Genetic diversity among hexaploid wheat landraces with different geographical origins revealed by microsatellites: comparison with AFLP, and RAPD data

P. Strelchenko[1], Kenneth Street[2], O. Mitrofanova[1], M. Mackay[3] and F. Balfourier[4]

[1]All-Russian Research Institute of Plant Industry (VIR), St. Petersburg, Russia, www.vir.nw.ru
Email p.strelchenko@vir.nw.ru
[2]Genetic Resources Unit, International Centre for Agricultural Research in the Dry Areas, Aleppo, Syria
[3]Australian Winter Cereals Collection, Tamworth, Australia
[4]INRA Genetics and Plant Breeding Station, Clermont-Ferrand, France

Abstract

To investigate the genetic diversity of 78 landraces that originated from 22 countries and the relationships between them the polymorphism of 20 wheat microsatellite loci was estimated. Data obtained was later compared with the results of a more recent study of the same set of landraces using AFLP and RAPD methods. Each of the three approaches was useful for characterizing genetic diversity and defining relationships between landraces with different geographical origins. However SSRs were more effective genetic markers for detecting genotypes containing combinations of rare alleles. Multivariate statistics techniques were applied to identify groups of genetically similar landraces on the basis of the band data generated by both SSR and combined set of data. This analysis divided the landraces into European and Asian wheats. Within the Asian set the groups largely consisted of landraces originated from the same region. Distinctions between European and Asian wheats, and also between groups within Asian wheats, were mainly determined by differences in the frequency of a considerable number of SSR alleles or AFLP and RAPD bands. The incongruity between the taxonomic division of hexaploid wheats and relationships of the landraces revealed by DNA markers is discussed.

Media summary

Microsatellite markers are more effective than AFLPs and RAPDs for the characterization of wheat landrace genetic diversity and for detecting genotypes containing combinations of rare alleles.

Key Words

DNA polymorphism, wheat classification.

Introduction

Over thousands of years landraces of hexaploid wheats, with a genome composition of AABBDD, have developed under a variety of different edaphic and climatic environments. This has led to the evolution of a large number of ecotypes adapted to specific local environments. In the past attempts have been made to describe the eco-geographical differentiation of wheat using morphological and agronomical traits (Palmova 1935; Vavilov, 1964). Recent developments

using PCR based methods have allowed fast and effective approaches for examining plant polymorphism at the DNA level. For example, up to now more than one thousand microsatellite markers (SSRs) have been developed and mapped in the wheat genome. The information generated can be applied for tagging resistance genes, identifying QTLs, marker-assisted selection and improvement of germplasm management. In this study SSRs were applied to characterize the genetic relationships between hexaploid wheat landraces with different geographical origins and compare the results with a previous study (Strelchenko et al., 2003) carried out using AFLP and RAPD markers.

Methods

Seventy eight landraces of hexaploid wheat, originating from 22 countries, were selected from the VIR germplasm collection. They included *T. aestivum* L., *T. compactum* Host, *T. sphaerococcum* Perc., *T. petropavlovskyi* Udach. et Migush., *T. spelta* L., *T. macha* Dek. et Men. and *T. vavilovii* (Tum.) Jakubz. In both the SSR and AFLP analyses the same genotypes of each landrace were studied. In the RAPD analysis 15-25 seedlings per each landrace were investigated. Classifications of landraces based on SSR, AFLP, RAPD and combined set (ALL) of data were performed by applying cluster and principal component analyses using the STATISTICA 6.0 software. Dice coefficient and Ward's algorithm were used.

Results and discussion

Nineteen primer pairs: *Xgwm2, 44, 46, 135, 190, 234, 251, 257, 260, 261, 312, 337, 341, 413, 427, 566, 626* and *cfd71* (amplifies two independent loci) were used (Röder et al., 1998; Guyomarc'h et al., 2002). These primers characterized 20 SSR loci and represented one locus from each chromosome (excluding chromosome 6D). A total of 271 alleles were detected. The frequency of them varied from 0.01 to 0.69 with an average of 0.07. Of all the alleles detected, 146 alleles had a frequency lower than 0.05 and thus were considered as rare. The number of alleles per locus ranged from four for *Xgwm415* to 21 for *Xgwm312*. These figures are in agreement with those recently published by Roussel et al (2004) on a set of 559 French bread wheat accessions. Five AFLP primer combinations (*Pst*I-ACC/*Mse*I-AGC, *Pst*I-ACC/*Mse*I-CAC, *Pst*I-ACC/*Mse*I-CCT, *Pst*I-GGG/*Mse*I-GAA and *Pst*I-GGC/*Mse*I-ACC) revealed 90 usable bands, which was 19% of the total number of detectable bands. A total of 125 polymorphic RAPD bands were analyzed using 28 random primers. The average frequencies of AFLP and RAPD bands were 0.46 and 0.39, respectively. In both analyses only 15% of the bands were identified as rare and 7% as very frequent (more than 0.95).

In our study, SSRs proved to be more effective genetic markers than AFLPs and RAPDs because they were fast and reliable, locus-specific, multi-allelic, and therefore convenient for a variety of purposes including detection of unique genotypes and assessment of genetic diversity. Among the landraces analysed the distribution of rare alleles ranged from 0 to 11, with a mean of 3.1 per genotype. However genotypes of *T. aestivum* from India (k-23974) and Pakistan (k-30665), *T. compactum* from China (k-44098) and *T. spelta* from Switzerland (k-24709) were the most unique and were characterised by a combination of more than six rare alleles. On the whole, distribution of the rare alleles over regions and among species was different. For example, they were two times more frequent among landraces from India and China than from Transcaucasia. Among the different species of wheat, *T. spelta* carried the most rare alleles at 4.4 per accession, followed by *T. compactum* with 2.8, while *T. aestivum* from Europe and Asia carried 2 and 2.6, rare alleles per accession, respectively.

The similarity matrices computed for all pairs of accessions based on SSR analysis using the Dice coefficient ranged from 0.00 (between 36 pairs of landraces) to 0.95 (between the two *T. vavilovii* genotypes) and averaged 0.19. Variation limits of the coefficient were 0.43-0.98, 0.43-0.93 and 0.43-0.95 with means of 0.69, 0.67 and 0.60 for AFLP, RAPD and ALL data analyses, respectively. Both Cluster Analysis (CA) and Principal Component Analysis (PCA) of SSR data indicated a complex pattern of genetic relationships between landraces. Using PCA, 24 groups were distinguished while CA identified 18 clusters (Fig.1). At the highest hierarchical level two families of clusters were revealed combining landraces either from Europe or Asia. Similar division of the accessions was revealed in our previous study (Strelchenko et al., 2003) using AFLP and RAPD markers. The SSR based dendrogram shows that, at a finer level, genotypes of European wheat combined into five clusters (A–E). These clusters included all landraces from the PCA based groups 1, 10, 12, 14-16, 19 and also landraces from groups 11, 13 and 18 which had maximal values for their factor loadings. However the majority of these landraces were not differentiated based on AFLP, RAPD and ALL data and are included into one group (1). Using SSR data the landraces from Asia are combined into 13 clusters (F–R) using CA and 14 groups (2–9, 17, 20–24) using PCA. There was a good similarity between separate clusters and majority of the PCA based groups (cluster F corresponded to group 2, G – 4, H – 5, J – 9, K – 17, L – 3, N – 20, P – 22 and R – 24). Fig.1 also illustrates the similarity between some of the groups revealed by different DNA markers (group 2 included landraces of *T. aestivum* and *T. sphaerococcum* from India; 4 – *T. macha*; 5 and 8 – *T. spelta* from Switzerland and Central Asia, respectively; 9 – *T. aestivum* from China). The total number of PCA based groups decreased as follow: SSR (24 groups) – AFLP (16) – RAPD (12) – ALL (10).

All three types of DNA markers allowed for a classification of the landraces in a way that reflected their geographic origin. For example, all marker systems showed a division between European and Asian wheats. Such a division has similarities to the classification of common wheat being divided into European and Asian "race groups" performed by Vavilov (1922-1923) or into subspecies – by Flaksberger (1935) and Dorofeev et al. (1979). It indicates primary evolutionary directions related to the historic and geographic distribution of hexaploid wheat in Europe and Asia. Furthermore, within Asia, wheat landraces originating from India, China and Central Asia all grouped independently.

The classification based on the combined set of 486 bands from different marker systems was the simplest and divided the landraces with a high degree of concurrence to geographical origin.
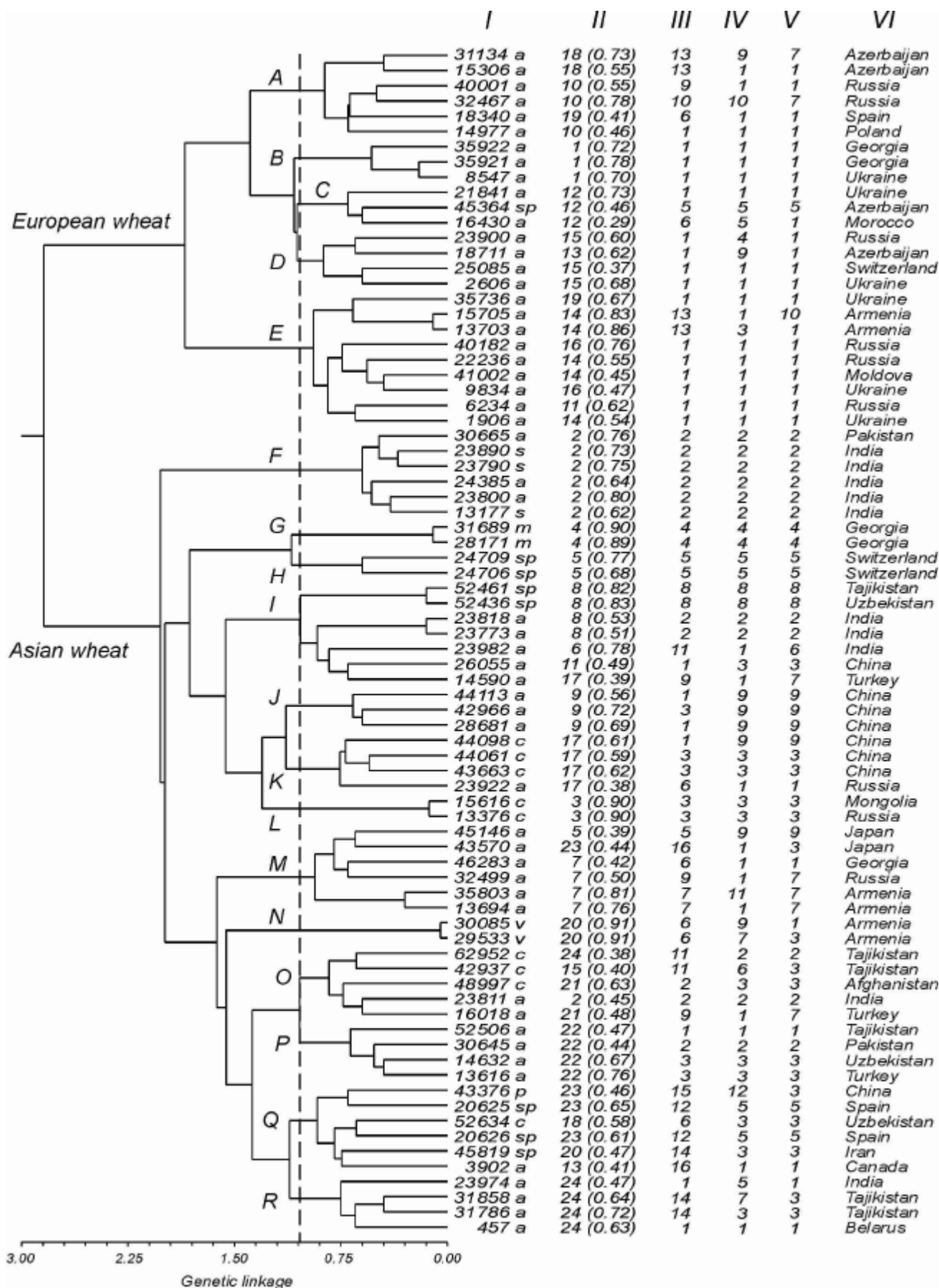
	I	II	III	IV	V	VI
A	31134 a	18 (0.73)	13	9	7	Azerbaijan
	15306 a	18 (0.55)	13	1	1	Azerbaijan
	40001 a	10 (0.55)	9	1	1	Russia
	32467 a	10 (0.78)	10	10	7	Russia
	18340 a	19 (0.41)	6	1	1	Spain
	14977 a	10 (0.46)	1	1	1	Poland
B	35922 a	1 (0.72)	1	1	1	Georgia
	35921 a	1 (0.78)	1	1	1	Georgia
	8547 a	1 (0.70)	1	1	1	Ukraine
C	21841 a	12 (0.73)	1	1	1	Ukraine
	45364 sp	12 (0.46)	5	5	5	Azerbaijan
	16430 a	12 (0.29)	6	5	1	Morocco
	23900 a	15 (0.60)	1	4	1	Russia
D	18711 a	13 (0.62)	1	9	1	Azerbaijan
	25085 a	15 (0.37)	1	1	1	Switzerland
	2606 a	15 (0.68)	1	1	1	Ukraine
	35736 a	19 (0.67)	1	1	1	Ukraine
E	15705 a	14 (0.83)	13	1	10	Armenia
	13703 a	14 (0.86)	13	3	1	Armenia
	40182 a	16 (0.76)	1	1	1	Russia
	22236 a	14 (0.55)	1	1	1	Russia
	41002 a	14 (0.45)	1	1	1	Moldova
	9834 a	16 (0.47)	1	1	1	Ukraine
	6234 a	11 (0.62)	1	1	1	Russia
	1906 a	14 (0.54)	1	1	1	Ukraine
F	30665 a	2 (0.76)	2	2	2	Pakistan
	23890 s	2 (0.73)	2	2	2	India
	23790 s	2 (0.75)	2	2	2	India
	24385 a	2 (0.64)	2	2	2	India
	23800 a	2 (0.80)	2	2	2	India
	13177 s	2 (0.62)	2	2	2	India
G	31689 m	4 (0.90)	4	4	4	Georgia
	28171 m	4 (0.89)	4	4	4	Georgia
H	24709 sp	5 (0.77)	5	5	5	Switzerland
	24706 sp	5 (0.68)	5	5	5	Switzerland
I	52461 sp	8 (0.82)	8	8	8	Tajikistan
	52436 sp	8 (0.83)	8	8	8	Uzbekistan
	23818 a	8 (0.53)	2	2	2	India
	23773 a	8 (0.51)	2	2	2	India
	23982 a	6 (0.78)	11	1	6	India
	26055 a	11 (0.49)	1	3	3	China
	14590 a	17 (0.39)	9	1	7	Turkey
J	44113 a	9 (0.56)	1	9	9	China
	42966 a	9 (0.72)	3	9	9	China
	28681 a	9 (0.69)	1	9	9	China
	44098 c	17 (0.61)	1	9	9	China
K	44061 c	17 (0.59)	3	3	3	China
	43663 c	17 (0.62)	3	3	3	China
	23922 a	17 (0.38)	6	1	1	Russia
L	15616 c	3 (0.90)	3	3	3	Mongolia
	13376 c	3 (0.90)	3	3	3	Russia
	45146 a	5 (0.39)	5	9	9	Japan
M	43570 a	23 (0.44)	16	1	3	Japan
	46283 a	7 (0.42)	6	1	1	Georgia
	32499 a	7 (0.50)	9	1	7	Russia
	35803 a	7 (0.81)	7	11	7	Armenia
N	13694 a	7 (0.76)	7	1	7	Armenia
	30085 v	20 (0.91)	6	9	1	Armenia
	29533 v	20 (0.91)	6	7	3	Armenia
O	62952 c	24 (0.38)	11	2	2	Tajikistan
	42937 c	15 (0.40)	11	6	3	Tajikistan
	48997 c	21 (0.63)	2	3	3	Afghanistan
	23811 a	2 (0.45)	2	2	2	India
	16018 a	21 (0.48)	9	1	7	Turkey
P	52506 a	22 (0.47)	1	1	1	Tajikistan
	30645 a	22 (0.44)	2	2	2	Pakistan
	14632 a	22 (0.67)	3	3	3	Uzbekistan
	13616 a	22 (0.76)	3	3	3	Turkey
	43376 p	23 (0.46)	15	12	3	China
	20625 sp	23 (0.65)	12	5	5	Spain
Q	52634 c	18 (0.58)	6	3	3	Uzbekistan
	20626 sp	23 (0.61)	12	5	5	Spain
	45819 sp	20 (0.47)	14	3	3	Iran
	3902 a	13 (0.41)	16	1	1	Canada
	23974 a	24 (0.47)	1	5	1	India
R	31858 a	24 (0.64)	14	7	3	Tajikistan
	31786 a	24 (0.72)	14	3	3	Tajikistan
	457 a	24 (0.63)	1	1	1	Belarus

European wheat

Asian wheat

Genetic linkage: 3.00 2.25 1.50 0.75 0.00

Figure 1. Dendrogram based on SSR data shows the distribution of 78 hexaploid wheat landraces.
(*I*) VIR catalogue numbers and species codes: a – *T. aestivum*, c – *T. compactum*, m – *T. macha*, p – *T. petropavlovskyi*, s – *T. sphaerococcum*, sp – *T. spelta* and v – *T. vavilovii*; (*II*) groups defined using PCA based on SSR data and factor loadings; (*III – V*) groups derived using PCA based on AFLP, RAPD and ALL data, respectively; (*IV*) countries of origin.

Thus group 1 included mainly landraces of *T. aestivum* from Europe, 2 – *T. aestivum* and *T. sphaerococcum* from India and Pakistan, 3 – *T. aestivum* and *T. compactum* from Central Asia, 4 – *T. macha*, 5 – *T. spelta* from Europe, 7 – *T. aestivum* from Caucasus mountains region, 8 – *T. spelta* from Central Asia, 9 – *T. aestivum* and *T. compactum* from China and Japan. In all cases the distinctions between clusters and groups were determined by differences in the frequency of a considerable number of alleles or amplified DNA fragments. Classifications based on DNA markers differ from existing taxonomic divisions of hexaploid wheats (Mac Key 1966; Dorofeev et al.1979; Miller 1987). Probably they reflect the genetic differentiation of hexaploid wheats more correctly because they assess distinctions revealed on the large number of genome loci while botanical classifications take into consideration morphological traits controlled by single genes.

Conclusion

The present study demonstrated the usefulness of SSR, AFLP, RAPD and a combined set of markers for the assessment of genetic diversity and the relationships between wheat landraces. However the microsatellite markers were shown to be superior for detecting genotypes with combinations of rare alleles. It was clearly demonstrated that genetic differentiation was closely correlated with geographical origin. The accessions were combined into groups of either European or Asian origin. Differences between them and clusters of accessions within Asian wheat were mainly defined by differences in the frequency of the SSR defined alleles and amplified DNA fragments in AFLP or RAPD analyses. It was pointed out that there was incongruity between the taxonomic division of hexaploid wheats and relationships of the landraces revealed by DNA markers.

Acknowledgements

References

Dorofeev VF, Filatenko AA, Migushova EF, Udachin RA and Jakubtziner MM (1979). Flora of cultivated plants. (Eds. V.F. Dorofeev and O.N. Korovina) Vol. 1. Wheat. (Kolos, Leningrad). (in Russian)

Flaksberger KA (1935). Flora of cultivated plants. (Ed. E.V. Wulff) Vol. I. Cereals. Wheat. (State agricultural publishing company, Moscow and Leningrad). (in Russian)

Guyomarc'h H, Sourdille P, Charmet G, Edwards KJ and Bernard M (2002). Characterisation of polymorphic markers from *T. tauschii* and transferability to the D-genome of bread wheat. Theor Appl Genet 104: 1164-1172.

Mac Key J (1966). Species relationships in *Triticum*. Hereditas [Suppl.] 2: 237-276.

Miller TE (1987). Systematics and evolution. In 'Wheat Breeding: Its Scientific Basis'. (Ed. F. G. H. Lupton), pp. 1-30. (Chapman and Hall, London).

Palmova EF (1935). Introduction in wheats ecology. (Ogiz Selkhozgiz, Leningrad and Moscow). (in Russian)

Röder MS,Korzun V, Wendehake K, Plaschke J, Tixier MH, Leroy P and Ganal MW (1998) A microsatellite map of wheat. Genetics 149: 2007-2023.

Roussel V., Koenig J., Beckert, M. and Balfourier F (2004) Molecular diversity in French bread wheat accessions related to temporal trends and breeding programmes. Theor. Appl. Genet. (published on line)

Strelchenko P, Street K, Mitrofanova O, Mackay M, Chabane K, Valkoun J (2003). The genetic relationships between hexaploid wheat landraces from different geographical origin. In 'Proc 10th Int Wheat Gen Symp. September 1-6, 2003, Paestum, Italy: vol 2, pp. 637-640.

Vavilov NI (1922-1923). A contribution to the classification of soft wheats – *Triticum vulgare* Vill. (Essays on systematics and geography of cereals). Bull Appl Bot Plant Breed, Petrograd (St. Petersburg) 13: 215-257.

Vavilov NI (1964). World resources of cereals, legumes, flax cultivars and their utilization in breeding. Wheat. (Nauka, Moskow and Leningrad). (in Russian)