**Table 1.** Descriptive statistics of agronomic traits

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Traits** | **Trait code** | **Reference** | **Unit** | **Range** | **Mean** | **Standard deviation** | **CV(%)** |
| **Quantitative traits** | Dried seed weight | DSW | IPGRI and NBPGR 2004 | g | 153.58 | 32.82 | 24.32 | 74.09 |
| Harvest index | HI | IPGRI and NBPGR 2004 | - | 0.48 | 0.12 | 0.09 | 72.39 |
| Number of branches per plant | BNU | UPOV 2014 | count | 42 | 5.47 | 3.91 | 71.53 |
| Number of capsules per plant | CNU | IPGRI and NBPGR 2004 | count | 1195 | 113.21 | 73.01 | 64.49 |
| Number of capsule per leaf axil | CNLA | IPGRI and NBPGR 2004 | count | 2 | 1.41 | 0.76 | 53.87 |
| Dried Biomass weight | DBI | Defined in this study | g | 1141 | 302.63 | 138.80 | 45.86 |
| Productive axis length | PAL | UPOV 2014 | cm | 205 | 77.23 | 29.51 | 38.20 |
| Stem diameter | DIA | Defined in this study | mm | 53.40 | 15.98 | 5.58 | 34.95 |
| Days to 50% flowering | FLO | IPGRI and NBPGR 2004 | day | 83 | 51.50 | 12.86 | 24.97 |
| Plant height | PHE | IPGRI and NBPGR 2004 | cm | 224 | 155.36 | 35.66 | 22.95 |
| Days from 50% flowering to 50% physiological maturity | FTM | Defined in this study | day | 50 | 48.95 | 8.73 | 17.85 |
| Capsule length | CLE | IPGRI and NBPGR 2004 | mm | 37 | 25.08 | 4.47 | 17.83 |
| 1000-seeds weight | TSW | IPGRI and NBPGR 2004 | g | 2.86 | 2.76 | 0.48 | 17.62 |
| Capsule width | CWI | IPGRI and NBPGR 2004 | mm | 15.2 | 7.75 | 1.18 | 15.28 |
| Days to 50% maturity | MAT | Defined in this study | day | 99 | 100.4679443 | 13.15873489 | 13.09 |
| Number of locules per capsule | NLC | IPGRI and NBPGR 2004 | 1-3 scale | Four locules (99.21%), 6 locules (0.39%), 8 locules (0.39%) | | | |
| **Qualitative traits** | Branching type | Branching\_type | Defined in this study | 1-6 scale | Highly branched [number of branches > 10] (61.73%), Moderately branched [3 < number of branches < 9] (18.93%),Extremely branched [Bushy-type] (10.84%), Bi-branched [two branches] (5.12%), Mono-branched [1 branch] (3.15%), unbranched (0.2%) | | | |
| Capsule hairiness | Cap\_Hairiness | IPGRI and NBPGR 2004 | 1-4 scale | Sparse (81.06%), Medium (16.96%), Profuse (1.18%), Glaborous (0.78%) | | | |
| Flower colour | Flower\_Colour | Defined in this study | 1-3 scale | White (97.04%), Pink (2.76%), Purple (0.2%) | | | |
| Inflorescence type | Inflorescence\_type | IPGRI and NBPGR 2004 | 1-2 scale | Indeterminate (99.60%), determinate (0.4%) | | | |
| Seed colour | Seed\_Colour | Defined in this study | 1-10 scale | White (33.72%), Brown (21.49%), Composite (20.31%), Black (9.66%), Light Brown (7.88%), Light gray (2.76%), Light yellow (1.57%), Dark brown (1.57%), Red brown (0.39%), Yellow (0.39%) | | | |

**Table 2.** Shanon-Weiner and Simpson diversity index for five qualitative traits

|  |  |  |
| --- | --- | --- |
| Traits | Shanon-Weiner Index (H) | Simpson's Index (D) |
| Inflorescence type | 0.026 | 0.008 |
| Flower color | 0.141 | 0.057 |
| Seed color | 1.722 | 0.781 |
| Branching type | 1.127 | 0.568 |
| Capsule hairiness | 0.562 | 0.314 |

**Table 3.** Direct and indirect effect of various traits on dried seed weight showed by path coefficient analysis

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| PHE | **-0.091** | -0.030 | -0.021 | -0.012 | -0.007 | 0.199 | -0.578 | -0.007 | -0.314 | 0.392 | -0.075 | 0.001 | 0.014 | 0.050 | -0.001 |
| CNU | 0.016 | **0.177** | -0.007 | -0.001 | 0.044 | 0.007 | 0.215 | 0.006 | 0.157 | -0.162 | 0.006 | 0.000 | 0.009 | -0.008 | 0.000 |
| BNU | -0.033 | 0.020 | **-0.060** | -0.009 | -0.051 | 0.148 | -0.396 | -0.005 | -0.191 | 0.196 | 0.006 | 0.001 | 0.029 | 0.056 | -0.001 |
| DIA | -0.062 | 0.009 | -0.030 | **-0.017** | -0.040 | 0.228 | -0.537 | -0.007 | -0.314 | 0.346 | -0.017 | 0.001 | 0.019 | 0.043 | -0.001 |
| PAL | 0.005 | 0.057 | 0.022 | 0.005 | **0.139** | -0.090 | 0.330 | 0.011 | 0.269 | -0.312 | 0.034 | -0.001 | -0.018 | -0.030 | 0.000 |
| DBI | -0.050 | 0.004 | -0.024 | -0.011 | -0.035 | **0.361** | -0.438 | -0.006 | -0.280 | 0.335 | -0.046 | 0.001 | 0.019 | 0.041 | 0.000 |
| HIN | 0.064 | 0.046 | 0.029 | 0.011 | 0.055 | -0.192 | **0.826** | 0.010 | 0.359 | -0.450 | 0.080 | -0.001 | -0.017 | -0.060 | 0.001 |
| TSW | 0.029 | 0.050 | 0.013 | 0.006 | 0.067 | -0.101 | 0.355 | **0.022** | 0.336 | -0.358 | 0.017 | -0.001 | -0.019 | -0.025 | -0.001 |
| FLO | -0.026 | -0.025 | -0.010 | -0.005 | -0.033 | 0.090 | -0.264 | -0.007 | **-1.121** | 0.993 | 0.121 | 0.001 | 0.014 | 0.032 | 0.000 |
| MAT | -0.031 | -0.025 | -0.010 | -0.005 | -0.037 | 0.105 | -0.322 | -0.007 | -0.964 | **1.154** | -0.184 | 0.001 | 0.015 | 0.041 | 0.000 |
| FTM | -0.012 | -0.002 | 0.001 | -0.001 | -0.008 | 0.029 | -0.116 | -0.001 | 0.235 | 0.369 | **-0.574** | 0.000 | 0.002 | 0.016 | 0.000 |
| CLE | 0.026 | 0.005 | 0.024 | 0.007 | 0.061 | -0.112 | 0.396 | 0.009 | 0.258 | -0.289 | 0.029 | **-0.003** | -0.023 | -0.039 | -0.001 |
| CWI | 0.021 | -0.025 | 0.027 | 0.005 | 0.040 | -0.112 | 0.231 | 0.007 | 0.258 | -0.277 | 0.023 | -0.001 | **-0.062** | -0.047 | 0.002 |
| CNLA | 0.043 | 0.012 | 0.031 | 0.007 | 0.039 | -0.137 | 0.462 | 0.005 | 0.336 | -0.439 | 0.086 | -0.001 | -0.027 | **-0.107** | -0.001 |
| NLC | 0.006 | -0.002 | 0.003 | 0.001 | -0.003 | -0.004 | 0.025 | -0.001 | -0.011 | 0.023 | -0.006 | 0.000 | -0.005 | 0.003 | **0.020** |

In bold and positive: direct effect; in bold and negative: indirect effect; See Table 1 for trait details.

**Table 4.** Quantitative traits associated to each cluster from the worldwide panel following the v test

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Clusters** | **Traits** | **Mean in the cluster** | **sd in the cluster** | **Overall mean** | **Overall sd** | **v test** | **p-value** |
| Cluster 1 | DIA | 18.91 | 3.25 | 16.08 | 3.66 | 12.86 | p < 0.01 |
| PHE | 178.45 | 19.32 | 156.09 | 29.86 | 12.46 | p < 0.01 |
| DBI | 393.75 | 124.73 | 316.25 | 127.62 | 10.10 | p < 0.01 |
| FLO | 60.82 | 10.69 | 54.16 | 11.20 | 9.88 | p < 0.01 |
| MAT | 107.43 | 10.56 | 100.88 | 11.46 | 9.49 | p < 0.01 |
| BNU | 6.86 | 2.32 | 5.50 | 2.48 | 9.04 | p < 0.01 |
| Cluster 2 | CNU | 137.95 | 45.94 | 112.68 | 45.63 | 10.49 | p < 0.01 |
| DSW | 38.07 | 18.81 | 28.93 | 20.22 | 8.56 | p < 0.01 |
| TSW | 2.81 | 0.32 | 2.64 | 0.43 | 7.45 | p < 0.01 |
| PAL | 81.62 | 17.12 | 73.38 | 21.53 | 7.24 | p < 0.01 |
| FTM | 48.03 | 6.66 | 46.71 | 6.04 | 4.13 | p < 0.01 |
| HIN | 0.12 | 0.05 | 0.11 | 0.08 | 3.08 | p < 0.01 |
| Cluster 3 | CNLA | 2.44 | 0.82 | 1.42 | 0.76 | 16.41 | p < 0.01 |
| HIN | 0.21 | 0.07 | 0.11 | 0.08 | 14.46 | p < 0.01 |
| CWI | 8.63 | 1.05 | 7.72 | 0.93 | 12.12 | p < 0.01 |
| CLE | 28.16 | 3.37 | 24.95 | 3.73 | 10.61 | p < 0.01 |
| DSW | 40.30 | 14.57 | 28.93 | 20.22 | 6.93 | p < 0.01 |
| TSW | 2.87 | 0.28 | 2.64 | 0.43 | 6.77 | p < 0.01 |
| PAL | 84.06 | 18.33 | 73.38 | 21.53 | 6.11 | p < 0.01 |

sd: standard deviation, See Table 1 for trait codes details

**Table 5.** Metrics showing the quality of the inferred core collection

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Range** | | | **CV** | | | **Mean** | | | **Variance** | | |  |
| **Traits** | **Whole** | **Core** | **CRR (%)** | **Whole** | **Core** | **VR(%)** | **Whole** | **Core** | **MDP(%)** | **Whole** | **Core** | **VDP(%)** | **Whole versus core means comparison( p-value)** |
| PHE | 171.30 | 168.5 | 98.37 | 0.19 | 0.21 | 108.38 | 156.12 | 151.00 | 3.39 | 891.91 | 980.00 | 8.99 | p = 0.18 |
| CNU | 394.70 | 231.7 | 58.70 | 0.41 | 0.36 | 89.65 | 112.80 | 111.17 | 1.47 | 2089.04 | 1630.70 | 28.11 | p = 0.98 |
| BNU | 15.50 | 11.36 | 73.28 | 0.45 | 0.52 | 114.42 | 5.51 | 5.05 | 8.99 | 6.18 | 6.81 | 9.27 | p = 0.10 |
| DIA | 21.20 | 19.49 | 91.94 | 0.23 | 0.25 | 107.83 | 16.08 | 15.59 | 3.19 | 13.40 | 14.63 | 8.42 | p = 0.18 |
| PAL | 123.22 | 102 | 82.78 | 0.29 | 0.34 | 114.42 | 73.41 | 71.02 | 3.36 | 464.06 | 568.71 | 18.40 | p = 0.35 |
| DBI | 757.50 | 549.50 | 72.54 | 0.40 | 0.37 | 90.74 | 316.27 | 303.22 | 4.30 | 16287.15 | 12328.18 | 32.11 | p = 0.52 |
| DSW | 114.36 | 102.81 | 89.90 | 0.70 | 0.64 | 91.13 | 28.90 | 30.62 | 5.65 | 409.34 | 381.83 | 7.20 | p = 0.38 |
| HIN | 0.39 | 0.31 | 77.94 | 0.75 | 0.70 | 92.80 | 0.11 | 0.12 | 10.91 | 0.01 | 0.01 | 7.83 | p = 0.71 |
| TSW | 2.42 | 2.10 | 86.59 | 0.16 | 0.15 | 93.02 | 2.64 | 2.63 | 0.53 | 0.18 | 0.16 | 16.81 | p = 0.51 |
| FLO | 53 | 36 | 67.92 | 0.21 | 0.19 | 90.75 | 54.15 | 54.97 | 1.49 | 125.63 | 106.62 | 17.82 | p = 0.31 |
| MAT | 59.50 | 42 | 70.59 | 0.11 | 0.11 | 95.92 | 100.88 | 101.47 | 0.58 | 131.36 | 122.27 | 7.43 | p = 0.62 |
| FTM | 50 | 40.50 | 81 | 0.13 | 0.15 | 116.87 | 46.72 | 46.49 | 0.49 | 36.56 | 49.45 | 26.06 | p = 0.34 |
| CLE | 27 | 19.40 | 71.85 | 0.15 | 0.15 | 97.10 | 24.94 | 24.96 | 0.06 | 13.91 | 13.13 | 5.94 | p = 0.71 |
| CWI | 7.38 | 5.55 | 75.25 | 0.12 | 0.13 | 105.22 | 7.72 | 7.89 | 2.21 | 0.87 | 1.00 | 13.62 | p = 0.07 |
| CNLA | 2 | 2 | 100 | 0.54 | 0.54 | 100.51 | 1.42 | 1.59 | 10.71 | 0.58 | 0.74 | 21.09 | p = 0.20 |
| NLC | 4 | 2 | 50 | 0.07 | 0.07 | 99.08 | 4.02 | 4.04 | 0.38 | 0.08 | 0.08 | 1.08 | p = 0.94 |
|  |  | | 78.04 |  | | 100.49 |  | | 3.61 |  | | 14.39 | - |

CV: Coefficient of variation, CRR: Coincidence rate of range, VR: Variable rate, VDP: Variance difference percentage

**Table 6.** Quantitative traits associated with each cluster from the core collection

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Clusters** | **Traits** | **Mean in the cluster** | **sd in the cluster** | **Overall mean** | **Overall sd** | **v test** | **p-value** |
| Cluster1 | CNLA | 2.59 | 0.67 | 1.74 | 0.90 | 6.55 | p < 0.01 |
| HIN | 0.22 | 0.05 | 0.15 | 0.07 | 6.05 | p < 0.01 |
| CWI | 8.83 | 0.79 | 8.01 | 1.06 | 5.31 | p < 0.01 |
| Protein | 28.90 | 1.23 | 26.97 | 2.57 | 5.19 | p < 0.01 |
| Oleic\_acid | 44.78 | 2.52 | 42.08 | 3.73 | 5.01 | p < 0.01 |
| CLE | 28.03 | 2.23 | 25.80 | 3.23 | 4.78 | p < 0.01 |
| TSW | 2.86 | 0.22 | 2.73 | 0.34 | 2.72 | p < 0.01 |
| Alpha\_linoleic\_acid | 0.41 | 0.19 | 0.35 | 0.16 | 2.50 | p < 0.01 |
| PAL | 83.59 | 20.12 | 77.71 | 19.88 | 2.05 | p < 0.01 |
| Cluster 2 | CNU | 150.52 | 30.91 | 122.07 | 33.67 | 4.42 | p < 0.01 |
| BNU | 5.76 | 2.00 | 4.39 | 2.37 | 3.01 | p < 0.01 |
| DSW | 45.82 | 18.51 | 36.85 | 16.85 | 2.78 | p < 0.01 |
| DBI | 325.84 | 83.47 | 280.53 | 101.93 | 2.32 | p < 0.01 |
| Cluster3 | PHE | 173.35 | 19.11 | 146.12 | 28.25 | 5.56 | p < 0.01 |
| Linoleic\_acid | 46.37 | 2.54 | 43.01 | 3.59 | 5.39 | p < 0.01 |
| MAT | 107.03 | 9.68 | 99.75 | 10.48 | 4.01 | p < 0.01 |
| Oil | 52.85 | 3.29 | 50.15 | 3.88 | 4.01 | p < 0.01 |
| DIA | 17.11 | 3.75 | 14.75 | 3.46 | 3.94 | p < 0.01 |
| Sesamin | 4.39 | 1.67 | 3.08 | 1.99 | 3.83 | p < 0.01 |
| FTM | 50.72 | 7.88 | 46.29 | 6.75 | 3.79 | p < 0.01 |
| DBI | 340.18 | 96.74 | 280.53 | 101.93 | 3.38 | p < 0.01 |
| BNU | 5.66 | 1.79 | 4.39 | 2.37 | 3.10 | p < 0.01 |

sd: standard deviation, See Table 1 for trait codes details