

Environmentally-related genotypic, phenotypic and metabolic diversity of oat (*Avena sativa* L.) landraces based on 67 Polish accessions

Maja Boczkowska · Jacek Zebrowski · Jarosław Nowosielski · Izabela Kordulasińska · Dorota Nowosielska · Wiesław Podyma

Received: 30 April 2017 / Accepted: 14 August 2017 / Published online: 30 August 2017
© The Author(s) 2017. This article is an open access publication

Abstract Here we focused on Polish oat landraces developed in the moderate temperature zone attempting to understand their genetic and phenotypic variability in context of the pressure of environmental constraints at the locations where they had evolved. Besides molecular markers and diverse phenotypic traits the Fourier transform infrared (FTIR) spectroscopy, enabling throughput metabolic fingerprinting, was used for the description of the germplasm collection. We used sophisticated multivariable analysis to integrate the data for structure analysis and to establish mutual relationships. We found that the accessions diversity across the eco-geographical variables was manifested differently depending on selected phenotypic, genetic, and metabolic criteria.

Grain metabolic fingerprint derived from FTIR spectroscopy revealed the highest degree of germplasm diversity among all of the examined traits, particularly for spectral bands assigned to lipids. The landraces, which were collected in close geographical proximity, showed clear morphological and metabolic resemblances, although they represented quite different genetic backgrounds. All three levels of analysis showed the presence of selection resulting from environmental pressures and more specifically from the temperature at the landrace origin site. This research also proved that coupling of genetic polymorphism with the FTIR fingerprinting markedly extended the description of the oat landraces diversity.

Keywords *Avena sativa* L. · Environment · FTIR spectroscopy · Genetic diversity · Metabolic fingerprint · Phenotypic description

Electronic supplementary material The online version of this article (doi:10.1007/s10722-017-0555-8) contains supplementary material, which is available to authorized users.

M. Boczkowska (✉) · W. Podyma
Polish Academy of Sciences, Botanical Garden - Center for Biological Diversity Conservation in Powsin, Warsaw, Poland
e-mail: maja.boczkowska@gmail.com

J. Nowosielski · I. Kordulasińska · D. Nowosielska
Plant Breeding and Acclimatization Institute (IHAR) - National Research Institute, Radzików, Poland

J. Zebrowski
Institute of Biotechnology, University of Rzeszów, Rzeszów, Poland

Introduction

Landraces are a dynamic, heterogeneous population of crops that have evolved in response to natural selection for the local environment, mutations, migrations and genetic drift, but also preferences and agricultural practices of local farmers (Boggini et al. 1997; Frankel et al. 1995; Mohammadi et al. 2014; Pusadee et al. 2014; Villa et al. 2005). They are associated with traditional farming systems and are not the result of premediated crop improvement (Villa

et al. 2005). Current elite cultivars have significantly lower genetic diversity than landraces or wild relatives (Boczkowska et al. 2015; Boczkowska and Onyśk 2016). Beside the genetic diversity aspect, landraces have been a significant part of traditional knowledge and history of the local community, which is still underestimated. Poland was one of the very few European countries that maintained the on-farm cultivation of cereal landraces for a relatively long time due to the considerable fragmentation of farms and, consequently, extensive land use. Such crop management was especially typical for eastern and southern parts of the country (Hammer et al. 1981; Hanelt and Hammer 1977; Kulpa and Jastrzębski 1986). Currently, cereal landraces are rarely collected during expeditions due to significant changes in agriculture.

The total area of arable land under sown crops decreased by about 4 million hectares in Poland during the last 50 years, which is over 25% total area. The decline was particularly drastic after 1990, when economic factors caused the set aside of large arable land. In the period of <20 years (1990–2008) the area of agricultural lands decreased by over 2.5 million ha. This decline was due to the transfer of land to non-agricultural purposes, including afforestation. Poor and very poor soils constituted more than 60% of land allocated for non-agricultural purposes (Krasowicz and Kuś 2010). Also many farms, especially small ones, have resigned in recent years from plant production. The average area of farms is growing, although still among 1.4 million farms 1 million has the size of 1–10 ha. There are also very large changes in the area of cultivation of some cereal species, this is particularly evident in the case of rye and oats. During last 50 years, the area of oats has decreased more than 3 times, from about 1.7 million ha to 0.5 million ha (Sułek and Leszczyńska 2004). Significant decline in the stock of horses and unfavourable prices for seeds result in a systematic decline in the oat growing area. Because traditionally oat grain is mainly used as a fodder, and only in a small part is used for consumption and industrial processing. However, in submontane and in the north especially in the north-eastern part of the country, oats still play an important role in the sowing structure (Sułek and Leszczyńska 2004).

The main aim of this study was a comprehensive analysis of the diversity of common oat landraces

acquired from different regions of Poland employing multiple phenotypic, genetic and metabolic approaches. The focus was given to establish mutual relationships between germplasm features determined at various description levels and to assess the impact of eco-geographic conditions of landrace place of origin on the descriptor patterns. The biochemical characterization of germplasm was performed using infrared spectroscopy, a convenient approach for high-throughput analysis, which provided a metabolic fingerprint of a sample. The feasibility of this method for discriminating and taxonomic applications have been successfully proved in researches with various plant organs or parts including seeds (Demir et al. 2015; Liu and Yu 2011), leaves (Kim et al. 2004; Lu et al. 2008), petals (Li et al. 2015), roots (Naumann et al. 2010), or pollens (Zimmermann and Kohler 2014).

To the best of our knowledge, this is the first attempt to associate spectral features of the same biological material to the extended range of characteristics including eco-geographic data, multiple phenotypic traits and genetic background.

Methods

Plant material

A germplasm assembly of 67 common oat landraces collected from 1973 to 1999 in Poland was obtained from the National Centre for Plant Genetic Resources (NCPGR). The list of tested accessions is provided in the supplementary file Appendix Table 1. The botanical varieties of all accessions were determined according to the classification of Rodionova (Rodionova et al. 1994) and was in detail described by Boczkowska and Tarczyk (2013).

Eco-geography

The eco-geographic conditions of collection sites were defined on the basis of geographic coordinates in the passport data (Nowosielska and Nowosielski 2008) (Appendix Table 1). Oat (*Avena sativa* L.) landraces were collected in five regions of Poland. The geographic coordinate ranges of the collection sites were 49.21–54.20 N latitude and 19.16–23.36 E longitude. The altitude of the collection sites was in the range from 100 to 913 meters above mean sea level (m a. s.

1.). Based on statistical meteorological data, annual average temperature in Poland is estimated at approximately 8 °C; however, in mountainous regions, it is significantly below the average. Annual average temperature in the collection sites was between 5.7 and 8.8 °C with the mean 7.7 °C. Average annual precipitation in Poland is about 630 mm. In the collection sites it ranged from 535 to 1066 mm with the average approximately equal to 730 mm, which resulted from typically much higher precipitation in mountain areas.

Genetic analysis

Each accession was represented by a bulk sample of 25 randomly chosen plants. The genetic analysis was described in detail in our previous studies i.e. AFLP (Boczkowska et al. 2014) and ISSR (Boczkowska and Tarczyk 2013) and is also enclosed in the supplementary file (Appendix).

Phenotypic analysis

A phenotype of examined accessions was described using both qualitative (morphological) and quantitative (evaluation of agriculturally relevant traits) characteristics. The morphological description was based on 25 traits (Appendix Table 2), which were assessed according to Boczkowska et al. (2014). Twelve phenological and agronomical quantitative traits, hereinafter called ‘evaluation’ were described on the basis of field experiments during three subsequent seasons, using standardized scales and descriptors (Appendix Table 3). The trial was carried out on an experimental field of the Plant Breeding and Acclimatization Institute at Radzików, Poland in 2009–2011. Each accession was represented by 1000 grains which were sown on the 2.5 m² plots. The characters were scored at an appropriate growth stage of the oat plants following the UPOV (1994). The traits were measured as follow: the number of days to heading was counted from the sowing date to 50% of panicles fully emerged; the number of days to maturity was counted from the sowing date to harvest ripeness; the shoot height was determined by the average of three plants measured from ground to the tip of panicle; the lodging score was determined in mature stage according to a scale where: 1 means all plants lodged and 9 no

lodging at all; the susceptibility to diseases was also determined based on percentage of affected plants (natural infections); the grain yield was assessed as the weight of grain per plot; the thousand seed weight was estimated based on the average weight of three samples of 100 fully filled grains; the number of panicles (number of fertile tillers) was an average of 10 plants.

Metabolic analysis

About fifty randomly selected dehusked oat grains were ground and homogenized with the laboratory ball mill (MM 301, Retsch GmbH, Germany) into a fine powder. After drying at 35 °C for 24 h, the material was transferred to an exicator and kept until the measurements. FTIR spectroscopy measurements were performed at ambient temperature using an FTIR spectrometer (Nicolet iZ 10 module, Thermo Scientific, USA) equipped with KBr/Ge beam splitter, a deuterated triglycine sulfate (DTGS) detector and the Smart Orbit Attenuated Total Reflectance (ATR) accessory endowed with one-bounce diamond crystal. Sixty-seven interferograms were collected for three biological replications at a resolution of 4 cm⁻¹ and co-added within the wavenumber range between 4000 and 400 cm⁻¹. Spectra were background, vapor absorption and rubber band baseline corrected, smoothed, area-normalized and averaged using OMNIC software (v. 8.1, Thermo Fischer Scientific Inc.).

Data analysis

For establishing mutual correlations and interrelationships between various types of data, measured parameters were converted to binary matrices. The length of AFLP and ISSR products were scored as 1 in their presence and 0 in the absence. The morphological qualitative traits were converted in the similar manner i.e. the presence or absence of a particular state of a trait was coded as 1 or 0, respectively. FTIR spectral data collected for the whole range of wavenumbers between 4000 and 400 cm⁻¹, the same as the quantitative evaluation scores (mean of 3 years), were normalized and then binarized. The normalization was defined by:

$$a'_i = \frac{a_i - a_{i_min}}{a_{i_max} - a_{i_min}}$$

where a_i is an original value, a_{i_max} is the maxima among all the data points, a_{i_min} is the minima among all the data points and a'_i is the data point i normalized between 0 and 1. The resulting data were binarized as: $[0, 0.5] = 0$ and $(0.5, 1] = 1$. In addition, spectral bands of 1700–1480 and 2990–2830 cm^{-1} wavelengths assigned to protein and lipid, respectively, were extracted and considered as further two matrices. Furthermore, the matrices of genetic (AFLP + ISSR), phenotypic (morphology + evaluation) and spectral metabolic data (the whole range) were prepared. Subsequent analyses were carried out in the same manner for all of the matrices. The Nei's coefficient (Lynch and Milligan 1994) was calculated to estimate genetic variation within the groups of accessions based of eco-geographic data, morphological traits and the performed method. One-way ANOVA was used to confirm the significance of differences in the coefficient values between groups. *Post-hoc* Fisher's least significant difference (LSD) was applied to explore all possible pair-wise comparisons of means. The genetic distance was calculated based on Nei's formula (Nei 1978). The geographic distance was calculated based on geographic coordinates of collection sites. Matrices of absolute values of environmental conditions were created based on eco-geographic data. The Mantel test (Mantel 1967) with 999 permutations was conducted to compare dissimilarity, geographic distance and absolute value of difference in environment conditions matrices. The Principal Coordinate Analysis (PCoA) was aimed at graphical representation of a resemblance matrix between landraces. Generalized Procrustes Analysis (GPA) was used to obtain a consensus configuration of all analysis types. Redundancy Analysis (RDA) allowed the relationship between PCoA matrices and eco-graphic and phenotypic data to be studied. The two-way ANOVA was conducted for agronomic traits and included in evaluation part of phenotypic description. All analyses were performed using GenAlex 6.5 (Peakall and Smouse 2012) and Microsoft® Excel 2010/XLSTAT©-Pro software (Version 2013.4.07, Addinsoft, Inc., Brooklyn, NY, USA).

Results

Phenotype description

Qualitative traits (morphological description)

The morphological traits differed in terms of variation calculated as Nei's coefficient (H_j) within the collection (Appendix Fig. 1). The highest variability among tested landraces was observed for the angle between the leaves and culm axis ($H_j = 0.426$). Both erectness of spikelet and hairiness of lemma showed a total lack of variation within the set of landraces. All of the plants had drooping spikelet and hairless lemma. The remaining 22 traits revealed medium and faint diversity.

Quantitative traits (phenological and agricultural values)

Yield-forming parameters, as well as the disease resistance scores, showed great variability in dependence on accession (Appendix Table 3). The two-way ANOVA indicated that the heading date was the sole trait determined by genotype, while other characteristics were influenced by environmental conditions to varying degrees.

FTIR spectroscopy based—metabolic fingerprint

Dehusked oat grains showed a few distinct and prominent peaks in the spectral region characteristic for functional groups between 4000 and 1500 cm^{-1} and several overlapping peaks in the fingerprint region between 1500 and 900 cm^{-1} (Appendix Fig. 2a). Broad and strong band centered at 3316 cm^{-1} reflected the N–H and the O–H stretching vibrations due to presence mostly of proteins and polysaccharides. It is also shaped to some extent by the presence of liquid water in grains, particularly through the contribution of symmetric ($\sim 3270 \text{ cm}^{-1}$) and asymmetric (3490 cm^{-1}) stretching and the overtone of bending vibrations of water molecules ($\sim 3290 \text{ cm}^{-1}$) (Max and Chapados 2009; Venyaminov and Prendergast 1997). Broadening of the band results from the extensive hydrogen binding of water and the biomolecules containing –NH and –OH groups (Stuart 2005). The strong peaks at 2926 and 2854 cm^{-1} were

assigned to methylene $-\text{CH}_2$ asymmetric and symmetric stretching vibrations, respectively. The bands specific for methyl $-\text{CH}_3$ groups occurred as minor shoulders at 2962 and 2863 cm^{-1} (Appendix Fig. 2b), and showed insubstantial absorbance relative to the methylene groups ($-\text{CH}_2$). This indicates the prevailing contribution of long-chain fatty acids with rather minor impact of side chains in proteins to the spectral features in this region (Silverstein et al. 2005). The spectral features of this region were used in the present paper for discriminating accessions in terms of lipid fraction content and quality. The presence of a shoulder due to carbonyl $\text{C}=\text{O}$ stretching vibrations of the ester group located at 1740 cm^{-1} and the lack of signals generated by free acids at about 1710 cm^{-1} indicate that the fatty acids were largely in esterified forms (Vlachos et al. 2006). Protein fraction of the oat grains was spectrally characterized by the shape and height of bands centered at 1652 and 1544 cm^{-1} assigned as amide I and amide II bands, respectively (Appendix Fig. 2a) (Barth 2007). The features of this region were considered here as spectral characteristics of the grain proteins.

The wavenumber range between 1200 and 900 cm^{-1} with a prominent peak at 1023 cm^{-1} and two distinct shoulders at 1152 and 1079 cm^{-1} on the higher frequency side and a minor shoulder at 960 cm^{-1} on a lower frequency side are considered spectral characteristics of carbohydrates (Kačuráková and Wilson 2001) (Appendix Fig. 2c). This region consists of many overlapping peaks due to various sugar compounds and linkages and to some extent may show contribution of nucleic acids (Tajmir-Riahi et al. 2009). The distinct bands in FTIR spectrum from oat grains (upper) at 1152, 932, 858, and 763 cm^{-1} were located at the wavelengths characteristic for spectrum of starch, taken herein as the experimental reference (lower), and thus reflected the contribution from this component. However, the most prominent and broad band centered at 1023 cm^{-1} only partially corresponded to the starch spectral features and should thus be considered the result of overlapping bands attributed to remaining polysaccharide components, mostly from (1 \rightarrow 3) (1 \rightarrow 4)- β -glucan, and to a much lower extent from cellulose, occurring at relatively minor content, which typically show peaks around 1030 cm^{-1} (Guillon et al. 2011). The absorbance from (1 \rightarrow 3) (1 \rightarrow 4)- β -glucan also contributed to the bands

at 1152 and 1080 cm^{-1} . In turn, the band of highly branched arabinoxylans, prominent in wheat at around 1040 cm^{-1} (Toole et al. 2007) was not present in the oat landraces. Overall, the whole-range infrared spectrum taken in this paper as one of the accession characteristics gave insight into a metabolic profile of dehusked grains, reflecting relative contribution of main biochemical components including storage and structural polysaccharides, proteins, lipids and some minor chemical compounds e.g. phenolics.

Diversity

Comparison of the diversity within genetic, phenotypic and metabolic data (Appendix Fig. 3), showed that the highest differentiation occurred for the latter ($H_j = 0.378$), while the lowest was obtained for genetic data ($H_j = 0.217$). Variation detected by particular techniques was in the range from 0.205 for AFLP to 0.358 for FTIR spectrometry of lipids. The level of diversity was also determined within the groups (Appendix Fig. 4) created based on passport and eco-geographic data.

Regions

According to the genetic and metabolic datasets, the highest differentiation occurred in the South region, while phenotypic data indicated the highest variation in the South-East. Both genetic and phenotypic data showed that the minimum variation was attributed to the central part of Poland, but metabolic data revealed the lowest value in the North-East. Each of the extracted techniques, except for the quantitative traits approach, showed a minimum diversity in the center of the country (Appendix Fig. 4).

Altitude

The accessions collected above 600 mamsl had higher metabolic variation compared to those acquired from lower altitudes. The genetic and phenotypic results presented quite opposite trends. Analysis of data acquired by particular techniques showed that the differentiation of accessions increased along with altitude for both lipids and proteins (i.e. FTIR ranges corresponding to lipids and proteins), except for the group >800 mamsl (Appendix Fig. 4).

Average annual precipitation

The highest metabolic variation was observed within the group with the highest annual rainfall. The genotypic data presented the opposite trend. For the phenotype, no statistically significant differences between groups in the H_j level were found, regardless of the method used for plant description (Appendix Fig. 4).

Average annual temperature

The diversity differed significantly exclusively for metabolic data. Interestingly, the level of this variation declined as the temperature increased (Appendix Fig. 4).

Mantel test

Pairwise Mantel test was performed for all of combinations of distance matrices (Appendix Table 4). The statistically significant correlation was observed for the comparison of genetic and phenotypic distance matrices (0.272, $p = 0.05$). The metabolome correlations with both genotype and phenotype were insignificant ($p > 0.05$). On the other hand, comparison of particular techniques allowed a few other significant correlations to be identified, including ISSR *vs* both morphology and proteins, and morphology *vs* both evaluation and proteins.

In the next step, all of the developed distance matrices were also compared with the geographical distance and absolute differences in altitude, precipitation and temperature (Appendix Table 4). The phenotype matrix correlated significantly but at a low level with all of the tested matrices, while the genotype was not correlated with geographical distance. At the same time, no significant correlation was found between the metabolome and matrices describing the collection site. Further relationships are given in more detail in Appendix file.

Grouping analysis

Combination of PCoA results obtained from AFLP, ISSR, phenotypic qualitative and quantitative traits, and protein and lipid spectroscopic profiles was done using the Generalized Procrustes Analysis (GPA). The consensus proportion (i.e. Rc = the proportion of total

variance explained by the consensus matrix) for the data set was rather low (0.297) but permutation test ($n = 10,000$) indicated that it was statistically significant ($p < 0.05$) so the agreement among six methods was not strong. Based on pairwise Rc values, it was clear that AFLP and evaluation gave the less consensual results (Appendix Table 5). The first two PCA dimensions (F1 and F2) accounted for 58% of the variance in consensus matrix (Fig. 1). At first glance, based on the bi-plot of the first two dimensions, no distinct group of accession was seen, but a thorough analysis of the scatter plots showed that the grouping pattern was related to yield (RDA correlation = 0.661) and to average annual temperature in collection site (RDA correlation = 0.561) (Table 1; Fig. 1).

Discussion

In this paper, we evaluated the impact of the ecosystem attributes defined by eco-geographical conditions on germplasm biodiversity using 67 oat landraces. An immense data set, determined by a number of various techniques including genetic (1260 loci), phenotypic (37 traits) and metabolomics approaches (several hundred variables), was employed for combined and comparative analysis and establishing regularities that characterized the landraces.

Analysis of the phenotypic data showed that morphologically similar accessions had a very different agronomically relevant characteristic. Moreover, the environment influenced both the richness of phenotypes and also favored some morphotypes specific for a given location. The interdependence found between stalk morphology and the thousand seed weight additionally confirmed the deliberate selection by farmers. Interestingly, white lemma landraces mostly occurred in mountain regions with the highest total annual precipitation. This might be explained by generally lower tolerance of white lemma genotypes to the drought (Lewicki and Mazurek 1967). Furthermore, accessions collected in close proximity were characterized by a very similar appearance but quite different scores of the evaluation data that, in turn, may indicate preferences of farmers for maintaining specific morphological forms, specifically in particular regions. The morphological diversity of examined populations, ranked in our test on the

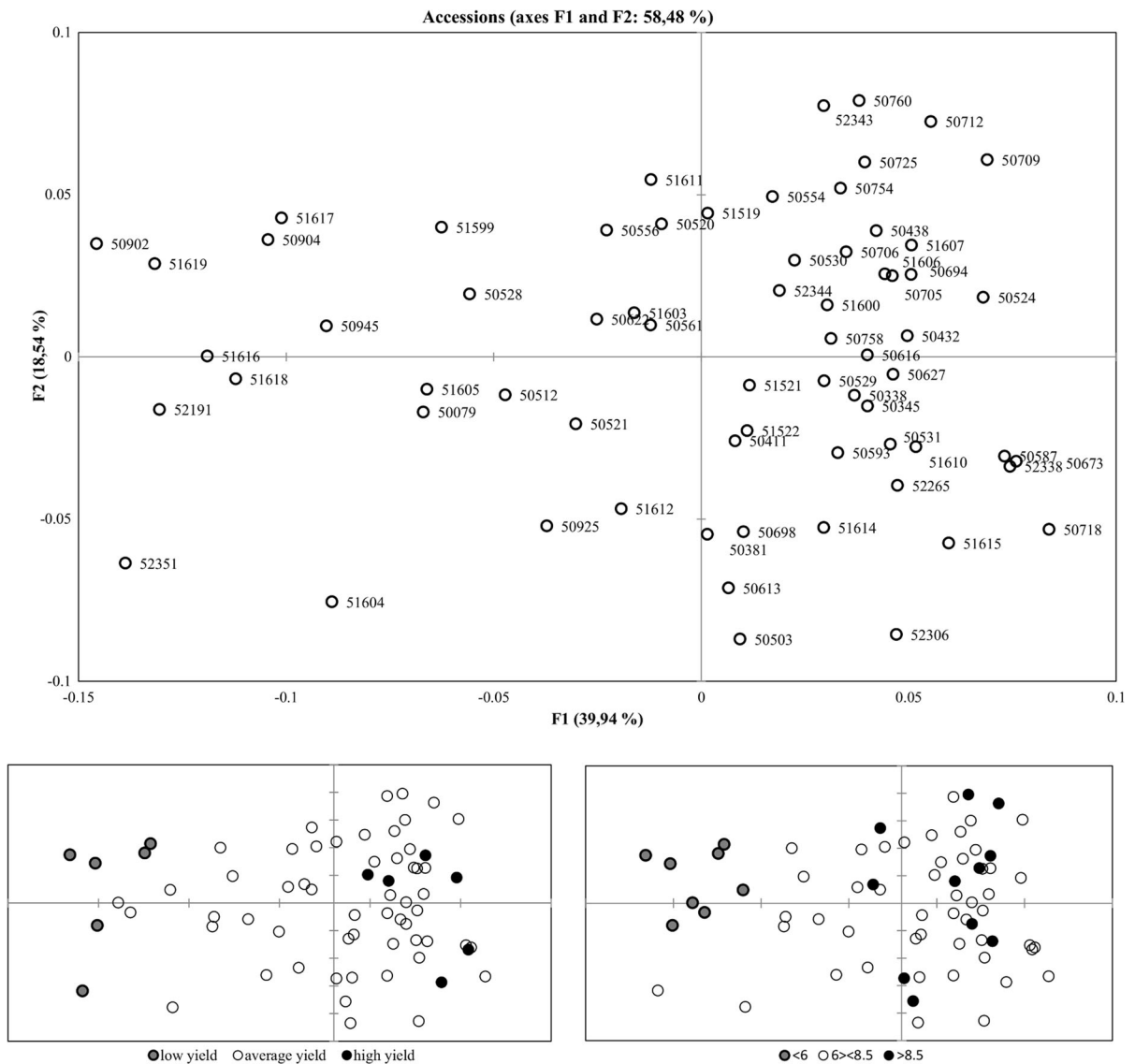


Fig. 1 Biplots of 1 versus 2 axis of GPA. The *upper panel* provides a general plot with the accession numbers. Two small ones below present grouping pattern related with yield (*left*) and with average annual temperature in collection site (*right*)

average level, was higher than estimated on the basis of the same criteria for oat cultivars grown in Poland before 1939 (Boczkowska et al. 2014). All of the analyzed landraces showed relatively high resistance to diseases such as crown and stem rust and powdery mildew. It could be a consequence of the buffering effect of disease spreading in genetically heterogeneous populations such as landraces (Frankel et al. 1995) or due to the presence of different resistance alleles.

Thorough analysis also revealed that the accessions, which were collected in a close proximity, were completely different genetically, despite morphological similarities. This indicates a different origin of landraces grown in particular regions and clearly proved preservation of the accessions’ genetic distinctiveness. Small albeit statistically significant correlation of the genetic markers with altitude, precipitation and temperature was clearly visible. Consequently, the specific genotype patterns of the oat

Table 1 Results of RDA analysis

	AFLP	ISSR	Morphology	Evaluation	Proteins	Lipids	GPA
Passport data							
Results of the permutation test							
Pseudo F	0.189	0.255	0.287	0.201	0.368	0.298	0.313
<i>p</i> value	0.012	0.004	<0.0001	0.010	<0.0001	0.002	<0.0001
Standardized canonical coefficients							
Altitude	−1.164	−0.363	−0.433	−0.967	0.337	−0.092	−0.203
Annual precipitation	0.215	0.218	0.310	0.141	−0.136	−0.119	0.141
Annual average temperature	−0.988	0.487	0.497	−1.078	−0.352	0.304	0.561
Evaluation data							
Results of the permutation test							
Pseudo F	0.112	0.320	0.771	x	0.550	0.291	1.354
<i>p</i> value	0.050	0.092	<0.0001	x	0.004	0.220	<0.0001
Standardized canonical coefficients							
Days to heading	ns	ns	0.014	x	0.005	ns	0.077
Lodging at mature stage	ns	ns	−0.093	x	0.122	ns	0.090
Crop height	ns	ns	−0.133	x	−0.057	ns	0.050
Panicle length	ns	ns	−0.187	x	−0.252	ns	−0.036
Powdery mildew	ns	ns	−0.087	x	−0.033	ns	0.003
Crown rust	ns	ns	0.001	x	−0.247	ns	0.077
Steam rust	ns	ns	−0.073	x	0.117	ns	−0.044
Number of panicles	ns	ns	0.002	x	0.328	ns	−0.245
Days to maturity	ns	ns	0.106	x	−0.045	ns	0.016
Yield	ns	ns	0.212	x	−0.500	ns	0.661
Thousand seed weight	ns	ns	0.471	x	−0.211	ns	0.305
Test weight	ns	ns	−0.122	x	−0.140	ns	0.097
Morphological data							
Results of the permutation test							
Pseudo F	0.869	1.051	x	1.169	1.280	0.651	3.208
<i>p</i> value	0.006	<0.0001	x	<0.0001	<0.0001	0.147	<0.0001
Standardized canonical coefficients							
Intensity of leaves spirality	0.323	−0.164	x	0.186	−0.061	ns	−0.049
Spirality of leaves	−0.228	0.270	x	0.091	0.024	ns	0.156
Color of leaves	0.502	0.219	x	−0.233	−0.096	ns	0.251
Color of panicle	−0.014	−0.046	x	0.019	0.083	ns	−0.074
Lemma color	−0.022	−0.014	x	0.171	0.208	ns	−0.185
Axis nodes in panicle	−0.121	−0.307	x	0.197	0.180	ns	−0.190
Number of grains in spikelet	−0.346	−0.011	x	−0.090	−0.275	ns	0.091
Hairiness of leaf margin	−0.465	0.091	x	0.397	−0.036	ns	0.105
Awn type	−0.336	−0.101	x	−0.297	−0.107	ns	0.217
Angle to culm of leaves	−0.145	−0.007	x	0.164	0.026	ns	−0.011
Angle of flag leaf to culm	0.186	−0.320	x	0.098	−0.006	ns	−0.197
Rigidity of leaves	0.154	−0.096	x	0.109	−0.113	ns	0.009
Growth habit	−0.183	0.066	x	0.041	0.201	ns	−0.166
Shape of panicle	0.133	−0.385	x	−0.007	0.511	ns	−0.461

Table 1 continued

	AFLP	ISSR	Morphology	Evaluation	Proteins	Lipids	GPA
Shape of primary grain	−0.242	0.049	x	0.100	0.053	ns	−0.113
Nodes hairiness	−0.161	−0.171	x	0.025	−0.083	ns	−0.087
Hairiness of basal part of the primary grain	−0.003	−0.233	x	−0.016	0.053	ns	−0.059
Hairiness of rachilla	−0.116	−0.231	x	−0.038	0.136	ns	−0.167
Length of rachilla	0.135	−0.118	x	0.047	−0.097	ns	−0.014
Hairiness of leaf sheath	−0.376	0.185	x	0.231	−0.108	ns	0.071
Type of panicle	−0.316	0.182	x	−0.325	−0.237	ns	0.199
Shape of basal part of the primary grain	−0.019	−0.036	x	−0.011	−0.188	ns	0.223
Awedness	0.268	−0.008	x	−0.046	0.338	ns	−0.232

landraces could be influenced or even created, to a great extent, by the local environment. Based on molecular data, it cannot be unequivocally confirmed whether these changes are related to the coding or non-coding genome regions. However, the comprehensive analysis presented in this paper indisputably demonstrated that some of them were located in coding regions, and were reflected in both the morphological traits and protein spectral features.

Cereal grains are biochemically highly heterogeneous organs, containing starch, proteins and lipids as the major compounds, (Klose and Arendt 2012; Redaelli et al. 2003; Shewry and Halford 2002) and a relatively high fraction of cell wall components, including heteroxylans, beta-mixed glucans and cellulose (Biel et al. 2014; Conciatori 2000; Shewry et al. 2008; Tiwari and Cummins 2009). All of these compounds and their relative contents affect the average metabolic profile of a grain providing a unique biochemical signature to particular cereal crop accessions, strains, cultivars, or varieties. Applying FTIR spectroscopy analysis, we expected to reveal the selective pressure of local environment on the grain metabolic fingerprint, particularly related to lipids and proteins, two main grain compounds which were pivotal in the biological improvement of oats. We found that FTIR spectroscopy revealed a much higher level of diversity of the oat landraces compared to the remaining examined features. This particularly concerned data extracted from the whole mid-infrared range of spectra and from the frequency range characteristic for lipids. Protein-related vibrational data showed a polymorphism comparable to that of genetic markers. The total spectral profile, which

corresponded to almost all functional groups of organic compounds, did not display any statistically significant correlations with both the genotype and phenotype. The lack of a clear relationship between the whole metabolome and AFLP and RFLP markers has been previously reported for the sesame (Laurentin et al. 2008) and rice (Mochida et al. 2009) germplasm, respectively. Truncation of the spectral data to frequencies assigned specifically to proteins or lipids gave a much more interesting insight into the landrace diversity and enabled the direct association of traits with the environment of the place of origin. Particularly, FTIR technique revealed landrace variability in relation to the altitude, annual temperature, and annual precipitation in collection sites, some of which could not be detected with other traits. For the groups classified according to the mean annual temperature, the spectral data exclusively generated significant diversity of the landraces and its level declined as the temperature increased. Since all accessions were examined in the same field trials and were thus exposed to identical environment, the observed metabolomics diversity reflected specific, adaptive modifications of landrace to habitats in the place of their origin. In other studies, the biochemical traits based on non-targeted spectroscopic data have been used as markers for accession identity (Dumlupinar et al. 2011; Luthria et al. 2008), for investigation of the effect of environmental factors on the cereal grain quality (Andersson and Börjesdotter 2011; Shao et al. 2015; Yan et al. 2007), or establishing phylogenetic relationships (Demir et al. 2015).

Oat grains contain lipids at relatively high concentration, mainly in triacylglycerol form (Klose and

Arendt 2012), and show high genotypic variability in this trait (Banaś et al. 2007; Biel et al. 2014; Danytė 2012). They possess exceptionally high capacity, among other cereals, to accumulate the oil in endosperm. The lipidic FTIR-based fingerprint showed the highest variability among all tested traits. This feature was also highly diverse when accessions were grouped according to the annual average temperature, region of landrace origin and altitude as classification criteria. The variability pattern was, however, distinct from that for proteins. Interestingly, a weak but statistically significant correlation was established between FTIR signatures and the ISSR markers.

The diversity level of oat landraces derived from amide I and II band assignment to proteins (Barth 2007) was comparable to that obtained from the genetic polymorphism analysis. A significant correlation occurred between the protein-related spectra and both genotype and phenotypic qualitative traits. Moreover, a large group of landraces with a very high internal similarity of protein and lipid spectral profiles was identified. Interestingly, accessions collected in the same or very close location had a similar lipid profile, whereas the considerable differences were observed in proteins.

This study also revealed interesting relationships indicating the plasticity of the grain biochemical profile dependent on the location of origin and thus prevailing environmental conditions in the natural habitat. This adaptive evolution coincided with the shift in the genetic background of accessions and possibly altered metabolic pathways shaping the pattern of accumulation of the major compounds in grains that was reflected in the FTIR spectral fingerprints. The biochemical diversity of the tested landraces might be a result of both their adaptive fitting to local habitats and the farmer-driven activity (Leclerc and Coppens d'Eeckenbrugge 2011; Watson and Eyzaguirre 2002).

The grains of oats are superior among other cereals due to relatively high content of lipids, health beneficial lipid and protein composition, the high beta-glucan to arabinoxylan ratio and the abundance of various phytonutrients that renders them valuable source of dietary compounds (Araus and Cairns 2014; Banaś et al. 2007; Biel et al. 2009; Kueger et al. 2012; Loskutov and Rines 2011; Vernocchi et al. 2011; Vilmane et al. 2015). Exploration of natural variability in these traits is important for the selection of valuable gene donors for breeding oriented to improved grain

quality. Landraces can also be a superior source of genes determining advantageous response to disadvantage the environment, which may attract attention in nearest future with priorities for the sustainability of crop production and challenging climate change.

Acknowledgements We would like to express our utmost gratitude to Zofia Bulińska-Radomska, Bogusław Łapiński and to Elżbieta Tarczyk for their invaluable help during experiments and manuscript preparation.

Author contributions Designed and performed the experiments: MB, JZ, JN, IK, DN, WP. Analyzed the data: MB. Wrote the paper: MB, JZ. Gave critical comments: WP.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical standards The authors certify that they have NO affiliations with or involvement in any organization or entity with any financial interest, or non-financial interest in the subject matter or materials discussed in this manuscript.

Human and animal rights This article does not contain any studies with human participants or animals performed by any of the authors.

Informed consent For this type of study formal consent is not required.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Andersson AAM, Börjesdotter D (2011) Effects of environment and variety on content and molecular weight of β -glucan in oats. *J Cereal Sci* 54:122–128. doi:10.1016/j.jcs.2011.03.003
- Araus JL, Cairns JE (2014) Field high-throughput phenotyping: the new crop breeding frontier. *Trends Pl Sci* 19:52–61
- Banaś A et al (2007) Lipids in grain tissues of oat (*Avena sativa*): differences in content, time of deposition, and fatty acid composition. *J Exp Bot* 58:2463–2470
- Barth A (2007) Infrared spectroscopy of proteins. *Biochimica et Biophys Acta (BBA) Bioenerg* 1767:1073–1101
- Biel W, Bobko K, Maciorowski R (2009) Chemical composition and nutritive value of husked and naked oats grain. *J Cereal Sci* 49:413–418

- Biel W, Jacyno E, Kawęcka M (2014) Chemical composition of hulled, dehulled and naked oat grains South African. *J Anim Sci* 44:189–197
- Boczkowska M, Onyśk A (2016) Unused genetic resources: a case study of Polish common oat germplasm. *Ann Appl Biol*. doi:[10.1111/aab.12289](https://doi.org/10.1111/aab.12289)
- Boczkowska M, Tarczyk E (2013) Genetic diversity among Polish landraces of common oat (*Avena sativa* L.). *Genet Resour Crop Evol* 60:2157–2169
- Boczkowska M, Nowosielski J, Nowosielska D, Podyma W (2014) Assessing genetic diversity in 23 early Polish oat cultivars based on molecular and morphological studies. *Genet Resour Crop Evol* 61:927–941
- Boczkowska M, Harasimiuk M, Onyśk A (2015) Studies on genetic variation within old Polish cultivars of common oat. *Cereal Res Commun*. doi:[10.1556/CRC.2014.0025](https://doi.org/10.1556/CRC.2014.0025)
- Boggini G, Doust MA, Annicchiarico P, Pecetti L (1997) Yielding ability, yield stability, and quality of exotic durum wheat germplasm in Sicily. *Pl Breed* 116:541–545. doi:[10.1111/j.1439-0523.1997.tb02187.x](https://doi.org/10.1111/j.1439-0523.1997.tb02187.x)
- Conciatori A, Stefanis ED, Redaelli R, Sgrulletta D (2000) Chemical characterization of some kernel traits in naked oat genotypes. *J Genet Breed* 54:299–302
- Danytė V (2012) Comparison of advanced oat breeding lines in Lithuania. In: Proceedings of the Latvian academy of sciences. Section B. Natural, Exact, and Applied Sciences, vol 4–5, pp 172–176
- Demir P, Onde S, Severcan F (2015) Phylogeny of cultivated and wild wheat species using ATR–FTIR spectroscopy. *Spectrochim Acta A Mol Biomol Spectrosc* 135:757–763
- Dumlupınar Z, Dokuyucu T, Akkaya A (2011) Identification of Turkish oat landraces (*Avena sativa* L.) based on avenin proteins by SDS-Page technique. *Turk J Field Crops* 16:76–83
- Frankel O, Brown A, Burdon J (1995) The conservation of plant biodiversity. Cambridge University Press, Cambridge
- Guillon F et al (2011) Brachypodium distachyon grain: characterization of endosperm cell walls. *J Exp Bot* 62:1001–1015
- Hammer K, Górski M, Hanelt P, Kühn F, Kulpa W, Schultze-Motel J (1981) Variability of wheat land-races from Czechoslovakia and Poland. *Kulturpflanze* 29:91–97
- Hanelt P, Hammer K (1977) Bericht über eine Reise nach der VR Polen 1976 zur Sammlung autochthoner Sippen von Kulturpflanzen. *Kulturpflanze* 27:33–44
- Kačuráková M, Wilson R (2001) Developments in mid-infrared FT-IR spectroscopy of selected carbohydrates. *Carbohydr Polym* 44:291–303
- Kim S et al (2004) Taxonomic discrimination of flowering plants by multivariate analysis of Fourier transform infrared spectroscopy data. *Pl Cell Rep* 23:246–250
- Klose C, Arendt EK (2012) Proteins in oats; their synthesis and changes during germination: a review. *Crit Rev Food Sci Nutr* 52:629–639
- Krasowicz S, Kus J (2010) Kierunki zmian w produkcji rolniczej w Polsce do roku 2020—próba prognozy (General directions of the agricultural production in Poland until the year 2020—a prognosis). *Zagadnienia Ekonomiki Rolnej* 3:5–18
- Kueger S, Steinhauser D, Willmitzer L, Giavalisco P (2012) High-resolution plant metabolomics: from mass spectral features to metabolites and from whole-cell analysis to subcellular metabolite distributions. *Pl J* 70:39–50
- Kulpa W, Jastrzębski A (1986) Zasoby miejscowych form roślin uprawnych Cz.I. Wyniki eksploracji Płaskowyżu Kolbuszowskiego, Pogórza Karpackiego i Beskidów w latach 1976 i 1978. *Biuletyn IHAR* 160:27–45
- Laurentin H, Ratzinger A, Karlovsky P (2008) Relationship between metabolic and genomic diversity in sesame (*Sesamum indicum* L.). *BMC Genom* 9:250. doi:[10.1186/1471-2164-9-250](https://doi.org/10.1186/1471-2164-9-250)
- Leclerc C, Coppens d'Eeckenbrugge G (2011) Social organization of crop genetic diversity. The G × E × S interaction model. *Diversity* 4:1–32
- Lewicki S, Mazurek J (1967) Owies. Państwowe Wydawnictwo Rolnicze i Lesneś, Warszawa
- Li X, Yi S, Zheng Y, He S (2015) Analysis of phylogenetic relationships of main citrus germplasms based on fir spectra of petals. *Intell Autom Soft Comput* 21:295–304
- Liu N, Yu P (2011) Molecular clustering, interrelationships and carbohydrate conformation in hull and seeds among barley cultivars. *J Cereal Sci* 53:379–383
- Loskutov I, Rines H (2011) *Avena*. In: Kole C (ed) Wild crop relatives: genomic and breeding resources. Springer, Berlin, pp 109–183. doi:[10.1007/978-3-642-14228-4_3](https://doi.org/10.1007/978-3-642-14228-4_3)
- Lu H-F, Shen J-B, Lin X-Y, Fu J-L (2008) Relevance of Fourier transform infrared spectroscopy and leaf anatomy for species classification in *Camellia* (*Theaceae*). *Taxon* 57:1274
- Luthria DL, Lin L-Z, Robbins RJ, Finley JW, Banuelos GS, Harnly JM (2008) Discriminating between cultivars and treatments of broccoli using mass spectral fingerprinting and analysis of variance–principal component analysis. *J Agric Food Chem* 56:9819–9827
- Lynch M, Milligan BG (1994) Analysis of population genetic-structure with RAPD markers. *Mol Ecol* 3:91–99. doi:[10.1111/j.1365-294X.1994.tb00109.x](https://doi.org/10.1111/j.1365-294X.1994.tb00109.x)
- Mantel N (1967) The detection of disease clustering and a generalised regression approach. *Cancer Res* 27:209–220
- Max J-J, Chapados C (2009) Isotope effects in liquid water by infrared spectroscopy. III. H₂O and D₂O spectra from 6000 to 0 cm⁽⁻¹⁾. *J Chem Phys* 131:184505
- Mochida K, Furuta T, Ebana K, Shinozaki K, Kikuchi J (2009) Correlation exploration of metabolic and genomic diversity in rice. *BCM Genom* 10:568
- Mohammadi R, Haghparast R, Sadeghzadeh B, Ahmadi H, Solimani K, Amri A (2014) Adaptation patterns and yield stability of durum wheat landraces to highland cold rainfed areas of Iran. *Crop Sci* 54:944–954. doi:[10.2135/cropsci2013.05.0343](https://doi.org/10.2135/cropsci2013.05.0343)
- Naumann A, Heine G, Rauber R (2010) Efficient discrimination of oat and pea roots by cluster analysis of Fourier transform infrared (FTIR) spectra. *Field Crops Res* 119:78–84
- Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89:583–590
- Nowosielska D, Nowosielski J (2008) Morphological diversity and DNA polymorphism of common oat (*Avena sativa* L.) landraces cultivated in Poland. *Plant Breed Seed Sci* 58:11–22
- Peakall R, Smouse PE (2012) GenA1Ex 6.5: genetic analysis in Excel. Population genetic software for teaching and

- research-an update. *Bioinformatics* 28:2537–2539. doi:[10.1093/bioinformatics/bts460](https://doi.org/10.1093/bioinformatics/bts460)
- Pusadee T, Oupkaew P, Rerkasem B, Jamjod S, Schaal B (2014) Natural and human-mediated selection in a landrace of Thai rice (*Oryza sativa*). *Ann Appl Biol* 165:280–292
- Redaelli R, Sgrulletta D, De Stefanis E (2003) Genetic variability for chemical components in sixty European oat (*Avena sativa* L.) cultivars. *Cereal Res Commun* 31:185–192
- Rodionova N, Soldatov V, Merezko V, Jarosh N, Kobyljanskij V (1994) Flora of cultivated plants, vol 2. Kolos, Moscow
- Shao Y, Xu F, Chen Y, Huang Y, Beta T, Bao J (2015) Analysis of genotype, environment, and their interaction effects on the phytochemicals and antioxidant capacities of red rice (*Oryza sativa* L.). *Cereal Chem* 92:204–210
- Shewry PR, Halford NG (2002) Cereal seed storage proteins: structures, properties and role in grain utilization. *J Exp Bot* 53:947–958
- Shewry PR et al (2008) Phytochemical and fiber components in oat varieties in the HEALTHGRAIN diversity screen. *J Agric Food Chem* 56:9777–9784
- Silverstein R, Bassler G, Morrill T (2005) Spectrometric identification of organic compounds, 7th edn. Wiley, New York
- Stuart B (2005) Infrared Spectroscopy. In: Kirk R, Othmer D (ed) *Encyclopedia of Chemical Technology*. Wiley. doi:[10.1002/0471238961.0914061810151405.a01.pub2](https://doi.org/10.1002/0471238961.0914061810151405.a01.pub2)
- Sulek A, Leszczyńska D (2004) Stan aktualny i perspektywy uprawy owsa w Polsce Komunikat (Actual state and perspectives of oats cultivation in Poland Short communication). *Biuletyn Instytutu Hodowli i Aklimatyzacji Roślin* 231:387–395
- Tajmir-Riahi H, N'Soukpoe-Kossi C, Joly D (2009) Structural analysis of protein–DNA and protein–RNA interactions by FTIR, UV-visible and CD spectroscopic methods. *J Spectrosc* 23:81–101
- Tiwari U, Cummins E (2009) Simulation of the factors affecting β -glucan levels during the cultivation of oats. *J Cereal Sci* 50:175–183
- Toole G, Wilson R, Parker M, Wellner N, Wheeler T, Shewry P, Mills E (2007) The effect of environment on endosperm cell-wall development in *Triticum aestivum* during grain filling: an infrared spectroscopic imaging study. *Planta* 225:1393–1403
- UPOV (1994) International convention for the protection of new varieties of plants. UPOV, Geneva
- Venyaminov SY, Prendergast FG (1997) Water (H_2O and D_2O) molar absorptivity in the 1000–4000 cm^{-1} range and quantitative infrared spectroscopy of aqueous solutions. *Anal Biochem* 248:234–245. doi:[10.1006/abio.1997.2136](https://doi.org/10.1006/abio.1997.2136)
- Vernocchi P, Vannini L, Gottardi D, Del Chierico F, Serazanetti DI, Ndagijimana M, Guerzoni ME (2011) Integration of datasets from different analytical techniques to assess the impact of nutrition on human metabolome. *Front Cell Infect Microbiol* 2:156
- Villa TCC, Maxted N, Scholten M, Ford-Lloyd B (2005) Defining and identifying crop landraces. *Pl Genet Res Charact Util* 3:373–384
- Vilmane L, Zute S, Straumīte E, Galoburda R (2015) Protein, amino acid and gluten content in oat (*Avena sativa* L.) grown in Latvia. In: *Proceedings of the Latvian academy of sciences. Section B. Natural, exact, and applied sciences. vol 4*, pp 170–177
- Vlachos N, Skopelitis Y, Psaroudaki M, Konstantinidou V, Chatzilazarou A, Tegou E (2006) Applications of Fourier transform-infrared spectroscopy to edible oils. *Anal Chim Acta* 573:459–465
- Watson JW, Eyzaguirre PB (2002) Home gardens and in situ conservation of plant genetic resources in farming systems. *Bioversity International*, Rome
- Yan W, Molnar SJ, Fregeau-Reid J, McElroy A, Tinker NA (2007) Associations among oat traits and their responses to the environment. *J Crop Improv* 20:1–29
- Zimmermann B, Kohler A (2014) Infrared spectroscopy of pollen identifies plant species and genus as well as environmental conditions. *PLoS One* 9:e95417