Canonical correlations reveal adaptive loci and phenotypic responses to climate in perennial ryegrass

Blanco-Pastor, José Luis; Barre, Philippe; Keep, Thomas; Ledauphin, Thomas; Escobar-Gutiérrez, Abraham; Roschanski, Anna Maria; Willner, Evelyn; Dehmer, Klaus J.; Hegarty, Matthew; Muylle, Hilde; Veeckman, Elisabeth; Vandepoele, Klaas; Ruttink, Tom; Roldán-Ruiz, Isabel; Manel, Stéphanie; Sampoux, Jean Paul

Published in:
Molecular Ecology Resources

DOI:
10.1111/1755-0998.13289

Publication date:
2021

Citation for published version (APA):
Blanco-Pastor, J. L., Barre, P., Keep, T., Ledauphin, T., Escobar-Gutiérrez, A., Roschanski, A. M., Willner, E., Dehmer, K. J., Hegarty, M., Muylle, H., Veeckman, E., Vandepoele, K., Ruttink, T., Roldán-Ruiz, I., Manel, S., & Sampoux, J. P. (2021). Canonical correlations reveal adaptive loci and phenotypic responses to climate in perennial ryegrass. *Molecular Ecology Resources*, *21*(3), 849-870. https://doi.org/10.1111/1755-0998.13289

MOLECULAR ECOLOGY
RESOURCES

Supplemental Information for:

Canonical correlations reveal adaptive loci and phenotypic responses to climate in perennial ryegrass

J.L. Blanco-Pastor*[1], P. Barre[1], T. Keep[1], T. Ledauphin[1], A. Escobar-Gutiérrez[1], A.M. Roschanski[2], E. Willner[2], K. J. Dehmer[2], M. Hegarty[3], H. Muylle[4], E. Veeckman[4, 5, 6], K. Vandepoele[4, 5, 7], T. Ruttink[4], I. Roldán-Ruiz[4, 6], S. Manel[8], J.P. Sampoux[1]

1. INRAE, UR4 (URP3F), Centre Nouvelle-Aquitaine-Poitiers, 86600 Lusignan, France;
2. Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Inselstr. 9, 23999 Malchow/Poel, Germany;
3. Institute of Biological, Environmental and Rural Sciences (IBERS), Aberystwyth University, Aberystwyth, SY23 3FL Ceredigion, UK;
4. Flanders Research Institute for Agriculture, Fisheries and Food (ILVO) - Plant Sciences Unit, Caritasstraat 39, 9090 Melle, Belgium;
5. Bioinformatics Institute Ghent, Ghent University, 9052 Ghent, Belgium;
6. Department of Plant Biotechnology and Bioinformatics, Ghent University, 9052 Ghent, Belgium;
7. Center for Plant Systems Biology, VIB, 9052 Ghent, Belgium;
8. EPHE, PSL Research University, CNRS, UM, UPV, IRD, UMR 5175 CEFE, 34293 Montpellier, France.

Contents

Methods S1: HiPlex SNP set

HiPlex primers were designed with Primer3 (Untergasser *et al.*, 2012) using prior knowledge of allelic sequence variation (Veeckman *et al.*, 2018). The primer pairs were then divided into two highly multiplex PCR-reactions according to their amplification efficiency (Supporting Information, Table S2). DNA was extracted using the CTAB method of Murray & Thompson (1980) and DNA concentration was measured using the Quantus double-stranded DNA assay (Promega, Madison, WI, USA). Per sample, the final DNA concentration was adjusted to 40 ng/µL and the amplicons were PCR-amplified while adding sample-specific indices. Libraries were prepared using the KAPA Hyper Prep PCR-free Kit according to manufacturer directions (Kapa Biosystems, USA). HiPlex amplification reactions and library preparations were outsourced to Floodlight Genomics LLC (Knoxville, TN, USA). The libraries were sequenced with 2x150 PE on a HiSeq2500 instrument in rapid run mode. Paired-end reads were merged with PEAR (v0.9.8) (Zhang *et al.*, 2013) and adapter sequences were removed.

Methods S2: Environmental variables (climate-related variables and soil variables)

See Table S4 for a complete list of environmental variables and for the values of these variables per site of origin of perennial ryegrass populations. Variable names in bold italic font indicate environmental variables used in analyses described in Methods S4, S5 and S7.

Base climate variables

Daily mean temperature (*tas*, in °C), minimum temperature (*tasmin*, in °C) and maximum temperature (*tasmax*, in °C) and daily cumulated rainfall (*pr*, in mm) were extracted from EURO4M-MESAN (1989-2010) NetCDF grids (0.05° longitude and latitude resolution) (http://www.euro4m.eu/). Surface incident shortwave solar radiation per day (*sis*, in W m^{-2}) was extracted from EUMETSAT CM SAF (1989-2013) NetCDF grids (0.11° longitude and latitude resolution) (https://www.cmsaf.eu/).

The 365-days year was broken down into 25 year-slices of 14 days and a last one of 15 days. Norms over the 1989-2010 period for *tas*, *tasmin*, *tasmax* and *pr* and over the 1989-2013 period for *sis* were computed for average values per year-slice by CERFACS (http://cerfacs.fr). Norms of potential evapotranspiration (*pet*) were also computed over the 1989-2010 period by CERFACS at the resolution of the 0.05° EURO4M-MESAN NetCDF grid for average value per year-slice using the formula of Turc (1961):

pet = 0,013 x number of days in slice x (*tas*/(*tas*+15)) x (*gR* + 50)

where *gR* = global radiation in Cal cm^{-2} day^{-1}.

From these base climatic variables, we computed climate descriptors at the resolution of the 0.05° EURO4M-MESAN NetCDF grid as described in the next four paragraphs. Values of climate descriptors at sites of origin of perennial ryegrass populations were set as the values of grid cells containing sites of origin of populations.

Seasonal climate descriptors

We used the preceding norms delivered by CERFACS to compute seasonal norms for *tas*, *tasmin*, *tasmax*, *pr*, *sis*, *pet* and *pr-pet*. For this purpose, the 365-days year was broken down into seasons as follows: spring from 26 February to 03 June, summer from 04 June to 09 September, autumn from 10 September to 02 December, winter from 03 December to 25 February. The following descriptors were computed:

- *Spring season descriptors*: ***tas_sp***, ***tasmin_sp***, ***tasmax_sp***, ***pr_sp***, ***sis_sp***, ***pet_sp***, ***pr_pet_sp***
- *Summer season descriptors*: ***tas_su***, ***tasmin_su***, ***tasmax_su***, ***pr_su***, ***sis_su***, ***pet_su***, ***pr_pet_su***
- *Autumn season descriptors*: ***tas_au***, ***tasmin_au***, ***tasmax_au***, ***pr_au***, ***sis_au***, ***pet_au***, ***pr_pet_au***
- *Winter season descriptors*: ***tas_wi***, ***tasmin_wi***, ***tasmax_wi***, ***pr_wi***, ***sis_wi***, ***pet_wi***, ***pr_pet_wi***.

ETCCDI derived indices

The joint CCl/CLIVAR/JCOMM Expert Team (ET) on Climate Change Detection and Indices (ETCCDI) set up a panel of 27 climatic indices dedicated to the analysis of climate extremes in a context of changing climate:

(http://etccdi.pacificclimate.org/

http://www.ecad.eu/documents/WCDMP_72_TD_1500_en_1.pdf).

Using daily values of base climate variables from the EURO4M-MESAN (1989-2010) NetCDF grids, CERFACS computed 1989-2010 norms of these indices for the 26 year-slices we defined. For the present study, we used the following 10 indices:

1. ***fd***, *number of frost days*: count of days when *tasmin* (daily minimum temperature) < 0°C.
2. ***su***, *number of summer days*: count of days when *tasmax* (daily maximum temperature) > 25°C.
3. ***id***, *number of icing days*: count of days when *tasmax* (daily maximum temperature) < 0°C.
4. ***tr***, *number of tropical nights*: count of days when *tasmin* (daily minimum temperature) > 20°C.
6. ***txx***, *maximum value of daily maximum temperature (tasmax)*
9. ***tnn***, *minimum value of daily minimum temperature (tasmin)*
16. ***dtr***, *daily temperature range: mean value of (tasmax-tasmin)*
17. ***rx1day***, *maximum 1-day precipitation*
19. ***sdii***, *simple precipitation intensity index*: mean daily precipitation amount on wet days (*i.e.* on days when precipitation ≥ 1mm)
22. ***r01mm***, *count of days when precipitation ≥ 1mm*

The index numbers are their rank in the list of 27 indices published by the ETCCDI (http://etccdi.pacificclimate.org/list_27_indices.shtml). More information about the computation of these indices can be found following the preceding link. Note that ETCCDI experts propose to compute *fd, su*, *id*, *tr* and *r01mm* on an annual span basis, *txx*, *tnn*, *dtr* and *rx1day* per month and do not propose any span length basis for *sdii*. From the norms computed for these indices for 14/15 days year-slices by CERFACS, we computed *fd*, *su*, *id* and *tr* values on an annual span basis and values of the six other indices per season, with seasons delineated as reported in the preceding paragraph.

BIOCLIM derived variables

We used the norms of *tas*, *tasmin*, *tasmax* and *pr* set up over the 1989-2010 period for year-slices to compute the bioclimatic variables (*bio1* to *bio19*) usually derived in ecological sciences from the *WorldClim* database (http://www.worldclim.org/bioclim). While the computation of these variables from the *WorldClim* data uses norms set up on a monthly time span, our computation used the 14/15 days time span of our 26 year-slices. Consequently, our computation of *bio5*, *bio6*, *bio13* and *bio14* addressed the warmest, coldest, wettest and driest year-slice of 14/15 days, respectively, instead of the warmest, coldest, wettest and driest month when computed from *WorldClim* data. Similarly, our computation of *bio2* addressed the mean of 14/15 days slices instead of the mean of monthly periods.

The BIOCLIM derived variables were thus set as:

- ***bio1***: annual mean temperature
- ***bio2***: mean diurnal range (mean of 14/15 days year-slices (*tasmax* – *tasmin*))
- ***bio3***: isothermality (*bio2*/*bio7*×100)
- ***bio4***: temperature seasonality (standard deviation of average daily mean temperature per year-slice × 100)
- ***bio5***: average daily maximum temperature (*tasmax*) of the warmest 14/15 days year-slice
- ***bio6***: average daily minimum temperature (*tasmin*) of the coldest 14/15 days year-slice
- ***bio7***: temperature annual range (*bio5* – *bio6*)
- ***bio8***: mean temperature of wettest quarter
- ***bio9***: mean temperature of driest quarter
- ***bio10***: mean temperature of warmest quarter
- ***bio11***: mean temperature of coldest quarter
- ***bio12***: annual precipitation
- ***bio13***: precipitation of wettest 14/15 days year-slice
- ***bio14***: precipitation of driest 14/15 days year-slice
- ***bio15.mod***: precipitation seasonality (standard deviation of cumulated precipitation per year-slice × 100, instead of coefficient of variation for the actual BIOCLIM variable *bio15*)
- ***bio16***: precipitation of wettest quarter
- ***bio17***: precipitation of driest quarter
- ***bio18***: precipitation of warmest quarter
- ***bio19***: precipitation of coldest quarter.

Since potential evapotranspiration norms were also available to us, we computed several additional 'BIOCLIM like' variables quantifying seasonal variations of potential evapotranspiration (*pet*) and climatic water balance (precipitation *minus* evapotranspiration or *pr* - *pet*):

- ***bio.ad.20***: (*pr-pet*) of wettest quarter
- ***bio.ad.21***: (*pr-pet*) of driest quarter
- ***bio.ad.22***: (*pr-pet*) of warmest quarter
- ***bio.ad.23***: (*pr-pet*) of coldest quarter
- ***bio.ad.24***: *pet* of wettest quarter
- ***bio.ad.25***: *pet* of driest quarter
- ***bio.ad.26***: *pet* of warmest quarter
- ***bio.ad.27***: *pet* of coldest quarter.

Ecophysiological indices

We used the norms of *tas*, *tasmin*, *tasmax*, *pr*, *sis* and *pet* set up for the 26 year-slices of 14/15 days to compute a panel of ecophysiological variables that may be relevant for the climatic adaptation of natural perennial ryegrass.

We considered that this perennial grass benefits from two periods within the year in which the climatic conditions make possible the biomass growth, one in spring and another one in autumn. Spring growth is expected to start when temperature and incident radiation do not fall anymore below a certain threshold and to be stopped by summer drought. Conversely, autumn growth is expected to start at the end of the summer drought period and to end when temperature and incident radiation begin to fall below a certain threshold. Natural selection is expected to select genotypes with phenological and growth schedules compatible with the position and length of the spring and autumn growth periods within the year. It is however conceivable that adaptive genetic variability exists for the minimum temperature and incident radiation necessary for biomass growth as well as for the minimum soil water content below which plants suffer from drought.

Period of occurrence of sufficient temperature and incident radiation for biomass growth

We considered that the daily minimum temperature should not fall anymore below 0°C (Zaka *et al.*, 2017) and the incident radiation below 60 W m^{-2} (Laboisse, 2018) to enable the start of spring growth. Conversely, we considered that the autumn growth ends when one of these two variables falls below the preceding thresholds.

Thus, we set the three following variables:

- ***dgrb***: first day of the spring growth period (*i.e.* first day of the first 14/15 days year-slice when tasmin > O°C and sis > 60 W m^{-2})
- ***dgre***: last day of the autumn growth period (*i.e.* last day of the last 14/15 days year-slice when tasmin > O°C and sis > 60 W m^{-2})
- ***lmgr***: length of the maximum potential growing period (*dgre* – *dgrb* + 1).

Water-stress variables

We set the following variables:

- swc_j the soil water content at the beginning of the year-slice *j*
- pet_j the potential evapotranspiration during the year-slice *j*
- etr_j the actual evapotranspiration during the year-slice *j*
- swc_{j+1} the soil water content at the beginning of the year-slice *j+1*
- pr_j the cumulated rainfall during the year-slice *j*
- swc_{max} the maximum possible soil water content.

Drought is considered to occur when the soil water content falls to, or below, the permanent wilting point (*i.e.* 0.4 swc_{max}). If the soil water content is higher than the permanent wilting point, the actual evapotranspiration is considered as equal to the potential evapotranspiration; if it is equal to, or lower, than the permanent wilting point, the actual evapotranspiration is reduced in proportion of the ratio 'actual soil water content / soil water content at wilting point'.

Therefore, for the year-slice *j*:

- if $swc_j \geq 0.4\ swc_{max}$, $etr_j = pet_j$
- if $swc_j \leq 0.4\ swc_{max}$, $etr_j = pet_j \times swc_j / (0.4\ swc_{max})$.

then:

- if $swc_j + pr_j - etr_j < swc_{max}$, $swc_{j+1} = swc_j + pr_j - etr_j$
- if $swc_j + pr_j - etr_j \geq swc_{max}$, $swc_{j+1} = swc_{max}$.

We fixed swc_{max} = 150 mm to all grid cells of the 0.05° EURO-4M MESAN NetCDF grid and run the above sequence for each grid cell. We started running the sequence for the 26 year-slices of a first year, setting the soil water content of the first slice in the year equal to the maximum soil water content ($swc_1 = swc_{max}$) at all grid cells. For some grid cells (experiencing long substantial drought periods), the soil water content at the beginning of the first year-slice of the second year (swc_{27}) was less than swc_{max}. We thus continue to run the sequence for additional consecutive years until the soil water content at the beginning of the first year-slice converged to a steady value. Convergence was met after seven or eight consecutive years and we used the *swc* year-slice data of the 11th year to set up the water stress variables to use in data analyses:

- ***dwsb***: first day in the year of the water (drought) stress period when swc $\leq$ 0.4 swc_{max},
- ***dwse***: last day in the year of the water stress period when swc $\leq$ 0.4 swc_{max},
- ***lmws***: length of the water stress period

where *lmws* = *dwse* – *dwsb* +1,

- ***cum_ws***: cumulated water stress across the water stress period

where $cum\_ws = \Sigma_{d=1,W} (swc_d - 0.4\ swc_{max})$
and swc_d is the soil water content at the d^{th} day of the water stress period and *W* is the number of days of the water stress period,

- ***daily_ws***: mean daily water stress across the water stress period

where *daily_ws* = *cum_ws* / *W*.

<u>Definition of the spring and autumn biomass growth periods</u>
The length of the spring growth period was computed as:

lspring = *dwsb* – *dgrb* +1

and that of the autumn biomass growth period as:

laut = *dgre* – *dwse* +1.

For grid cells for which water stress never occurs (*swc* always greater than 0.4 swc_{max}), we looked for the year-slice when swc_j was the lowest, or if swc_j was steady over the maximum potential growth period, for the year-slice when $pr_j - pet_j$ was the lowest. We recorded the median day of this year-slice and used it as a replacement of *dwsb* and *dwse* in the two preceding computations of *lspring* and *laut*.

In addition, we computed sums of growing-degree-days for the spring and autumn growth periods:

stspring = $\Sigma_{d=1,S}\ tas_d$

where tas_d is the mean temperature of the d^{th} day of the spring period and *S* is the number of days of the spring growth period

and

staut = $\Sigma_{d=1,A}\ tas_d$

where tas_d is the mean temperature of the d^{th} day of the autumn period and *A* is the number of days of the autumn growth period.

Note that only $tas_d > 0$ should be taken into account to compute growing-degree-days for a temperate grass. All tas_d values are indeed positive since we considered that *tasmin* should be greater than 0 in the spring and autumn periods.

<u>Heat stress variables</u>
Heat stress was considered as occurring when the daily maximum temperature is higher than 30°C (Zaka *et al.*, 2017). We defined the heat stress period as follows:

dtsb: first day of the heat stress period, *i.e.* first day of the first year-slice when *tasmax* > 30°C
dtse: last day of the heat stress period, *i.e.* last day of the last year-slice when *tasmax* > 30°C
lmts: length of the heat stress period
where *lmts* = *dtse* – *dtsb* +1.

Vernalization period variable

We considered that temperatures favorable for vernalization were those between 0°C and 10°C in autumn and winter seasons (Kleinendorst & Sonneveld, 1930). Thus, we computed the length of the period favorable for vernalization as:

lvr = number of year-slices when 0°C < tas_j < 10°C between 10 September and 25 February

where tas_j is the norm of the average daily mean temperature of the year-slice *j*.

Soil data

We used data made available by the European Soil Data Centre (ESDAC) (http://esdac.jrc.ec.europa.eu).

Data were extracted from the 1 km resolution raster layers of the *European Soil Database Derived Data* (http://esdac.jrc.ec.europa.eu/content/european-soil-database-derived-data) (Hiederer, 2013a,b) for the following variables:

root_soil: depth available to roots (cm)
clay_topsoil: topsoil clay content (%)
clay_subsoil: subsoil clay content (%)
sand_topsoil: topsoil sand content (%)
sand_subsoil: subsoil sand content (%)
silt_topsoil: topsoil silt content (%)
silt_subsoil: subsoil silt content (%)
oc_topsoil: topsoil organic carbon content (%)
oc_subsoil: subsoil organic content (%)
bd_topsoil: topsoil bulk density (g cm^{-3})
bd_subsoil: subsoil bulk density (g cm^{-3})
cf_topsoil: topsoil coarse fragment content (%)
cf_subsoil: subsoil coarse fragment content (%)
tawc_soil: total available water content from Pedo-Transfer-Function (mm).

Topsoil refers to soil above 30cm depth and subsoil to soil below 30cm depth.

Values of soil variables at sites of origin of accessions were set as the values of grid cells containing sites of origin.

pH_soil:
This variable was documented by extracting pHCaCl2 data from the 5km resolution ESDAC quantitative map of estimated soil pH values (http://esdac.jrc.ec.europa.eu/content/soil-ph-europe).

Methods S3: High throughput phenotyping

Experimental design

Three hundred and eighty five perennial ryegrass populations were sown in experimental gardens in three locations: Poel Island (PO) in Germany (53.990, 11.468) on 8th of April 2015, Melle (ME) in Belgium (50.976, 3.780) on 2nd of October 2015 and Lusignan (LU) in France (46.402, 0.082) on 9th of April 2015. 2g m^{-2} seeds of good germination quality (> 80%) is the seed density commonly used to sow dense meadows for forage usage. Because some genebank seed lots were quite old (more than 15 years old), their germination rate might be too low to sow at standard seed density. Therefore, in each of these three locations, each population was sown in three 1m² micro-swards with 2g, 4g or 6g seeds according to whether its previously checked germination rate was higher than 80%, between 80 and 60%, or smaller than 60%, respectively. The 385 populations were sown in three complete blocks (replicates) in each location. Trials were monitored until end of 2017 in PO and ME and until end of 2018 in LU. Micro-swards were cut (all aerial biomass higher than 7 cm above ground surface) regularly as to simulate common cutting regime of meadows used for green forage production or grazing. Cutting dates were 16/06/15, 06/08/15, 04/09/15, 12/10/15, 04/03/16, 01/06/16, 13/07/16, 31/08/16, 26/10/16, 10/03/17, 07/06/17, 19/07/17, 01/09/17, 13/10/17 at PO, 13/05/16, 08/07/16, 29/08/16, 13/10/16, 17/04/17, 31/05/17, 13/07/17, 24/08/17, 04/10/17 at ME and 30/06/15, 03/09/15, 30/10/15, 04/02/16, 08/06/16, 26/07/16, 01/02/17, 13/06/17, 07/09/17, 07/06/18, 27/08/18 at LU. Anti-dicotyledon herbicide was applied once in 2015 in each location and a second time in 2016 or 2017 depending on locations. In each location, a

nitrogen fertilization was applied with 60 kg N ha^{-1} two months after sowing and after each aerial biomass cut and with 80 kg N ha^{-1} after winters 2015-2016 and 2016-2017 at start of spring growth.

Weather conditions experienced at each trial location are displayed per season of each year in Table S6. At LU, drought stress was severe during summers 2016 and 2017. During these summer periods, the average soil water content fell below 20% of the soil water content at field capacity. At PO, drought stress remained small or negligible during both summers 2016 and 2017; however, the winter periods were colder (mean temperature below 3°C) than in the two other locations, especially at the end of the 2015-2016 winter period. At ME, periods of moderate drought stress occurred during summer and autumn periods, notably during summer 2017 when soil water content fell below 27% of the soil water content at field capacity.

Recorded phenotypic traits

Scores or measurements of phenotypic traits were recorded at the level of 1 m² micro-swards over all plants, *i.e.* without phenotyping individual plants within micro-swards. The recorded traits are described in detail hereafter. See also Table S5 for a complete list of phenotypic traits and for the values of these traits per perennial ryegrass population. Variable names in bold italic font indicate phenotypic variables used in analyses described in Methods S6 and Methods S7.

Traits related to vigor after sowing:

Days from sowing to emergence: Number of days between the sowing of the plot and the start of emergence. Recorded only at PO in 2015 (***DES_po15***)

Vigor after sowing: Visual score performed two months after sowing on a 1 (small and/or necrotized and/or chlorotic plants) to 9 (strong green plants, very good size, difficult to pull out) scale. Recorded in 2015 at the three trial locations (***VAS-lu15***, ***VAS_me15*** and ***VAS_po15***).

Regularity after sowing: Visual score performed two months after sowing and proportional to the percentage of emerged plants on a 1 (no emergence) to 9 (full stand, 100% plants emerged) scale. Recorded in 2015 at the three locations (***RAS_lu15***, ***RAS_me15*** and ***RAS_po15***).

Morphology of plants and sward density

Leaf lamina width: Visual score performed during spring growth before fertile stem elongation on a 1 (very thin) to 9 (very large) scale. Recorded at PO in 2016 (***LMW_po16***), ME in 2016 (***LMW_me16***) and LU in 2017 (**LMW_lu17**).

Growth habit: Visual score recorded on a 1 (most erect) to 9 (most prostrate) scale in 2016 at LU and PO. Because of high correlation between the two locations, only average values of populations over locations (***GRH_avg***) were used in data analyses (see later '*Computation of mean values and elaborate variables*').

Sward density: Visual score recorded in the early spring to assess density of vegetative tillers on a 1 (very poor density, few tillers, ground largely visible) to 9 (very dense, many tillers, no visible ground) scale. Recorded in April 2017 at LU (***DVG_04_lu17***).

Traits related to phenology

Proportion of plants heading in first year: Visual score reporting the density of elongated fertile (bearing spike) stems the year of sowing (*i.e.* without vernalization) on a 1 (no fertile stem) to 9 (100% plants with fertile stems) scale. Recorded in 2015 at LU (***HFY_lu15***) and PO (***HFY_po15***).

Heading (or spike emergence) date: In spring after a vernalization period, date when at least 20 spikes are arising at the top of tiller sheath in a micro-sward. This date was converted into growing-degree-days (base 0) starting from the first day when daily minimum temperature and incident shortwave global radiation do not fall anymore below 0°C and 60 W m^{-2}, respectively (*i.e.* from the start of vegetative spring growth, see Methods S2 - Ecophysiological indices). Recorded in 2016 at LU and PO (***HEA_lu16*** and ***HEA_po16***) and in 2017 at LU and PO (***HEA_lu17*** and ***HEA_po17***). Note that these four heading dates were highly correlated (correlations higher than 0.90).

Aftermath heading: After the cut of the first spring wave of elongated fertile stems, visual score reporting the intensity of afterwards recurring fertile stem elongation on a 1 (no fertile stem) to 9 (100% plants with fertile

stems) scale. Recorded in 2016 and 2017 at LU, ME and PO (***AHD_lu16***, ***AHD_lu17***, ***AHD_me16***, ***AHD_me17***, ***AHD_po16*** and ***AHD_po17***).

Investment in sexual reproduction

Density of elongated fertile stems: Visual score recorded for the first spring wave of elongated fertile stems on a 1 (no fertile stem) to 9 (maximum density of elongating fertile stems) scale. Recorded in 2017 at LU (***DST_lu17***) and PO (***DST_po17***).

Straw height: Length (in cm) of one average elongated fertile stem per micro-sward (from ground to base of spike). Recorded at LU in 2017 (***HST_lu17***).

Spike length (LSP): Length in mm of a single average spike per micro-sward. Recorded at LU in 2017 (***LSP_lu17***).

Spikelet length (LSL): Length in mm of a spikelet from a single average spike per micro-sward. Recorded at LU in 2017 (***LSL_lu17***).

Spikelet count (NSL): Number of spikelets from a single average spike per micro-sward. Recorded at LU in 2017 (***NSL_lu17***).

Dynamics of vegetative spring growth

Vegetative spring growth was monitored in 2016 and 2017 in each of the three trial locations from the start of spring growth to a couple of weeks before spike emergence (heading) in ME and to a couple of weeks after spike emergence in LU and PO (the trial was cut before spike emergence in ME to collect samples for a biochemical analysis of vegetative aerial biomass, see later 'Biochemistry of aerial biomass'). The monitoring was carried out by measuring the micro-sward canopy height once a week. To measure the canopy height, a 'herbometre®' tool (ARVALIS) was used. It is a rule along which a plate runs; the plate was let leaning on the canopy, except at PO in 2016 where the plate was maintained at the top of natural canopy height, and the canopy height was measured (in mm) as the distance from the ground to the plate (Powell, 1974). At each measurement date, the canopy height of a given micro-sward was measured twice at two different positions in the micro-sward. A Schnute growth model (Schnute, 1981) was afterwards fitted to model the spring canopy height growth of each population as a function of growing-degree-days (base 0) for each combination of location and year using the six observations available for a population at each measurement date. The growth curves were fitted in the time interval running from the start of spring growth (when daily minimum temperature and incident shortwave global radiation do not fall anymore below 0°C and 60 W m^{-2}, respectively, see Methods S2 - Ecophysiological indices) to spike emergence date of populations at LU and PO and to the date of end of weekly measurements at ME. These models were used to predict the canopy heights of populations at several thermal time dates (one predicted value per population for each combination of location and year). The predicted canopy heights were the followings:

- Canopy height at 300 growing-degree-days after the start of spring growth in 2016 and 2017 at LU, PO and ME (***CHs300_lu16***, ***CHs300_lu17***, ***CHs300_po16***, ***CHs300_po17***, ***CHs300_me16*** and ***CHs300_me17***)
- Canopy height at 500 growing-degree-days after the start of spring growth in 2016 and 2017 at LU and ME, and in 2017 at PO (***CHs500_lu16***, ***CHs500_lu17***, ***CHs500_po17***, ***CHs500_me16*** and ***CHs500_me17***)
- Canopy height at 300 growing-degree-days before spike emergence (heading) date in 2016 and 2017 at LU and PO (***CH300h_lu16***, ***CH300h_lu17***, ***CH300h_po16*** and ***CH300h_po17***)
- Canopy height at 400 growing-degree-days before spike emergence (heading) date in 2016 and 2017 at LU and PO (***CH400h_lu16***, ***CH400h_lu17***, ***CH400h_po16*** and ***CH400h_po17***)

Summer and autumn growth variables

Summer canopy height: Canopy height (in mm) after a summer period measured with the herbometre as previously described for the monitoring of the dynamics of spring growth. Recorded in 2016 at LU and ME (***SMH_lu16*** and ***SMH_me16***) and in 2017 at ME and PO (***SMH_me17*** and ***SMH_po17***).

Summer growth rate: The preceding variable divided by the number of growing-degree-days between the date of cut preceding the canopy height measurement and the date of the canopy height measurement. Computed for the same combinations of locations and years as *summer canopy height* (***SGR_lu16***, ***SGR_me16***, ***SGR_me17*** and ***SGR_po17***).

Autumn canopy height (AMH): Canopy height (in mm) after a period of autumn growth measured with the herbometre as previously described for the monitoring of the dynamics of spring growth. Recorded in 2017 at ME and PO (***AMH_me17*** and ***AMH_po17***).

Autumn growth rate (AGR): The preceding variable divided by the number of growing-degree-days between the date of cut preceding the canopy height measurement and the date of the canopy height measurement. Computed for the same combinations of locations and year as *autumn canopy height* (***AGR_me17*** and ***AGR_po17***).

Dynamics of regrowth after cutting

Vigour after cutting: Recovery after cutting scored five to ten days after the cut on a visual scale proportional to the aerial biomass regrowth from 1 (no regrowth) to 9 (strongest regrowth). This trait was recorded at LU on 21/06/2016, 21/09/2016, 05/07/2017 and 26/07/2017 and mean values per year were considered (***VAC_lu16*** and ***VAC_lu17***). It was recorded at PO on 27/06/2017 (***VAC_po17***).

Abiotic stresses

Drought stress symptoms: Visual score of susceptibility to drought recorded at the end of a drought period and reporting the percentage of sere leaves on a 1 (no damage, green growing plants) to 9 (all plants with completely sere foliage) scale. Recorded in 2016 at LU and PO (***DRO_lu16*** and ***DRO_po16***).

Winter damage: Visual score of damage on plants at the end of the winter period on a 1 (no damage, green plants) to 9 (all plants with necrotized foliage) scale. Recorded at PO in 2016 and 2017 (***WID_po16*** and ***WID_po17***).

Biotic stresses - disease damages

Drechslera spp. (Syn. Helminthosporium spp.) *susceptibility*: Visual score of *Drechslera* damages recorded after occurrence of the disease and reporting the proportion of foliage affected on a 1 (no symptoms) to 9 (all plants with highly affected foliage) scale. Recorded at LU in January and July 2016 (***DHE_01_lu16*** and ***DHE_07_lu16***) and in April 2017 (***DHE_04_lu17***).

Black rust (Puccinia graminis) susceptibility: Visual score of black rust damages recorded after occurrence of the disease and reporting the proportion of foliage affected on a 1 (no symptoms) to 9 (all plants with highly affected foliage) scale. Recorded at LU in 2015 and 2016. Average values of populations over record dates were computed and used in data analyses (***DRB_lu1516***).

Susceptibility to indeterminate diseases: Visual score of damages caused by indeterminate diseases or mixtures of diseases recorded after their occurrence and reporting the proportion of foliage affected on a 1 (no symptoms) to 9 (all plants with highly affected foliage) scale. Recorded at LU in 2015, 2016 and 2017 (***DIS_lu15***, ***DIS_lu16*** and ***DIS_lu17***), at ME in 2016 and 2017 (***DIS_me16*** and ***DIS_me17***) and at PO in 2015 and 2017 (***DIS_po15*** and ***DIS_po17***).

Dynamics of persistency over successive trial years

A visual score of soil coverage by plants was recorded in each location four months after sowing and then every couple of months. It was proportional to the soil surface covered by plant material and was scored on a 1 (no living plants on the micro-sward plot) to 9 (best soil coverage, *i.e.* micro-sward perfectly filled with strong living plants) scale. To assess the dynamics of change in population persistency between two dates of record of soil coverage in a given location, we computed the difference between the soil coverage record of micro-swards at the late date and at the early date (late date record – early date record). This difference was considered as an indicator of population persistency over the targeted time span.

Such differences were computed throughout targeted periods at the three trial locations as follows:

- Throughout summers 2015, 2016 and 2017 at LU (***SCD_su15_lu***, ***SCD_su16_lu*** and ***SCD_su17_lu***)
- Throughout summer 2017 at ME (***SCD_su17_me***)
- Throughout winters 2015-16, 2016-17 and 2017-18 at LU (***SCD_wi1516_lu***, ***SCD_wi1617_lu*** and ***SCD_wi1718_lu***)
- Throughout winter 2016-17 at ME (***SCD_wi1617_me***)
- Throughout winters 2015-16 and 2016-17 at PO (***SCD_wi1516_po*** and ***SCD_wi1617_po***)

- From end of spring 2015 to end of autumn 2018 at LU (***SCD_15to18_lu***)
- From end of spring 2015 to end of autumn 2017 at PO (***SCD_15to17_po***)
- From end of spring 2016 to end of autumn 2017 at ME (***SCD_16to17_me***)

Biochemistry of aerial biomass

At the April and October cuts in 2017 at ME, fresh samples of aerial biomass were collected, dried down at 60°C for 72 h and ground to pass a 1 mm sieve. Ground samples were analyzed by Near Infrared Reflectance Spectroscopy (NIRS) at ILVO to predict the following biochemical composition variables:

- *Acid Detergent Lignin* (ADL), *Acid detergent fiber* (ADF) and *neutral detergent fiber* (NDF) in dry matter (% DM) (Van Soest *et al.*, 1991) (***ADL_04_me17*** and ***ADL_10_me17, ADF_04_me17*** and ***ADF_10_me17***, ***NDF_04_me17*** and ***NDF_10_me17***)
- *Crude protein* content (ISO derived method 5983-2) in % DM (***PRT_04_me17*** and ***PRT_10_me17***)
- *Water-soluble-carbohydrate* content in % DM (Wiseman *et al.*, 1960) (***WSC_04_me17*** and ***WSC_10_me17***).
- *Organic matter digestibility* (De Boever *et al.*, 1988) (***OMD_04_me17*** and ***OMD_10_me17***)
- *In vitro neutral detergent fibre degradability* (DNDF) as per Dolstra and Medema (1990) (***DNDF_04_me17*** and ***DNDF_10_me17***).

Leaf lamina traits

Nitrogen content in leaf lamina was previously demonstrated to report for the fulfilment of plant nitrogen supply for optimal growth (Lemaire *et al.*, 1989; Farruggia *et al.*, 2004). Isotopic discrimination of ^{13}C ($\delta^{13}C$) is considered as a marker of water use efficiency (Condon *et al.*, 2002; Durand, 2007) and possibly of photosynthetic efficiency when water supply is not limiting (Condon *et al.*, 2007). A leaf lamina sample was collected on 30 plants per micro-sward at LU on April 2016. The 30 samples per micro-sward were pooled to a single batch which was dried down, ground and analyzed with a mass spectrometry tool (Flash 2000 Thermo-Fischer) at INRA to predict the nitrogen content of the dry leaf lamina tissue (%) (***NLI_lu16***) and the isotopic signature of ^{13}C (‰) (***^{13}C_lu16***).

Computation of mean values and elaborate phenotypic variables

Computation of mean values of populations

Models of analysis of variance (ANOVA) were used to check the accuracy of collected raw data and the significance of the population effect and to compute adjusted means of populations. Analyses of variance were performed using the R functions *lm* and *Anova* of the R *car* library.

For a single combination of one location and one trait, raw data of a given trait was analyzed using the following fixed effect model:

$y_{ij} = \mu + g_i + b_j + e_{ij}$ (1)

where y_{ij} is the observed value of population *i* in the complete block *j*, g_i is the (genetic) effect of population *i*, b_j is the effect of the complete block *j* and e_{ij} is the residual effect of the model.

F statistics of the population effect were highly significant (*p-value* < 0.005) for all combinations of traits, locations and record dates, which indicated a satisfactory accuracy of collected raw data. Adjusted means were computed using the *emmeans()* function of the *emmeans R* library. These adjusted means were used as population values for downstream analyses.

For traits for which it was relevant, an analysis of variance was also performed across tested combinations of locations and dates of record (*i.e.* environments) using the following model:

$y_{ijr} = \mu + g_i + env_j + g*env_{ij} + b/env_{jr} + e_{ijr}$ (2)

where y_{ijr} is the observed value of population *i* within the complete block *r* of environment *j*, g_i is the (genetic) effect of population *i*, env_j is the effect of environment *j*, $g*env_{ij}$ is the interaction between population *i* and environment *j*, b/env_{jr} is the effect of complete block *r* nested within environment *j* and e_{ijr} is the residual of the model.

Setting g_i as fixed effect and the other effects as random (mixed model), the analysis resulted in a highly significant F statistics (*p-value* < 0.005) for all traits. Additionally, setting all effects as fixed, the env_j effect and $g*env_{ij}$ were significant (*p-value* < 0.05) for all traits. These results made relevant to compute both values per

environment (adjusted means from model 1) and means of populations across environments (adjusted means from model 2) for downstream analyses. The mean of a trait across environments was identified in downstream analyses by adding to the acronym of the trait the suffix *avg* (***ADF_avg***, ***ADL_avg***, ***AHD_avg, DIS_avg, DNDF_avg***, ***DRO_avg, DST_avg***, ***GRH_avg***, ***HEA_avg, HFY_avg***, ***LMW_avg NDF_avg***, ***OMD_avg***, ***PRT_avg***, ***RAS_avg VAC_avg***, ***VAS_avg, WSC_avg***).

Computation of elaborate variables

Canopy heights in spring and density of elongated fertile stems appeared correlated to heading date whereas some variability of these traits was obvious at same heading earliness. To take into account this trend, we regressed the mean value of populations for these traits in each environmental condition on the mean value of populations for heading date averaged over environments (***HEA_avg***).

The Schnute model-predicted canopy heights were thus regressed on average heading date and regression residuals were kept as additional variables for downstream analyses:

- ***resCHLs300_lu16***, ***resCHLs300_lu17***, ***resCHLs300_po16***, ***resCHLs300_po17***, ***resCHLs300_me16 and resCHLs300_me17***
- ***resCHLs500_lu16***, ***resCHLs500_lu17***, ***resCHLs500_po17***, ***resCHLs500_me16 and resCHLs500_me17***
- ***resCHL300h_lu16***, ***resCHL300h_lu17***, ***resCHL300h_po16 and resCHL300h_po17***
- ***resCHL400h_lu16***, ***resCHL400h_lu17***, ***resCHL400h_po16 and resCHL400h_po17***

Similarly, mean value of populations for density of elongated fertile stems at LU was regressed on average heading date and the residual was kept as additional variable for downstream analyses (***resDST_lu17***).

Methods S4: GEA linear mixed models

We implemented the linear mixed model (3) to assess the association between environmental variables and outlier loci. This model used the environmental variable as the predictor and the population alternative allele frequencies of an outlier locus as the response variable.

$$y_e = Zg + L_l\, v_{el} + \epsilon \qquad (3)$$

In this model, y_e is the *n* length vector of population values for the environmental variable *e* (scaled data [0,1]). Z is an incidence matrix and the *n* length vector g models the genetic background of each population as a random effect with Var[g] = Qσ^2, where Q is the kinship matrix calculated from GBS and HiPlex alternative allele frequencies following Endelman and Jannink (2012). L_l is the *n* length vector of alternative allele frequencies at locus *l* for the *n* populations. The coefficient v_{el} is the additive fixed effect of the locus *l* on the environmental variable *e*. ϵ is the vector of residuals. Associations were considered significant using the liberal threshold FDR = 0.2 and the more conservative one FDR = 0.1.

Methods S5: GWAS linear mixed models

The linear mixed model (4) was used to assess individual locus effect on a given phenotypic trait.

$$y_t = Zg + L_l\, v_{tl} + \epsilon \qquad (4)$$

In this model, y_t is the *n* length vector of population adjusted mean values for trait *t*. Z is an incidence matrix. The variable g models the genetic background of each population as a random effect with Var[g] = Qσ^2 where Q is the kinship matrix calculated as above. L_l is the *n* length vector of alternative allele frequencies at locus *l* for the *n* populations. The coefficient v_{tl} is the additive fixed effect of the locus *l* on trait *t*. ϵ is the vector of residuals. Associations were considered significant using the liberal threshold FDR = 0.2 and the more conservative one FDR = 0.1.

Methods S6: CANCOR test

Our approach was inspired by the redundancy analysis (RDA) used to detect adaptive loci (Forester *et al.*, 2015, 2018). Here, we performed the following two-step analysis. In a first step, the alternative allele frequency of each locus was modelled as a linear function of each single environmental variable on the one hand and of each single phenotypic trait on the other hand by implementing the following regression models:

$$y_l = E_i\, \varphi_{li} + \epsilon \qquad (5)$$

$$y_l = P_t\, \psi_{lt} + \epsilon \qquad (6)$$

In model (5), y_l is the *n* length vector of population alternative allele frequencies for the candidate locus *l*, E_i is the *n* length vector containing the values of the environmental variable *i* for the *n* populations and φ_{li} is the additive fixed effect (regression slope) of the environmental variable *i* on allele frequency at locus *l*. In model (6), P_t is the *n* length vector containing the values of the phenotypic trait *t* for the *n* populations and ψ_{lt} is the additive fixed effect (regression slope) of the phenotypic trait *t* on allele frequency at locus *l*. ϵ is the vector of residuals in each model. Then, we set up a (*l, i*) table Y containing the regression slopes φ_{li} and a (*l*, *t*) table X containing the regression slopes ψ_{lt}. These two tables were used as input data to perform a Canonical Correlation Analysis (CCorA) (Hotelling, 2006) (see Fig. 2b, main text).

After performing the CCorA analysis, we followed the method of Luu et al. (2017) and Capblancq *et al.* (2018) to detect outlier loci. First, we recovered the loci loadings supplied by the 'Cx' element (phenotype effect matrix) of the CCorA function of the R package 'vegan'. A Mahalanobis distance D was then computed between each locus position and the barycentre of all loci in the space defined by the CCorA axes and loci showing extreme D values were considered as outliers. Let K be the number of CCorA dimensions to retain in the analysis. Mahalanobis distances have a chi-squared distribution with K degrees of freedom after correcting for the genomic inflation factor (Luu *et al.*, 2017; Capblancq *et al.*, 2018). We applied a chi-squared test (CANCOR test) to detect significant outlier loci. The number K and the threshold values of the minimum average minor allele frequency (min MAF) used to select the genotype matrix were determined by comparing the histogram of p-values for alternative values of K and min MAF, with the expectation that the distribution of p-values should be flat with enrichment only for the low ones (François *et al.*, 2016). The CANCOR test resulted in an excess of high p-values with K = 4 and 8, and min MAF = 0 and 0.05. In contrast, with K = 2 and min MAF = 0.1, the distribution of p-values was correct (Fig. S2). P-values obtained with the selected values of K and min MAF were adjusted for the FDR. A locus was considered as outlier at FDR = 0.1.

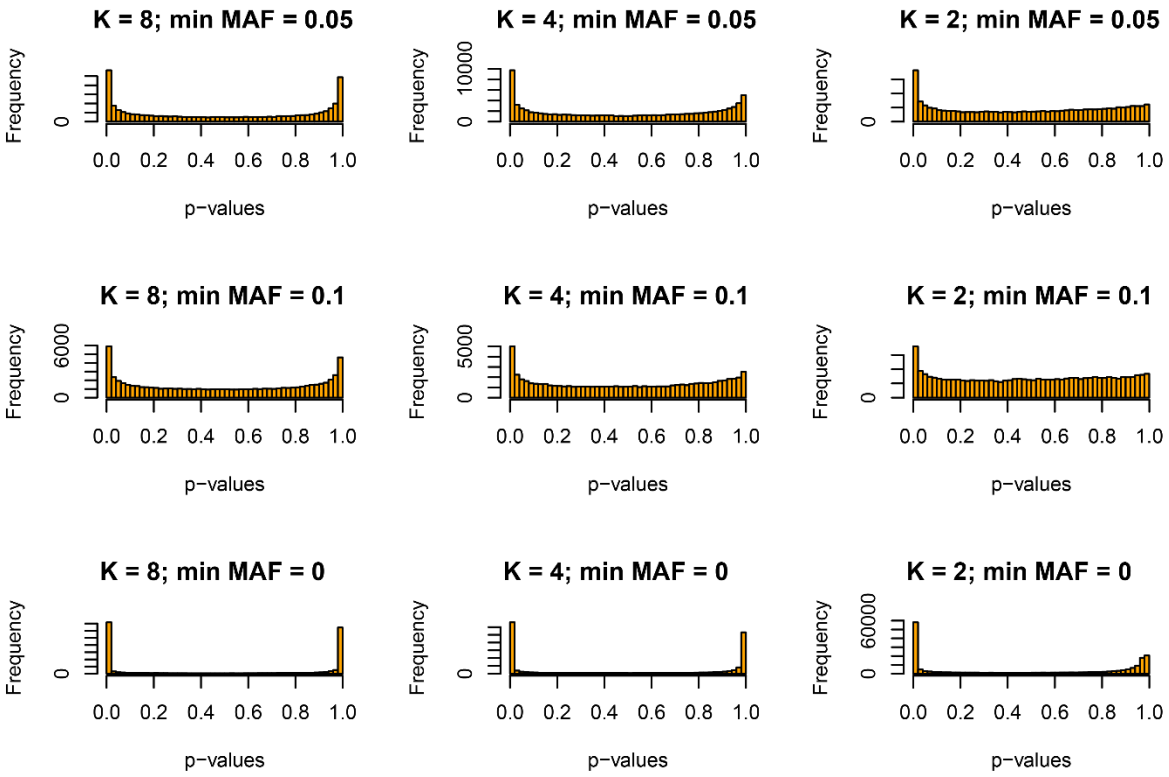

Fig. S1 – Distribution of p-values for alternative numbers of K dimensions and threshold values of the minimum average minor allele frequency (min MAF) in the CANCOR test.

Results S1: Outlier loci detected by the CANCOR and the GEA-GWAS approaches

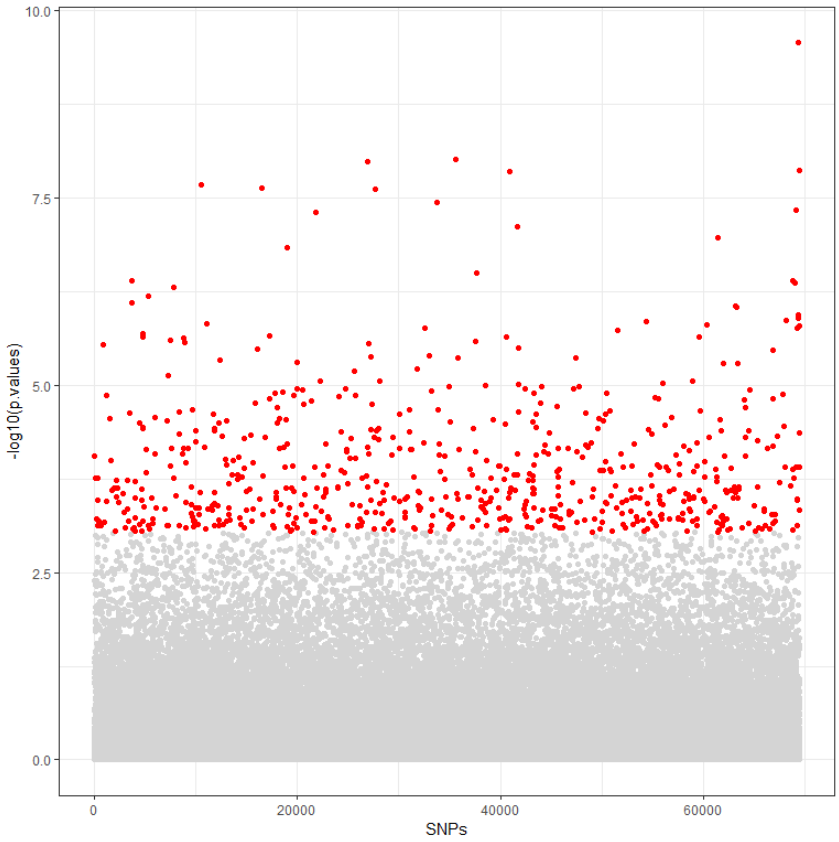

Fig. S2 – CANCOR test. Red points represent outlier loci passing the threshold at FDR = 0.1.

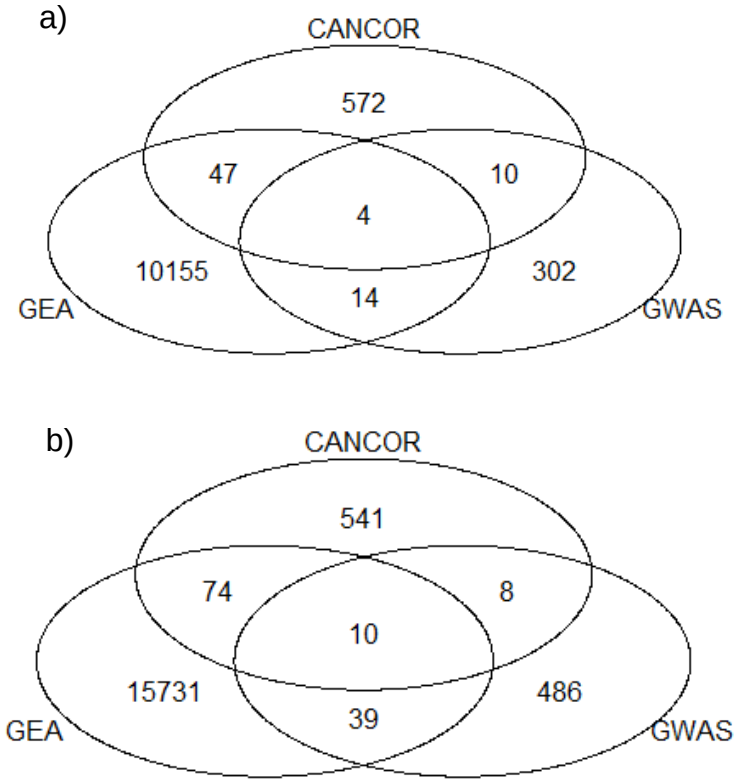

Fig. S3 – Venn diagram displaying the numbers of outlier loci detected by all analyses used in the present study. a) GEA-GWAS at FDR = 0.1 and CANCOR at FDR = 0.1, b) GEA-GWAS at FDR = 0.2 and CANCOR at FDR = 0.1.

Supplemental references

Capblancq T, Luu K, Blum MGB, Bazin E. **2018**. Evaluation of redundancy analysis to identify signatures of local adaptation. *Molecular Ecology Resources* **18**: 1223–1233.

Condon AG, Reynolds M, Rebetzke G, Van Ginkel M, Richards R, Farquhar G. **2007**. Using stomatal aperture-related traits to select for high yield potential in bread wheat. In: Wheat Production in Stressed Environments. Springer, 617–624.

Condon AG, Richards RA, Rebetzke GJ, Farquhar GD. **2002**. Improving intrinsic water-use efficiency and crop yield. *Crop Science* **42**: 122.

De Boever JL, Cottyn BG, Andries JI, Buysse FX, Vanacker JM. **1988**. The use of a cellulase technique to predict digestibility, metabolizable and net energy of forages. *Animal Feed Science and Technology* **19**: 247–260.

Dolstra O, Medema JH. **1990**. An effective screening method for genetic improvement of cell-wall digestibility in forage maize. In: Proceedings 15th congress maize and sorghum section of Eucarpia. 4–8.

Durand JL. **2007**. Les effets du déficit hydrique sur la plante: aspects physiologiques. *Fourrages* **190**: 181–195.

Endelman JB, Jannink J-L. **2012**. Shrinkage estimation of the realized relationship matrix. *G3: Genes, Genomes, Genetics* **2**: 1405–1413.

Farruggia A, Gastal F, Scholefield D. **2004**. Assessment of the nitrogen status of grassland. *Grass and Forage Science* **59**: 113–120.

Forester BR, Jones MR, Joost S, Landguth EL, Lasky JR. **2015**. Detecting spatial genetic signatures of local adaptation in heterogeneous landscapes. *Molecular Ecology* **25**: 104–120.

Forester BR, Lasky JR, Wagner HH, Urban DL. **2018**. Comparing methods for detecting multilocus adaptation with multivariate genotype–environment associations. *Molecular Ecology* **27**: 2215–2233.

François O, Martins H, Caye K, Schoville SD. **2016**. Controlling false discoveries in genome scans for selection. *Molecular Ecology* **25**: 454–469.

Hiederer R. **2013a**. Mapping soil properties for Europe—spatial representation of soil database attributes. *Publications Office of the European Union, EUR26082EN Scientific and Technical Research series. Luxembourg*.

Hiederer R. **2013b**. Mapping soil typologies—spatial decision support applied to European Soil Database. *EUR25932EN Scientific and Technical Research Series*.

Hotelling H. **2006**. Relations between two sets of variates. *Biometrika* **28**: 321.

Kleinendorst A, Sonneveld A. **1930**. Influence of light intensity and temperature before and after attaining the reproductive phase on the behaviour of perennial ryegrass *Lolium perenne* L.). *JAARBOEK 1966*: 19.

Laboisse S. **2018**. Analyse et modélisation des déterminants écophysiologiques intervenant dans la croissance des Poacées dans un milieu artificiel : l'exemple du stade de football. Doctoral Thesis. École doctorale Sciences pour l'environnement - Gay Lussac (La Rochelle) and Unité de recherche pluridisciplinaire prairies et plantes fourragères - (URP3F) INRA Nouvelle-Aquitaine-Poitiers.

Lemaire G, Gastal F, Salette J. **1989**. Analysis of the effect of N nutrition on dry matter yield of a sward by reference to potential yield and optimum N content. In: Proceedings of the XVI International Grassland Congress, 4–11.

Luu K, Bazin E, Blum MGB. **2017**. pcadapt: an R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources* **17**: 67–77.

Murray MG, Thompson WF. **1980**. Rapid isolation of high molecular weight plant DNA. *Nucleic acids research* **8**: 4321–4326.

Powell TL. **1974**. Evaluation of weighted disc meter for pasture yield estimation on intensively stocked dairy pasture. *New Zealand Journal of experimental agriculture* **2**: 237–241.

Schnute J. **1981**. A versatile growth model with statistically stable parameters. *Canadian Journal of Fisheries and Aquatic Sciences* **38**: 1128–1140.

Turc L. **1961**. Evaluation des besoins en eau d'irrigation, évapotranspiration potentielle. *Annales Agronomiques* **12**: 13–49.

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG. **2012**. Primer3—new capabilities and interfaces. *Nucleic acids research* **40**: e115–e115.

Van Soest P van, Robertson JB, Lewis BA. **1991**. Methods for dietary fiber, neutral detergent fiber, and nonstarch polysaccharides in relation to animal nutrition. *Journal of dairy science* **74**: 3583–3597.

Veeckman E, Van Glabeke S, Haegeman A, Muylle H, Van Parijs FR, Byrne SL, Asp T, Studer B, Rohde A, Roldán-Ruiz I. **2018**. Overcoming challenges in variant calling: exploring sequence diversity in candidate genes for plant development in perennial ryegrass (*Lolium perenne*). *DNA Research* **26**: 1–12.

Wiseman HG, Mallack JC, Jacobson WC. **1960**. Silage analysis, determination of sugar in silages and forages. *Journal of Agricultural and Food Chemistry* **8**: 78–80.

Zaka S, Ahmed LQ, Escobar-Gutiérrez AJ, Gastal F, Julier B, Louarn G. **2017**. How variable are non-linear developmental responses to temperature in two perennial forage species? *Agricultural and Forest Meteorology* **232**: 433–442.

Zhang J, Kobert K, Flouri T, Stamatakis A. **2013**. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* **30**: 614–620.

R packages

- base (R Core Team 2018)
- ggtree (Yu et al. 2017)
- doSNOW (Corporation and Weston 2017)
- snow (Tierney et al. 2018)
- doParallel (Corporation and Weston 2018)
- iterators (Analytics and Weston 2018)
- foreach (Microsoft and Weston 2017)
- sjPlot (Lüdecke 2018)
- splitstackshape (Mahto 2018)
- rgdal (R. Bivand, Keitt, and Rowlingson 2018)
- adespatial (Dray et al. 2018)
- spDataLarge (Nowosad and Lovelace 2018)
- spdep (R. Bivand and Wong 2018)
- spData (R. Bivand, Nowosad, and Lovelace 2018)
- Matrix (Bates and Maechler 2018)
- pcadapt (Luu, Blum, and Privé 2018)
- raster (Hijmans 2018)

- sp (Pebesma and Bivand 2005)
- varhandle (Mahmoudian 2018)
- rrBLUP (Endelman 2011)
- dummies (Brown 2012)
- psych (Revelle 2018)
- corrplot (**???**)
- tidyr (Wickham and Henry 2018)
- purrr (Henry and Wickham 2018)
- plyr (Wickham 2011)
- VennDiagram (Chen 2018)
- futile.logger (Rowe 2016)
- gplots (Warnes et al. 2016)
- stringr (Wickham 2019)
- qdap (Rinker 2019)
- RColorBrewer (Neuwirth 2014)
- qdapTools (Rinker 2015)
- qdapRegex (Rinker 2017)
- qdapDictionaries (Rinker 2013)
- ggplot2 (Wickham 2016)
- qvalue (Andrew J. Bass, Dabney, and Robinson 2015)
- robust (Wang et al. 2017)
- fit.models (Konis. 2017)
- reshape2 (Wickham 2007)
- dplyr (Wickham et al. 2018)
- bayou (Uyeda, Eastman, and Harmon 2018)
- coda (Plummer et al. 2006)
- phytools (Revell 2012)
- maps (Richard A. Becker, Ray Brownrigg. Enhancements by Thomas P Minka, and Deckmyn. 2018)
- geiger (Alfaro et al. 2009)
- ape (Paradis and Schliep 2018)
- vegan (Oksanen et al. 2018)
- lattice (Sarkar 2008)
- permute (Simpson 2016)
- grateful (Rodriguez-Sanchez 2018)
- knitr (Xie 2018)

R packages references

Alfaro, ME, F Santini, C Brock, H Alamillo, A Dornburg, DL Rabosky, G Carnevale, and LJ Harmon. 2009. "Nine Exceptional Radiations Plus High Turnover Explain Species Diversity in Jawed Vertebrates." Proceedings of the National Academy of Sciences of the United States of America 106: 13410–4.

Analytics, Revolution, and Steve Weston. 2018. Iterators: Provides Iterator Construct for R. https://CRAN.R-project.org/package=iterators.

Andrew J. Bass, John D. Storey with contributions from, Alan Dabney, and David Robinson. 2015. Qvalue: Q-Value Estimation for False Discovery Rate Control. http://github.com/jdstorey/qvalue.

Bates, Douglas, and Martin Maechler. 2018. Matrix: Sparse and Dense Matrix Classes and Methods. https://CRAN.R-project.org/package=Matrix.

Bivand, Roger, and David W. S. Wong. 2018. "Comparing Implementations of Global and Local Indicators of Spatial Association." TEST 27 (3): 716–48. https://doi.org/10.1007/s11749-018-0599-x.

Bivand, Roger, Tim Keitt, and Barry Rowlingson. 2018. Rgdal: Bindings for the 'Geospatial' Data Abstraction Library. https://CRAN.R-project.org/package=rgdal.

Bivand, Roger, Jakub Nowosad, and Robin Lovelace. 2018. SpData: Datasets for Spatial Analysis. https://CRAN.R-project.org/package=spData.

Brown, Christopher. 2012. Dummies: Create Dummy/Indicator Variables Flexibly and Efficiently. https://CRAN.R-project.org/package=dummies.

Chen, Hanbo. 2018. VennDiagram: Generate High-Resolution Venn and Euler Plots. https://CRAN.R-project.org/package=VennDiagram.

Corporation, Microsoft, and Stephen Weston. 2017. DoSNOW: Foreach Parallel Adaptor for the 'Snow' Package. https://CRAN.R-project.org/package=doSNOW.

Corporation, Microsoft, and Steve Weston. 2018. DoParallel: Foreach Parallel Adaptor for the 'Parallel' Package. https://CRAN.R-project.org/package=doParallel.

Dray, Stéphane, David Bauman, Guillaume Blanchet, Daniel Borcard, Sylvie Clappe, Guillaume Guenard, Thibaut Jombart, et al. 2018. Adespatial: Multivariate Multiscale Spatial Analysis. https://CRAN.R-project.org/package=adespatial.

Endelman, J. B. 2011. "Ridge Regression and Other Kernels for Genomic Selection with R Package rrBLUP." Plant Genome 4: 250–55.

Henry, Lionel, and Hadley Wickham. 2018. Purrr: Functional Programming Tools. https://CRAN.R-project.org/package=purrr.

Hijmans, Robert J. 2018. Raster: Geographic Data Analysis and Modeling. https://CRAN.R-project.org/package=raster.

Konis., Kjell. 2017. Fit.models: Compare Fitted Models. https://CRAN.R-project.org/package=fit.models.

Luu, Keurcien, Michael Blum, and Florian Privé. 2018. Pcadapt: Fast Principal Component Analysis for Outlier Detection.

Lüdecke, Daniel. 2018. SjPlot: Data Visualization for Statistics in Social Science. doi:10.5281/zenodo.1308157.

Mahmoudian, Mehrad. 2018. Varhandle: Functions for Robust Variable Handling. https://CRAN.R-project.org/package=varhandle.

Mahto, Ananda. 2018. Splitstackshape: Stack and Reshape Datasets After Splitting Concatenated Values. https://CRAN.R-project.org/package=splitstackshape.

Microsoft, and Steve Weston. 2017. Foreach: Provides Foreach Looping Construct for R. https://CRAN.R-project.org/package=foreach.

Neuwirth, Erich. 2014. RColorBrewer: ColorBrewer Palettes. https://CRAN.R-project.org/package=RColorBrewer.

Nowosad, Jakub, and Robin Lovelace. 2018. SpDataLarge: Large Datasets for Spatial Analysis. https://github.com/Nowosad/spDataLarge.

Oksanen, Jari, F. Guillaume Blanchet, Michael Friendly, Roeland Kindt, Pierre Legendre, Dan McGlinn, Peter R. Minchin, et al. 2018. Vegan: Community Ecology Package. https://CRAN.R-project.org/package=vegan.

Paradis, E., and K. Schliep. 2018. "Ape 5.0: An Environment for Modern Phylogenetics and Evolutionary Analyses in R." Bioinformatics xx: xxx–xxx.

Pebesma, Edzer J., and Roger S. Bivand. 2005. “Classes and Methods for Spatial Data in R.” R News 5 (2): 9–13. https://CRAN.R-project.org/doc/Rnews/.

Plummer, Martyn, Nicky Best, Kate Cowles, and Karen Vines. 2006. “CODA: Convergence Diagnosis and Output Analysis for Mcmc.” R News 6 (1): 7–11. https://journal.r-project.org/archive/.

R Core Team. 2018. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Revell, Liam J. 2012. “Phytools: An R Package for Phylogenetic Comparative Biology (and Other Things).” Methods in Ecology and Evolution 3: 217–23.

Revelle, William. 2018. Psych: Procedures for Psychological, Psychometric, and Personality Research. Evanston, Illinois: Northwestern University. https://CRAN.R-project.org/package=psych.

Richard A. Becker, Original S code by, Allan R. Wilks. R version by Ray Brownrigg. Enhancements by Thomas P Minka, and Alex Deckmyn. 2018. Maps: Draw Geographical Maps. https://CRAN.R-project.org/package=maps.

Rinker, Tyler W. 2013. qdapDictionaries: Dictionaries to Accompany the Qdap Package. Buffalo, New York: University at Buffalo/SUNY. http://github.com/trinker/qdapDictionaries.

———. 2015. qdapTools: Tools to Accompany the Qdap Package. Buffalo, New York: University at Buffalo/SUNY. http://github.com/trinker/qdapTools.

———. 2017. qdapRegex: Regular Expression Removal, Extraction, and Replacement Tools. Buffalo, New York: University at Buffalo/SUNY. http://github.com/trinker/qdapRegex.

———. 2019. qdap: Quantitative Discourse Analysis Package. Buffalo, New York. http://github.com/trinker/qdap.

Rodriguez-Sanchez, Francisco. 2018. Grateful: Facilitate Citation of R Packages. https://github.com/Pakillo/grateful.

Rowe, Brian Lee Yung. 2016. Futile.logger: A Logging Utility for R. https://CRAN.R-project.org/package=futile.logger.

Sarkar, Deepayan. 2008. Lattice: Multivariate Data Visualization with R. New York: Springer. http://lmdvr.r-forge.r-project.org.

Simpson, Gavin L. 2016. Permute: Functions for Generating Restricted Permutations of Data. https://CRAN.R-project.org/package=permute.

Tierney, Luke, A. J. Rossini, Na Li, and H. Sevcikova. 2018. Snow: Simple Network of Workstations. https://CRAN.R-project.org/package=snow.

Uyeda, Josef C., Jon Eastman, and Luke Harmon. 2018. Bayou: Bayesian Fitting of Ornstein-Uhlenbeck Models to Phylogenies. https://CRAN.R-project.org/package=bayou.

Wang, Jiahui, Ruben Zamar, Alfio Marazzi, Victor Yohai, Matias Salibian-Barrera, Ricardo Maronna, Eric Zivot, et al. 2017. Robust: Port of the S+ “Robust Library”. https://CRAN.R-project.org/package=robust.

Warnes, Gregory R., Ben Bolker, Lodewijk Bonebakker, Robert Gentleman, Wolfgang Huber Andy Liaw, Thomas Lumley, Martin Maechler, et al. 2016. Gplots: Various R Programming Tools for Plotting Data. https://CRAN.R-project.org/package=gplots.

Wickham, Hadley. 2007. “Reshaping Data with the reshape Package.” Journal of Statistical Software 21 (12): 1–20. http://www.jstatsoft.org/v21/i12/.

———. 2011. “The Split-Apply-Combine Strategy for Data Analysis.” Journal of Statistical Software 40 (1): 1–29. http://www.jstatsoft.org/v40/i01/.

———. 2016. Ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. https://ggplot2.tidyverse.org.

———. 2019. Stringr: Simple, Consistent Wrappers for Common String Operations. https://CRAN.R-project.org/package=stringr.

Wickham, Hadley, and Lionel Henry. 2018. Tidyr: Easily Tidy Data with ’Spread()’ and ’Gather()’ Functions. https://CRAN.R-project.org/package=tidyr.

Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2018. Dplyr: A Grammar of Data Manipulation. https://CRAN.R-project.org/package=dplyr.

Xie, Yihui. 2018. Knitr: A General-Purpose Package for Dynamic Report Generation in R. https://yihui.name/knitr/.

Yu, Guangchuang, David Smith, Huachen Zhu, Yi Guan, and Tommy Tsan-Yuk Lam. 2017. “Ggtree: An R Package for Visualization and Annotation of Phylogenetic Trees with Their Covariates and Other Associated Data.” Methods in Ecology and Evolution 8 (1): 28–36. doi:10.1111/2041-210X.12628.