

УДК 575.22:631.52:633.111.1

М. Н. ШАПТУРЕНКО<sup>1</sup>, С. В. ВАКУЛА<sup>1</sup>, В. КОРЗУН<sup>2</sup>, академик Л. В. ХОТЫЛЕВА<sup>1</sup>

## SNP-АНАЛИЗ ГЕНЕТИЧЕСКОГО РАЗНООБРАЗИЯ ПШЕНИЦЫ БЕЛАРУСИ

<sup>1</sup>Институт генетики и цитологии НАН Беларуси, Минск, Беларусь  
m.shapturenko@igc.by; svettera@yandex.ru; l.khotyleva@igc.by

<sup>2</sup>KWS LOCHOW GmbH, Берген, Германия  
viktor.korzun@kws-lochow.de

Проведено изучение генетического разнообразия пшеницы Беларуси на основе высокопропускного SNP маркирования. В целом из 384 использованных маркеров в исследуемой коллекции типирован 331 локус. Генофонды озимых и яровых сортов достоверно различаются по частотам 248 вариантов 174 SNP. Генетическая структура белорусских сортов обнаруживает значительное сходство с образцами российской и украинской селекции, но при этом обладает значительным запасом разнообразия, которое представляет хороший потенциал для создания новых высокопродуктивных адаптированных форм.

**Ключевые слова:** пшеница (*Triticum aestivum* L.), генетическое разнообразие, SNP-маркеры.

M. N. SHAPTURENKO<sup>1</sup>, S. V. VAKULA<sup>1</sup>, V. KORZUN<sup>2</sup>, L. V. KHOTYLEVA<sup>1</sup>

## HIGH-THROUGHPUT SNP ARRAY FOR GENETIC DIVERSITY EVALUATION WITHIN HEXAPLOID WHEAT IN BELARUS

<sup>1</sup>Institute of Genetics and Cytology of the National Academy of Sciences of Belarus, Minsk, Belarus  
m.shapturenko@igc.by; svettera@yandex.ru; l.khotyleva@igc.by

<sup>2</sup>KWS LOCHOW GmbH, Bergen, Germany  
viktor.korzun@kws-lochow.de

We used a high-throughput array to evaluate the diversity of hexaploid wheat growing in Belarus under the breeding program through 384 gene-associated SNPs. The gene pool of winter and spring varieties are significantly different in frequency of 248 variants from 174 SNPs. The genetic structure of the Belarusian population of wheat has appeared similar to the Russian and Ukrainian varieties and is essentially different from west-European varieties. But it has a high variability and, consequently, a good genetic potential for the improvement through breeding.

**Keywords:** wheat (*Triticum aestivum* L.), genetic diversity, SNP-markers.

**Введение.** Мягкая пшеница *Triticum aestivum* L. ( $2n = 42$ ) является естественным аллополиплоидом с геномной формулой ВВААDD [1]. Ее геном содержит около  $17,3 \cdot 10^9$  п. н. ДНК, из которых 80 % приходится на повторяющиеся последовательности [2].

Генетическое разнообразие пшеницы служит источником вариаций для селекции, обуславливая создание новых форм хозяйственно ценных растений с улучшенными свойствами. Узкая генетическая основа снижает эффективность селекции, поскольку не позволяет преодолеть уязвимость к неблагоприятным факторам и ограничивает возможности комбинаторики наследственного материала при гибридизации. В связи с этим важное значение приобретает сохранение генетических ресурсов и оценка их разнообразия для дальнейшего использования в практических целях [3].

В Беларуси пшеница является одной из основных зерновых культур наряду с ячменем, рожью и тритикале. Преимущественно сельскохозяйственное производство обеспечивает потребности государства в зерне, за исключением высококачественных элитных сортов. В последнее десятилетие структура зерновых культур значительно изменилась в пользу высокоурожайных сортов пшеницы на 19 % (с 20,1 до 32,1 %) за счет сокращения посевных площадей ржи [4]. Это вызвало изменения общей урожайности зерновых культур. В 2014 г. площадь под пшеницей в Беларуси составила около 500 тыс. га, две трети из которых были засеяны озимыми сортами.

В до- и послевоенные годы, вплоть до середины 1980-х годов, селекция пшеницы в Беларуси практически не развивалась [5], но ежегодно около 100 тыс. га засеивались российскими, украинскими и польскими сортами. Производство местной пшеницы имело низкую конкурентоспособность за счет экспорта высококачественного зерна сильной пшеницы из различных регионов бывшего СССР [6]. Селекция пшеницы в Беларуси начала интенсивно развиваться лишь после распада СССР. На основе гибридизации и отбора были созданы перспективные местные сорта, обогащенные ценным генетическим материалом зарубежных форм.

На сегодняшний день многочисленные местные и иностранные образцы составляют основу белорусского генофонда пшеницы, но при этом имеется недостаточно информации об их разнообразии и возможных генетических связях. Настоящая работа нацелена на оценку генетического пула возделываемой в Беларуси пшеницы с использованием высокопропускной технологии SNP-маркирования.

**Материалы и методы исследования.** Проводили сравнительный анализ 64 образцов гексаплоидной пшеницы белорусской селекции и 27 сортов, представляющих разнообразие 5 европейских стран (Украина, Германия, Россия, Польша, Великобритания).

Для оценки генетической структуры коллекции мы использовали 384 ген-ассоциированных SNP, отобранных из нескольких тысяч полиморфизмов, типизируемых с использованием технологии 9K iSelect по принципу равномерного распределения по геному и максимального отражения генетической дифференциации согласно Cavanagh и соавт. (2013) [7].

Для обработки первичных данных SNP-типирования использовали программный пакет GENEPOP [8]. Для биаллельных SNP маркеров рассчитывали: частоты минорных аллелей (MAF – minor allele frequency); наблюдаемую гетерозиготность (He); информативность (PIC – polymorphism information content); индивидуальное ( $1-Q_{intra}$ ) и групповое ( $\pi$ ,  $1-Q_{inter}$ ) генетическое разнообразие, индивидуальный (Fis) и субпопуляционный (Fst) индексы фиксации, логарифмическое правдоподобие ( $G$ ) [9–12].

Кластерный анализ выполняли при помощи программного пакета DarWin 3.0.

**Результаты и их обсуждение.** Согласно определению Brookes (1999) [13], SNPs (single nucleotide polymorphisms) – это однонуклеотидные позиции в геномной ДНК, для которых в некоторой популяции имеются различные варианты (аллели), причем редкий аллель встречается с частотой выше 1 %. Ограничения по частоте встречаемости противопоставляют SNP точечным мутациям и предполагают их использование в качестве генетических маркеров. По сравнению с другими молекулярными методами оценки полиморфизма (RAPD, SSR, ISSR, AFLP) использование SNP-маркеров является наиболее прогрессивным и информативным подходом, практически полная автоматизация которого дает возможность проанализировать одновременно несколько тысяч SNP.

В нашем исследовании из 384 использованных SNP в общей сложности генотипирован 331 (86 %). Из них 295 с известной локализацией, распределенные по 21 хромосоме, и 36 с неустановленной локализацией, которые в совокупности представили варибельность 97 % локусов (таблица). Выявленные полиморфизмы включали 80,8 % транзиций и 19,2 % трансверсий. Покрываемость хромосом маркерами ранжировалась от 1 для 4D до 26 для 5B с последующим распределением по геномам: А-геном – 129, В-геном – 138 и D-геном – 28 SNP. Распределение маркеров по гомеологичным группам хромосом в основном было равномерным, за исключением 4-й группы.

Большинство SNP выявляли высокий полиморфизм. Однако наблюдались некоторые различия в его изменчивости между яровой и озимой пшеницей. Самый высокий уровень полиморфизма выявлен в группе озимой пшеницы. Но наибольшая вариация по геномам и гомеологичным группам хромосом характеризует яровые формы.

Показатели информационного содержания, наблюдаемые в нашем исследовании, значимо колебались с высокой долей маркеров со значениями PIC, превышающими 0,35 (61,9 %) (рис. 1). Также наблюдалось их варьирование по хромосомам и геномам. Наиболее информативные SNP ( $PIC > 0,4$ ) были локализованы на хромосомах 7D, 6B и 3A.

Средние значения PIC для маркеров с известной хромосомной локализацией также были рассчитаны отдельно для каждой из подгрупп в соответствии с типом вегетации. В результате мы

**Распределение SNP маркеров по геномам и гомеологичным группам хромосом**

Хромосома	Кол-во SNP	% от общего числа	Яровая пшеница		Озимая пшеница		Общий полиморфизм	
			Полиморфные SNPs, %	Мономорфные SNPs, %	Полиморфные SNPs, %	Мономорфные SNPs, %		
А-геном	1	20	6,0	90	10	100	0	100
	2	14	4,2	71,4	28,6	100	0	100
	3	19	5,7	100	0	100	0	100
	4	15	4,5	100	0	100	0	100
	5	21	6,3	90,5	9,5	100	0	100
	6	17	5,1	88,2	11,8	100	0	100
	7	23	6,9	82,6	17,4	86,9	13	95,7
Итого	129	39,0	89,0	11,0	98,1	1,9	99,4	
В-геном	1	21	6,3	90,5	9,5	100	0	100
	2	25	7,6	84	16	96	4	96
	3	25	7,6	92	8	96	4	96
	4	11	3,3	100	0	100	0	100
	5	26	7,9	92,3	7,7	92,3	7,7	92,3
	6	17	5,1	94,1	5,9	100	0	100
	7	13	3,9	92,3	7,7	100	100	100
Итого	138	41,7	92,2	7,8	97,8	16,5	97,8	
D-геном	1	7	2,1	85,7	14,3	85,7	14,3	85,7
	2	6	1,8	83,3	16,7	100	0	100
	3	3	0,9	66,7	33,3	100	0	100
	4	1	0,3	100	0	100	0	100
	5	2	0,6	50	50	100	0	100
	6	4	1,2	50	50	75	25	75
	7	5	1,5	80	20	100	100	100
Итого	28	8,5	73,7	26,3	94,4	19,9	94,4	
НЛ	36	10,9	77,8	22,2	97,2	2,8	97,2	
Средняя			83,2	16,9	96,9	10,3	97,2	

Примечание. НЛ – маркеры с неустановленной локализацией.

выяснили, что наблюдаемый уровень информационного содержания варьирует не только в зависимости от хромосомной локализации, но также зависит от анализируемой подгруппы. Наиболее высокий уровень выявлен у озимой пшеницы для SNP, покрывающих 7D и 2A хромосомы, тогда как для яровой наиболее информативные маркеры были локализованы на 4A и 7A хромосомах.

Наблюдаемая частота минорных аллелей (MAF) была смещена в сторону аллелей с  $MAF > 0,25$ . Из 331 маркера, только 8 (2,4 %) имели  $MAF \leq 0,05$  и 236 (71,3 %) –  $MAF \geq 0,2$ , что подтвердило

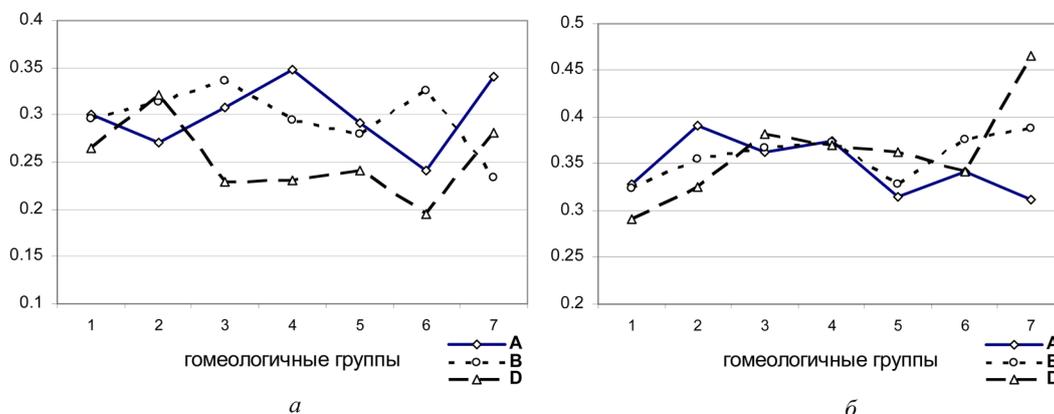


Рис. 1. Распределение средних значений PIC по гомеологичным группам хромосом и геномам у яровой (а) и озимой (б) пшеницы

их высокую дифференцирующую способность для тестирования сортов пшеницы. Основные локусы с редкими аллелями расположены на хромосомах 3D, 4B, 5A, 6D.

Распределение MAF в подгруппах также выявило различия (рис. 2): у озимой пшеницы, характеризующейся высоким общим полиморфизмом SNP, были представлены высокие частоты аллелей с MAF более 0,25, тогда как для яровой пшеницы отмечено преобладание мономорфных маркеров и распространенность SNP с MAF менее 0,25.

Распределение MAF в подгруппах и генеральной совокупности значимо коррелирует ( $r = 0,53$ ;  $0,85$  для яровой и озимой пшеницы соответственно). Слабая связь ( $r = 0,29$ ) MAF между озимыми и яровыми формами указывает на дифференциальное накопление SNP в результате длительной селекции этой культуры. Мы нашли 72,5 % совпадения аллелей озимых и яровых образцов, в то время как 27,5 % локусов выявляли переключение частот. Таким образом, наши результаты показали некоторые общие и специфические закономерности разнообразия гаплотипов у яровых и озимых образцов, что предполагает влияние различных сил отбора, которые, возможно, действовали в определенных областях генома пшеницы на протяжении селекции.

Усредненное генетическое разнообразие по индивидуальным сортам (1-Qintra) составило 0,19, статистика изменчивости по отдельным локусам (1-Qintra среднее на локус SNP) немного ниже, но, с учетом округления, также составляет 0,19. Незначительный перевес гетерозиготных локусов в геноме яровых пшениц в сравнении с озимыми отражает более низкое значение медианы признака ( $1-Qintra^{\text{Мед-Озимые}} = 0,05$ ;  $1-Qintra^{\text{Мед-Яровые}} = 0,04$ ).

Полученные результаты показывают, что генетическое разнообразие ( $\pi$ ) яровых пшениц ниже как общего значения всей совокупности (0,37), так и показателя озимых пшениц (0,36). В работе С. R. Cavanagh и соавт. (2013) значения  $\pi$  составляли для культурных сортов 0,36 и для ландрас – 0,34. Таким образом, генетическое разнообразие, представленное белорусским генофондом, согласуется с данными о «мировом» уровне генетического разнообразия пшеницы. Распределение индексов Fis соответствует более высокой доле низкочастотных минорных аллелей и высокой гетерозиготности яровых пшениц.

G-тест логарифмического правдоподобия позволил оценить достоверность различий распределения частот аллелей озимых и яровых сортов пшениц. Субпопуляции достоверно различаются по 248 (77 %) из 322 полиморфных локусов. Среднее значение индексов Fst и Fit составляют 0,16 и 0,53 соответственно. Генетическая дифференциация высокодостоверна и значима.

Оценка аллельного состава генома пшеницы и последующий кластерный анализ с использованием метода ближайших соседей (neighbor-joining) позволили провести простую дифференциацию, которая отражает деление образцов коллекции по типу озимость/яровость, что также указывает на связь отдельных SNP с данным признаком (рис. 3).

Большинство корней иерархического дерева имеют бутстрэп значение более 50 %. Кластер 1 (озимая пшеница) подразделяется на две подгруппы (I-a, I-б). Первая (I-a) представлена белорусскими сортами, которые были введены в сельскохозяйственное производство до 2000 г. Эта подгруппа также включает российский сорт Льговская-4 и несколько украинских образцов, которые формируют топологически дискретные подкластеры.

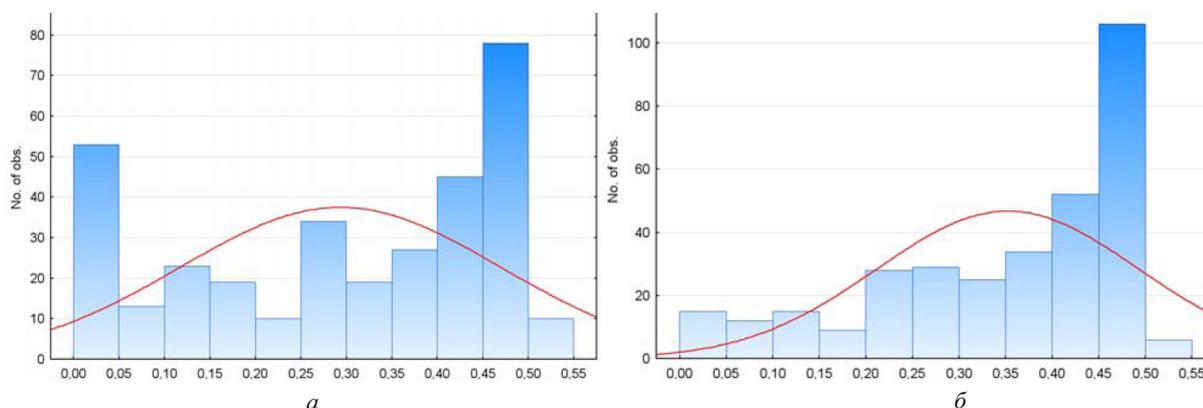


Рис. 2. Распределение частот минорных аллелей (MAF) у яровой (а) и озимой (б) пшеницы

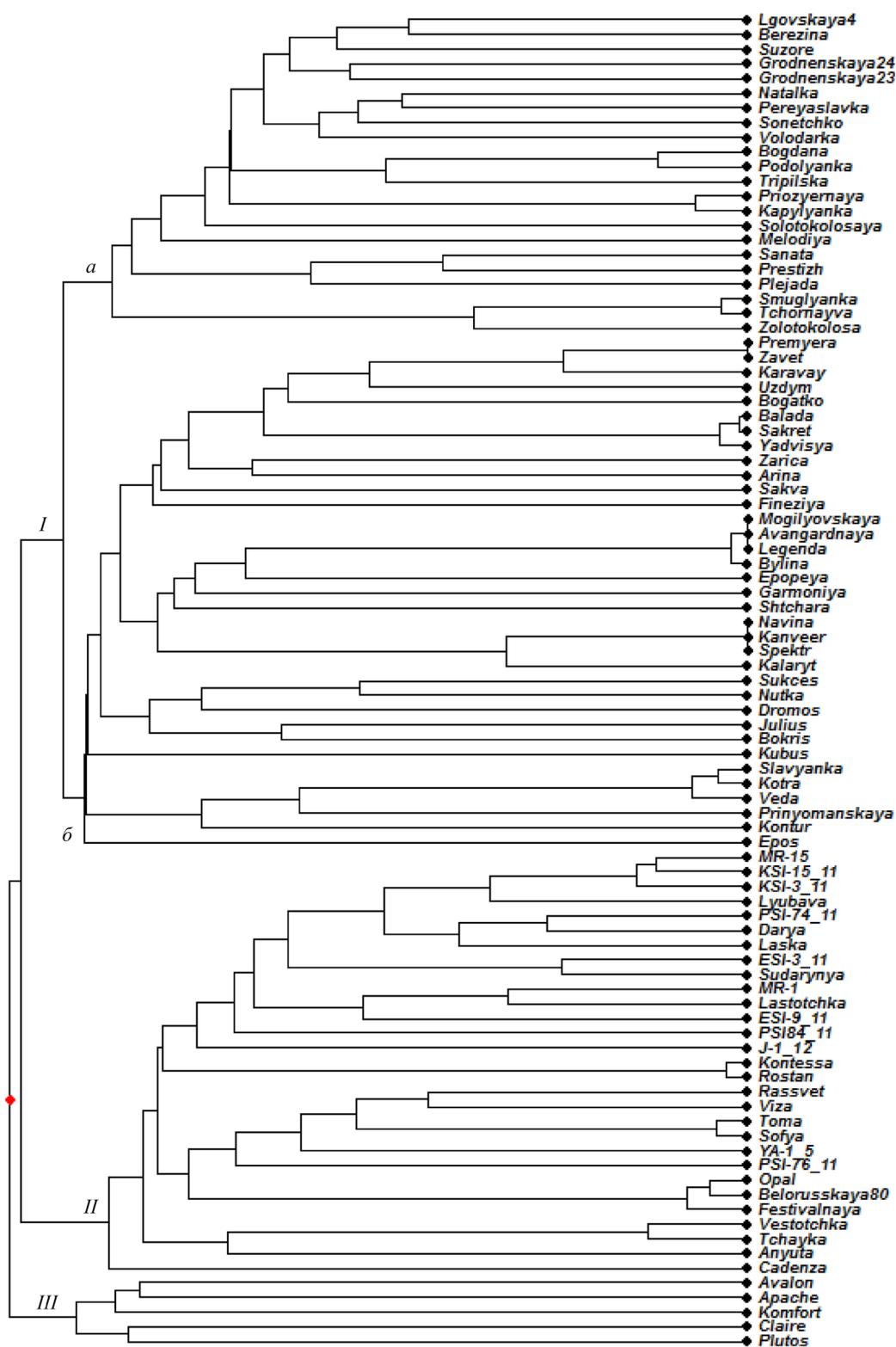


Рис. 3. Топологическая дифференциация образцов пшеницы, выполненная по данным SNP анализа

Вторую подгруппу (I-б) составляют белорусские сорта, введенные в сельскохозяйственное производство после 2000 г., а также образцы польской (Sukcess, Nutka, Кобра, Богатко, Саква, Finezia) и немецкой (Dromos, Bokris, Kubus) селекции. Интересен факт обнаружения генетически идентичных белорусских сортов, которые, возможно, получены из сестринских линий, характеризующихся высокой генетической общностью.

Кластер 2 включает разнообразие яровой пшеницы, преимущественно белорусской селекции, а также шотландский образец Cadenza и немецкий сорт Opal, который был использован при создании отечественных сортов Белорусская 80 и Фестивальная. Последний получен в результате отбора эуплоидного потомства моносомных растений пшеницы Opal.

Кластер 3 образуют современные высокоурожайные сорта английской, шотландской и немецкой селекции, а также один из недавно созданных белорусскими селекционерами сортов – Комфорт.

**Заключение.** Результаты проведенного исследования позволяют сделать вывод о высоком генетическом разнообразии пшеницы, возделываемой на территории Беларуси, что связано с особенностями отечественной селекции этой культуры, которая начала активно развиваться только в 1980-е годы. Посевные площади Беларуси того времени были заняты российскими и украинскими сортами, которые активно использовались в селекционной работе. Это подтверждается данными SNP маркирования, на основании которого отечественные сорта раннего периода (сортосмена 1985–2000 гг.) объединены в общую подгруппу (I-a) с украинскими сортами.

Период сортосмены 2000–2013 гг. характеризуется активным привлечением в селекционную работу генетического разнообразия Польши, Германии и других европейских стран. Вместе с тем проявляется тенденция использования близкородственного материала и, вероятно, сестринских линий, о чем свидетельствует наличие групп полной идентичности.

Действительно, на сегодняшний день селекционные центры Беларуси имеют хороший потенциал для улучшения пшеницы, но разнообразие должно быть систематизировано на платформе ДНК-типирования, которое может обеспечить более эффективное использование существующих генетических ресурсов в селекции.

### Список использованной литературы

1. *Feldman, M.* Origin of cultivated wheat / M. Feldman // The world wheat book: a history of wheat breeding / eds. by A. P. Bonjean, W. J. Angus. – Paris, France: Lavoisier Publishing, 2001. – P. 3–56.
2. *Bennett, M. D.* Nuclear DNA amounts in angiosperms / M. D. Bennett, I. J. Leitch // Ann. Bot. – 1995. – Vol. 76. – P. 113–176.
3. *Börner, A.* Preservation of plant genetic resources in the biotechnology era / A. Börner // Biotechnology J. – 2006. – № 1. – P. 1393–1404.
4. Перепланировка посевных площадей: суть и особенности реформы // Есть вопрос: архив интервью, статей и заметок [Электронный ресурс]. – Режим доступа: <http://vorposy.info/?p=675>. – Дата доступа: 15.08.2016.
5. *Гриб, С. И.* Прогресс в селекции яровой пшеницы в Беларуси / С. И. Гриб // Весці НАН Беларусі. Сер. аграрн. навук. – 2009. – № 3. – С. 37–41.
6. *Коптик, И. К.* Научно-методические подходы и результаты в селекции озимой мягкой пшеницы (*Triticum aestivum* L.) в Республике Беларусь / И. К. Коптик // Весці НАН Беларусі. Сер. аграрн. навук. – 2010. – № 1. – С. 47–54.
7. Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars / C. R. Cavanagh [et al.] // PNAS. – 2013. – Vol. 110. – P. 8057–8062.
8. *Rousset, F.* genepop'007: a complete re-implementation of the genepop software for Windows and Linux / F. Rousset // Mol. Ecol. Resources. – 2008. – Vol. 8. – P. 103–106.
9. *Wright, S.* Evolution and the Genetics of Populations / S. Wright. – Chicago: Univ. of Chicago Press, 1969. – Vol. 2: the Theory of Gene Frequencies. – P. 294–295.
10. *Bryc, K.* Novel Approach to Estimating Heterozygosity from Low-Coverage Genome Sequence / K. Bryc, N. Patterson, D. Reich // Genetics. – 2013. – Vol. 195. – P. 553–561.
11. Construction of a genetic linkage map in man using restriction fragment length polymorphisms / D. Botstein [et al.] // Am. J. Hum. Genet. – 1980. – Vol. 32. – P. 314–331.
12. *Weir, B. S.* Estimating F-Statistics for the Analysis of Population-Structure / B. S. Weir, C. C. Cockerham // Evolution. – 1984. – Vol. 38. – P. 1358–1370.
13. *Brookes, A. J.* The essence of SNPs / A. J. Brookes // Gene. – 1999. – Vol. 234. – P. 177–186.

Поступило в редакцию 13.06.2016