE Organic II COBRA. O.D.W., J.P.B. and M.R.F. additionally received funding from the project INSUSFAR (BMBF-Grant agreement FKZ031A350C). Current funding for O.D.W., A.H. and M.R.F. has been provided through the project BAKWERT (BÖL—Grant agreement 2819OE033). Current funding for P.M. has been provided through the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 771367 (ECOBREED).
Data Availability Statement: The data presented in this study are openly available in Dryad at https://doi.org/10.5061/dryad.x95x69pq3.

Acknowledgments: We thank both Geza Kovacs and Martin Wolfe for their contribution and dedication to breeding and research on heterogeneous populations. Their contributions to the subject have inspired and provided a guiding light for many researchers past, present and future.

Conflicts of Interest: The authors declare no conflict of interest.

References


42. Yan, W.; Rajcan, I. Biplot analysis of test sites and trait relations of soybean in Ontario. *Crop Sci.* **2002**, *42*, 11–20. [CrossRef]


Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.