

Английский текст <https://vavilov.elpub.ru/jour>

## Автоматическое фенотипирование морфологии колоса тетра- и гексаплоидных видов пшеницы методами компьютерного зрения

А.Ю. Пронозин<sup>1</sup>, А.А. Паулиш<sup>2</sup>, Е.А. Заварзин<sup>2</sup>, А.Ю. Приходько<sup>2</sup>, Н.М. Прохошин<sup>2</sup>, Ю.В. Кручинина<sup>1, 3</sup>, Н.П. Гончаров<sup>1, 4</sup>, Е.Г. Комышев<sup>1, 2, 3</sup>, М.А. Генаев<sup>1, 2, 3</sup> 

<sup>1</sup> Федеральное исследовательское учреждение Института цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия

<sup>2</sup> Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия

<sup>3</sup> Курчатовский геномный центр Института цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия

<sup>4</sup> Новосибирский государственный аграрный университет, Новосибирск, Россия

 [mag@bionet.nsc.ru](mailto:mag@bionet.nsc.ru)

**Аннотация.** Внутривидовая классификация культурных растений необходима для эффективного сохранения биологического разнообразия видов, изучения их происхождения, определения филогении и проведения межвидовой гибридизации при селекции. Современные возделываемые виды пшениц произошли от трех диких диплоидных предков в результате гибридизации и нескольких раундов удвоения геномов и представлены ди-, тетра- и гексаплоидными видами. Поэтому идентификация плоидности пшениц и определение их геномного состава являются одними из основных этапов их классификации на основе визуального анализа фенотипических признаков колоса. Цель работы – исследование морфологических характеристик колосов полиплоидных видов пшеницы методами высокопроизводительного фенотипирования. Выполнено фенотипирование количественных характеристик колоса 17 видов пшеницы (595 растений, 3348 изображений), включая восемь тетраплоидных: *Triticum aethiopicum*, *T. dicoccoides*, *T. dicoccum*, *T. durum*, *T. militinae*, *T. polonicum*, *T. timopheevii*, *T. turgidum* и девять гексаплоидных: *T. compactum*, *T. aestivum* (в том числе изогенная линия сорта Новосибирская 67 АНК-23), *T. antiquorum*, *T. spelta* (включая стародавний сорт *T. spelta* Rother Sommer Kolben), *T. petropavlovskiyi*, *T. yunnanense*, *T. macha*, *T. sphaerococcum*, *T. vavilovii*. Морфология колоса описана на основе девяти количественных признаков, включающих форму, размер и остистость. Признаки были получены в результате анализа цифровых изображений с помощью программы WERecognizer. Кластерный анализ растений по характеристикам формы колоса и сравнение их распределений у тетра- и гексаплоидных видов показали более высокую вариабельность признаков у гексаплоидных видов по сравнению с тетраплоидными. При этом сами виды в пространстве характеристик колоса формируют два кластера. К первому относятся преимущественно гексаплоидные виды, за исключением одного тетраплоидного, дикорастущего *T. dicoccoides*, ко второму – тетраплоидные, за исключением трех гексаплоидных, *T. compactum*, *T. antiquorum*, *T. sphaerococcum*, и i:АНК-23. Показано, что морфологические характеристики колосов для гекса- и тетраплоидных видов, полученные на основе компьютерного анализа изображений, демонстрируют различия, которые в дальнейшем могут быть использованы для разработки методики эффективной классификации растений по плоидности и их видовой принадлежности в автоматическом режиме. Ключевые слова: пшеница; морфология колоса; феномика; обработка изображений; компьютерное зрение; машинное обучение; биотехнологии.

**Для цитирования:** Пронозин А.Ю., Паулиш А.А., Заварзин Е.А., Приходько А.Ю., Прохошин Н.М., Кручинина Ю.В., Гончаров Н.П., Комышев Е.Г., Генаев М.А. Автоматическое фенотипирование морфологии колоса тетра- и гексаплоидных видов пшеницы методами компьютерного зрения. *Вавиловский журнал генетики и селекции*. 2021; 25(1):71-81. DOI 10.18699/VJ21.009

## Automatic morphology phenotyping of tetra- and hexaploid wheat spike using computer vision methods

A.Yu. Pronozin<sup>1</sup>, A.A. Paulish<sup>2</sup>, E.A. Zavarzin<sup>2</sup>, A.Yu. Prikhodko<sup>2</sup>, N.M. Prokhoshin<sup>2</sup>, Yu.V. Kruchinina<sup>1, 3</sup>, N.P. Goncharov<sup>1, 4</sup>, E.G. Komyshev<sup>1, 2, 3</sup>, M.A. Genayev<sup>1, 2, 3</sup> 

<sup>1</sup> Institute of Cytology and Genetics of Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia

<sup>2</sup> Novosibirsk State University, Novosibirsk, Russia

<sup>3</sup> Kurchatov Genomics Center of the Institute of Cytology and Genetics of Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia

<sup>4</sup> Novosibirsk State Agrarian University, Novosibirsk, Russia

 [mag@bionet.nsc.ru](mailto:mag@bionet.nsc.ru)

**Abstract.** Intraspecific classification of cultivated plants is necessary for the conservation of biological diversity, study of their origin and their phylogeny. The modern cultivated wheat species originated from three wild diploid ancestors as a result of several rounds of genome doubling and are represented by di-, tetra- and hexaploid species. The

identification of wheat ploidy level is one of the main stages of their taxonomy. Such classification is possible based on visual analysis of the wheat spike traits. The aim of this study is to investigate the morphological characteristics of spikes for hexa- and tetraploid wheat species based on the method of high-performance phenotyping. Phenotyping of the quantitative characteristics of the spike of 17 wheat species (595 plants, 3348 images), including eight tetraploids (*Triticum aethiopicum*, *T. dicoccoides*, *T. dicoccum*, *T. durum*, *T. militinae*, *T. polonicum*, *T. timopheevii*, and *T. turgidum*) and nine hexaploids (*T. compactum*, *T. aestivum*, i:ANK-23 (near-isogenic line of *T. aestivum* cv. Novosibirskaya 67), *T. antiquorum*, *T. spelta* (including cv. Rother Sommer Kolben), *T. petropavlovskiyi*, *T. yunnanense*, *T. macha*, *T. sphaerococcum*, and *T. vavilovii*), was performed. Wheat spike morphology was described on the basis of nine quantitative traits including shape, size and awns area of the spike. The traits were obtained as a result of image analysis using the WERecognizer program. A cluster analysis of plants according to the characteristics of the spike shape and comparison of their distributions in tetraploid and hexaploid species showed a higher variability of traits in hexaploid species compared to tetraploid ones. At the same time, the species themselves form two clusters in the visual characteristics of the spike. One type is predominantly hexaploid species (with the exception of one tetraploid, *T. dicoccoides*). The other group includes tetraploid ones (with the exception of three hexaploid ones, *T. compactum*, *T. antiquorum*, *T. sphaerococcum*, and i:ANK-23). Thus, it has been shown that the morphological characteristics of spikes for hexaploid and tetraploid wheat species, obtained on the basis of computer analysis of images, include differences, which are further used to develop methods for plant classifications by ploidy level and their species in an automatic mode.

Key words: wheat spike morphology; wheat; phenomics; image processing; computer vision; machine learning; biotechnology.

**For citation:** Pronozin A.Yu., Paulish A.A., Zavarzin E.A., Prikhodko A.Yu., Prokhoshin N.M., Kruchinina Yu.V., Goncharov N.P., Komyshev E.G., Genaev M.A. Automatic morphology phenotyping of tetra- and hexaploid wheat spike using computer vision methods. *Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding*. 2021;25(1):71-81. DOI 10.18699/VJ21.009

## Введение

Ряд важных вопросов, включая аспекты эффективного сохранения биологического разнообразия видов возделываемых растений, изучение их происхождения, определение их филогении, предполагает детальную разработку внутривидовых классификаций (Дорофеев и др., 1979; Goncharov, 2011). Создание таких классификаций, отражающих филогенез и генетическую структуру видов, следует считать основной целью современной таксономии (Hammer et al., 2011). При разработке классификации культурных растений предполагается максимальное полное описание всех существующих крупных и мелких форм (таксонов) (Синская, 1969). Это определяется удобством применения такого деления, с одной стороны, в экспериментальной работе, с другой, при селекции и апробации сельскохозяйственных растений.

Успех и эффективность исследовательской работы часто связаны с детальностью и полнотой экспериментальной проработки, которая зависит от того, каков материал и насколько подробно его следует изучать. В связи с этим исключительно важно, чтобы естественная дифференциация того или иного рода, взаимосвязи между видами с высокой точностью были отражены детально разработанной таксономией (Дорофеев, 1985). Следует заметить, что у большей части растений, важных для сельского хозяйства, до настоящего времени однозначно не определены объемы рода и видов (Родионов и др., 2019).

Серьезная проблема систематики культурных растений – аспект укрупнения–дробления таксонов, причем в случаях возделывания она проявляется особенно контрастно (Головнина и др., 2009; Goncharov, 2011). В то же время эффективное использование классификаций (систем родов) возделываемых растений в работе исследователей вызывает определенные сложности. И дихотомические таблицы (Дорофеев и др., 1979; Гончаров, 2009), и идеографические определители (Зуев и др., 2019) требуют определенных навыков, поэтому создание баз данных и

программного обеспечения, позволяющее по цифровым изображениям определять видовую принадлежность, – очень перспективное направление. Разработка данных методов основывается преимущественно на технологиях анализа цифровых изображений органов растений в рамках компьютерной феномики (Афонников и др., 2016; Zhang et al., 2019; Демидчик и др., 2020; Yang et al., 2020).

Пшеница – одна из важнейших мировых продовольственных культур. Современные возделываемые виды пшениц произошли от трех диких диплоидных предков в результате их гибридизации и нескольких раундов удвоения генома (полиплоидизации). В настоящее время культивируемые пшеницы представлены ди- ( $2n = 2x = 14$ , геном  $A^bA^b$ ), тетра- ( $2n = 4x = 28$ , геном  $BBA^uA^uDD$ ) и гексаплоидными ( $2n = 6x = 42$ , геном  $BBA^uA^uDD$ ) видами (Гончаров, Кондратенко, 2008). Основной возделываемый вид, мягкая пшеница (*Triticum aestivum* L.), является гексаплоидом (геномная формула  $BBA^uA^uDD$ ). Уровень плоидности служит одним из основных таксономических и классифицирующих признаков видов пшениц (Международный классификатор..., 1984; van Slageren, Payne, 2013). Его можно устанавливать цитогенетическими (Родионов и др., 2020), молекулярными методами, а также на основе сравнения морфологических характеристик растений (Международный классификатор..., 1984). В нашей работе проведено изучение морфологических характеристик колосьев растений тетра- и гексаплоидных видов пшеницы с применением метода высокопроизводительного фенотипирования.

Целью исследования было изучение распределения морфологических характеристик колосьев у тетра- и гексаплоидных видов пшениц и сравнение их распределений.

## Материалы и методы

**Биологический материал.** В работе изучено 17 видов полиплоидных пшениц: девять гексаплоидных (*Triticum compactum* Host, *T. aestivum* L., *T. antiquorum* Heer ex

**Таблица 1.** Описание используемых видов пшениц

| Вид  | Всего      |          |          |          |             | Плоидность             | Список вегетаций      |      |
|--|------------|----------|----------|----------|-------------|------------------------|-----------------------|------|
|  | фотографий | растений | образцов | на столе | на прищепке |                        |                       |      |
| <i>T. compactum</i> Host                                       | 472        | 101      | 10       | 177      | 295         | Hexaploid              | II18, IX16            |      |
| <i>T. aestivum</i> L.  | 456        | 80       | 8        | 166      | 290         |                        | II19, IX16, IX18, X14 |      |
| <i>T. antiquorum</i> Heer ex Udacz.                            | 184        | 37       | 4        | 116      | 68          |                        | II18, X14             |      |
| <i>T. spelta</i> L.  | 164        | 49       | 5        | 40       | 124         |                        | II18                  |      |
| <i>T. petropavlovskiyi</i> Udacz. et Migusch.                  | 374        | 75       | 6        | 74       | 300         |                        | II17, IX17, IX18      |      |
| i:АНК-23   | 50         | 10       | 1        | 14       | 36          |                        | IX16                  |      |
| <i>T. yunnanense</i> King ex S.L. Chen                         | 191        | 43       | 3        | 43       | 148         |                        | IX17, IX18            |      |
| <i>T. spelta</i> cv. Rother Sommer Kolben                      | 45         | 9        | 1        | 9        | 36          |                        | IX16, II18            |      |
| <i>T. macha</i> Dekapr. et Menabde                             | 46         | 10       | 1        | 10       | 36          |                        | IX17, IX18            |      |
| <i>T. sphaerococcum</i> Perciv.                                | 100        | 20       | 2        | 20       | 80          |                        | IX17                  |      |
| <i>T. vavilovii</i> (Thum.) Jakubz.                            | 15         | 3        | 1        | 3        | 12          |                        | II18                  |      |
| <i>T. aethiopicum</i> Jakubz.                                  | 595        | 119      | 12       | 119      | 476         |                        | Tetraploid            | X14  |
| <i>T. dicoccoides</i> (Körn. ex Aschers. et Graebn.) Schweinf. | 40         | 8        | 1        | 8        | 32          |                        |                       | II16 |
| <i>T. dicoccum</i> (Schränk) Schuebl.                          | 41         | 9        | 1        | 9        | 32          | II17                   |                       |      |
| <i>T. durum</i> Desf.  | 275        | 56       | 5        | 55       | 220         | II16, II17, II19, IX18 |                       |      |
| <i>T. militinae</i> Zhuk. et Migusch.                          | 40         | 8        | 1        | 8        | 32          | IX17                   |                       |      |
| <i>T. polonicum</i> L.   | 95         | 19       | 2        | 19       | 76          | II16, II19             |                       |      |
| <i>T. timopheevii</i> (Zhuk.) Zhuk.                            | 125        | 25       | 3        | 25       | 100         | II16, IX18             |                       |      |
| <i>T. turgidum</i> L.  | 40         | 8        | 1        | 8        | 32          | II15                   |                       |      |

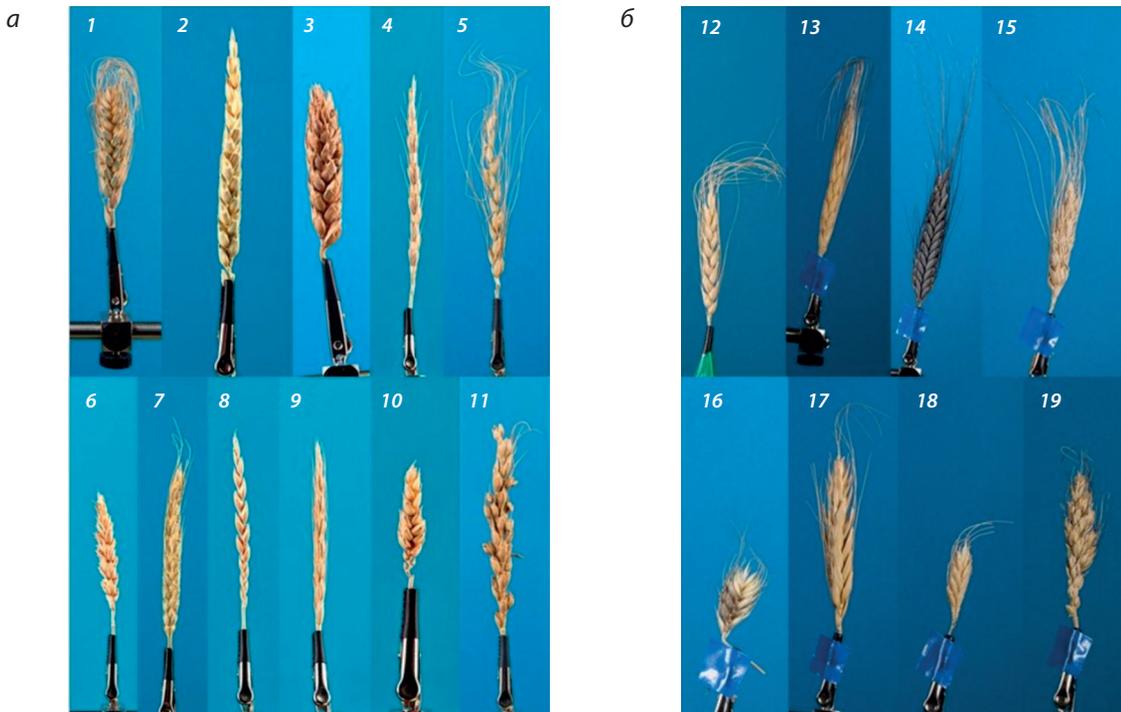
Udacz., *T. spelta* L. (включая стародавний сорт *T. spelta* Rother Sommer Kolben), *T. petropavlovskiyi* Udacz. et Migusch., *T. yunnanense* King ex S.L. Chen, *T. macha* Dekapr. et Menabde, *T. sphaerococcum* Perciv., *T. vavilovii* (Thum.) Jakubz.), изогенная линия сорта мягкой пшеницы Новосибирская 67 АНК-23 и восемь тетраплоидных (*T. aethiopicum* Jakubz., *T. dicoccoides* (Körn. ex Aschers. et Graebn.) Schweinf., *T. dicoccum* (Schränk) Schuebl., *T. durum* Desf., *T. militinae* Zhuk. et Migusch., *T. polonicum* L., *T. timopheevii* (Zhuk.) Zhuk., *T. turgidum* L.); выборка представлена колосьями 595 уникальных и выращенных в девяти вегетациях растений. Растения были выращены в 2014–2019 гг. в гидропонной теплице центра коллективного пользования «Лаборатория искусственного выращивания растений» ИЦиГ СО РАН (г. Новосибирск). Материал, использованный в работе, описан в табл. 1.

Нужно отметить, что ни в одном из крупных генетических банков мира не существует типовых коллекций по пшенице, поэтому выборки, как правило, либо отражают взгляд исследователей на проблему создания таких коллекций (Пальмова, 1935), либо определяются репрезентативностью доступного для исследования материала (Гончаров, Шумный, 2008). Стандартное таксономическое описание образцов есть в публичных базах данных на сайтах генбанков (<http://db.vir.nw.ru/virdb/maindb>).

**Получение цифровых изображений.** В работе использовали два протокола получения фотографий зрелых колосьев. Первый – колос располагается на стекле просветного столика, который находится на столе с поверхностью синего цвета (фон). Фотокамера фиксируется на

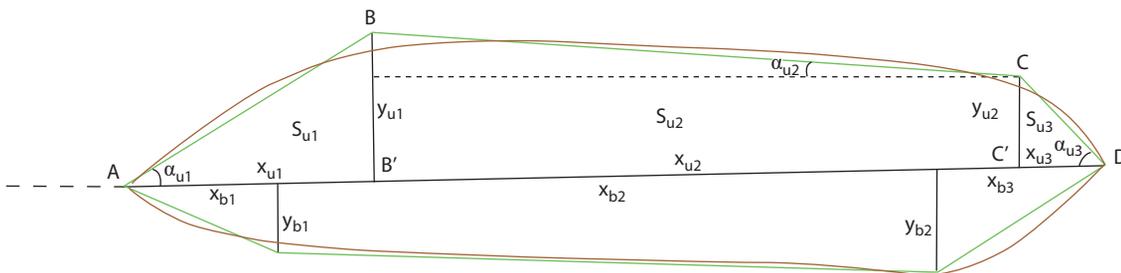
стойке над стеклом. С помощью данного метода можно производить съемку лицевой проекции колоса. Второй – колос располагается вертикально перед синим фоном. Опорой колоса служат прищепки, которые помещаются на штатив. Применяя этот метод, при вращении колоса относительно его оси, можно производить съемку колоса в четырех или более проекциях (Генаев и др., 2018). Согласно протоколам, на фотографиях должна присутствовать цветовая палитра (ColorChecker). Она нужна для нормализации цветов и определения масштаба. Одному растению в нашем наборе данных может соответствовать до пяти фотографий его колоса, снятых по разным протоколам и в разных проекциях. Примеры изображений колосьев (по одному для каждого вида) представлены на рис. 1. Всего с использованием двух протоколов было получено 3348 изображений колосьев в разных проекциях, 2097 из них относились к гексаплоидным видам, а 1251 – к тетраплоидным, из них 915 изображений получены по протоколу «на столе» и 2433 – «на прищепке».

**Оценка количественных характеристик колосьев.** Для всесторонней оценки таких характеристик на основе анализа изображений использовали программу WERecognizer (Genaev et al., 2019). Эта программа описывает колос пшеницы в виде модели двух четырехугольников (рис. 2). Геометрия данной модели описывается девятью независимыми параметрами. Для верхнего четырехугольника – это параметры  $x_{u1}$ ,  $x_{u2}$ ,  $y_{u1}$ ,  $y_{u2}$ ; для нижнего четырехугольника –  $x_{b1}$ ,  $x_{b2}$ ,  $y_{b1}$ ,  $y_{b2}$ ; общий параметр для двух четырехугольников – длина колоса. Дополнительно программа рассчитывает ряд общих ха-



**Рис. 1.** Изображения колосьев различных видов гексаплоидных (а) и тетраплоидных (б) пшениц.

1 – *T. compactum*; 2 – *T. aestivum*; 3 – *T. antiquorum*; 4 – *T. spelta*; 5 – *T. petropavlovskyi*; 6 – i:АНК-23; 7 – *T. yunnanense*; 8 – *T. spelta* cv. Rother Sommer Kolben; 9 – *T. macha*; 10 – *T. sphaerococcum*; 11 – *T. vavilovii*; 12 – *T. aethiopicum*; 13 – *T. dicoccoides*; 14 – *T. dicoccum*; 15 – *T. durum*; 16 – *T. militinae*; 17 – *T. polonicum*; 18 – *T. timopheevii*; 19 – *T. turgidum*.



**Рис. 2.** Представление формы колоса в виде двух четырехугольников (Genaev et al., 2019).

Черной горизонтальной линией обозначена осевая линия колоса. Его контур показан коричневой линией. Четырехугольники, аппроксимирующие контур колоса, отмечены зелеными линиями. Штриховая линия слева – его основание. Для верхнего четырехугольника указаны основные параметры, характеризующие его геометрию. Аналогичные параметры определяются и для нижнего четырехугольника.

рактических характеристик формы и размера колоса, а также параметры его остистости. Детали алгоритма извлечения признаков приведены в статье (Genaev et al., 2019).

Мы использовали признаки модели, которые были отобраны как наиболее информативные для предсказания индекса плотности колоса в нашем предыдущем исследовании (Genaev et al., 2019), а также общие признаки формы и остей. Эти признаки характеризуют комплексное представление о морфологии (фенотипе) колоса, описывая его форму (Circularity, Roundness), физические размеры тела колоса (Perimeter, Rachis length) и площадь остей (Awns area). Признаки, полученные в результате аппроксимации колоса двумя четырехугольниками, связаны с шириной ( $x_{u2}$ ,  $y_{bm}$ ) и длиной ( $x_{b2}$ ,  $y_{u2}$ ) отдельных сегментов колоса (табл. 2).

**Анализ данных.** Для того чтобы оценить распределение колосьев в пространстве изучаемых признаков, мы воспользовались нелинейным алгоритмом снижения размерности t-SNE (t-distributed stochastic neighbor embedding; Maaten, Hinton, 2008). Этот метод позволяет визуализировать многомерные данные путем отображения объектов в многомерном пространстве в пространство меньшей размерности (двух- или трехмерное). Основная идея t-SNE заключается в уменьшении размерности пространства при сохранении относительных попарных расстояний между объектами. Преимуществом метода t-SNE является склонность к локализации изолированных плотных пространственных структур произвольной геометрии. Метод t-SNE был применен для ординации изображений колосьев; при этом изображения каждой из

**Таблица 2.** Описание признаков колоса

| Название признака | Описание   | Размерность     |
|-------------------|--|-----------------|
| Awns area         | Площадь остей  | мм <sup>2</sup> |
| Circularity       | Индекс округлости равен отношению периметра окружности с площадью, равной площади контура, к периметру контура. Индекс отражает, насколько форма контура близка к форме окружности. Значение варьирует от 0 до 1 | Безразмерная    |
| Roundness         | Индекс закругленности равен отношению площади контура к площади окружности с диаметром, равным осевой линии колоса   |                 |
| Perimeter         | Периметр контура колоса без остей  | мм              |
| Rachis length     | Длина ломаной линии вдоль оси сложного колоса (линии оси колоса)   |                 |
| X <sub>u2</sub>   | Параметр модели четырехугольников, связанный с длиной левой центральной части колоса (вверху на рис. 2)  |                 |
| X <sub>b2</sub>   | Параметр модели четырехугольников, связанный с длиной правой центральной части колоса (внизу на рис. 2)  |                 |
| Y <sub>u2</sub>   | Расстояние от вершины С до ее проекции С' на основание AD (см. рис. 2)   |                 |
| Y <sub>bm</sub>   | Параметр модели четырехугольников. Среднее значение высоты правого (нижнего) четырехугольника  |                 |

проекций одного колоса рассматривались как отдельные объекты.

Для оценки сходства количественных характеристик колосьев для разных видов мы применили иерархическую кластеризацию (Johnson, 1967) 17 видов пшеницы по признакам, полученным в результате усреднения по всем колосьям одного и того же вида. При этом каждый вид был охарактеризован вектором признаков длины 9. В качестве метрики расстояния между видами было значение  $1 - r$ , где  $r$  – величина коэффициента корреляции Пирсона между значениями признаков (Müllner, 2011). Для кластеризации и построения дендрограммы использованы соответственно функции linkage (алгоритм UPGMA) и dendrogram из библиотеки SciPy (Virtanen et al., 2020).

Для сравнения дисперсий признаков у растений, относящихся к разным типам плоидности, мы применили  $F$ -статистику (Snedecor, Cochran, 1989), которая оценивает значимость различий дисперсий двух распределений. Данные нормированы функцией StandardScaler библиотеки scikit-learn (Pedregosa et al., 2011). Тест проводили независимо для каждого из девяти признаков, описанных в табл. 2. При выполнении этого теста для каждого растения взято одно изображение колоса, полученное в проекции по протоколу «на столе».

## Результаты и обсуждение

Средние значения, медиана, среднее квадратическое отклонение и дисперсия девяти признаков, рассчитанных для 17 видов пшениц, представлены в Приложении 1<sup>1</sup>.

Рассмотрим распределение колосьев нашей выборки растений по признаку «площадь остей». Чем выше этот параметр, тем больше остей было идентифицировано для колоса на изображении. По этой характеристике колосья гексаплоидных пшениц условно можно разделить на три класса: остистые (значение параметра выше 90), умеренно остистые (значение параметра от 30 до 90) и безостые

(значение параметра менее 30). К остистым по такому критерию относятся представители видов *T. compactum*, *T. spelta*, *T. petropavlovskiyi*, *T. vavilovii*; к умеренно остистым – *T. aestivum*, *T. yunnanense*, *T. macha*; к безостым – *T. antiquorum*, i:АНК-23, *T. spelta* cv. Rother Sommer Kolben, *T. sphaerococcum* (см. Приложение 1). Эти данные хорошо согласуются с внешним видом колосьев (см. рис. 1, а). Таким образом, образцы гексаплоидных пшениц демонстрируют значительное разнообразие по наличию/отсутствию остей.

Если применить указанную классификацию для тетраплоидных пшениц, то к категории безостых относится лишь вид *T. militinae* (среднее значение параметра 24.09 мм<sup>2</sup>). Четыре вида можно причислить по этому признаку к умеренно остистым: *T. dicoccoides*, *T. polonicum*, *T. timopheevii*, *T. turgidum*; три вида – к остистым: *T. aethiopicum*, *T. dicoccum*, *T. durum*. В целом представленность остистых видов (образцов) у тетраплоидных пшениц значительно выше, чем у гексаплоидных.

Анализ такой характеристики, как длина колоса, показывает, что колосья можно также разделить на три класса: длина менее 60 мм (короткие), от 60 до 90 мм (средние) и более 90 мм (длинные). По данной классификации виды гексаплоидных пшениц *T. spelta*, *T. petropavlovskiyi* и *T. vavilovii* можно отнести к длинноколосым; *T. aestivum*, *T. yunnanense*, *T. spelta* cv. Rother Sommer Kolben и *T. macha* – к среднеколосым, а *T. compactum*, *T. antiquorum*, *T. sphaerococcum* и изогенную линию АНК-23 – к короткоколосым. При этом граница между видами, характеризующимися длинными и средними колосьями, оказывается достаточно условной. У тетраплоидных видов мы не обнаружили ни одного вида, который попал бы по этому параметру в категорию длинноколосых. К категории среднеколосых можно отнести *T. aethiopicum*, *T. dicoccoides*, *T. polonicum* и *T. turgidum*, к категории короткоколосых – *T. dicoccum*, *T. durum*, *T. timopheevii* и *T. militinae*.

На рис. 3, а приведено распределение изученных образцов по длине колосьев для гекса- и тетраплоидных видов; на рис. 3, б показано распределение параметра, также

<sup>1</sup> Приложения 1–4 см. по адресу:  
<http://www.bionet.nsc.ru/vogis/download/pict-2021-25/appx3.pdf>

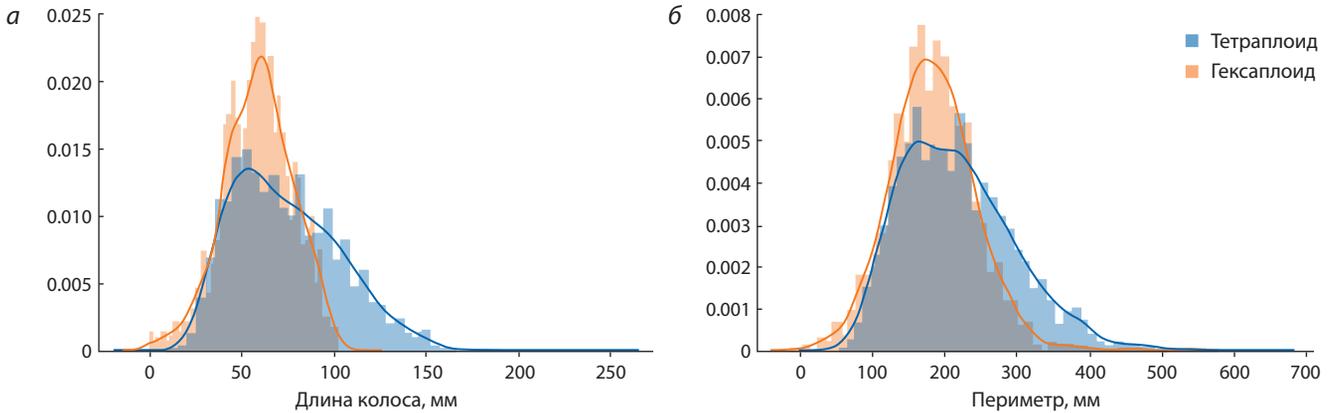


Рис. 3. Распределение длины (а) и периметра (б) колоса у тетраплоидных (синий цвет) и гексаплоидных (оранжевый цвет) видов пшениц.

характеризующего размер колоса, – периметра контура тела колоса на изображении.

Согласно рис. 3, распределения обоих параметров у гексаплоидных пшениц имеют более высокий разброс, при этом изменчивость этих признаков у гексаплоидных пшениц выше в основном за счет большей частоты встречаемости колосьев с большими значениями этих признаков.

Распределение проанализированных изображений колосьев в пространстве девяти характеристик мы визуализировали при помощи метода t-SNE, получив на его основе двумерное пространство параметров (компоненты 1 и 2). Результаты преобразования приведены на рис. 4. На диаграмме каждая точка представляет одно из проанализированных изображений колоса. На рис. 4, а точки окрашены в соответствии с типом плоидности растения (синий цвет – тетраплоидные виды пшеницы, оранжевый – гексаплоидные); на рис. 4, б цвет и форма каждой точки соответствуют определенному виду пшениц.

На диаграмме рис. 4, а показано, что области, которые занимают образцы гекса- и тетраплоидных видов пшениц, сильно перекрываются. Следовательно, по своим характеристикам колосья этих двух групп достаточно близки. Это согласуется с результатами, представленными в Приложениях 1 и 2, а также на рис. 3. Однако нужно отметить, что на диаграмме рис. 4, а образцы гексаплоидных видов занимают большую площадь, прежде всего, за счет преобладания соответствующих точек в правой части диаграммы. Видно, что в области при значениях компоненты 1 более  $-20$  преобладают оранжевые точки (гексаплоидные пшеницы), такое превалирование еще более заметно в верхнем правом углу диаграммы (значения компоненты 1 менее 0, а компоненты 2 – больше 20). Это означает, что ряд характеристик колоса наблюдается лишь у гексаплоидных видов, но не у тетраплоидных, что хорошо согласуется с результатом, приведенным на рис. 3. В частности, такие области могут соответствовать большим значениям параметров «периметр» и «длина колоса».

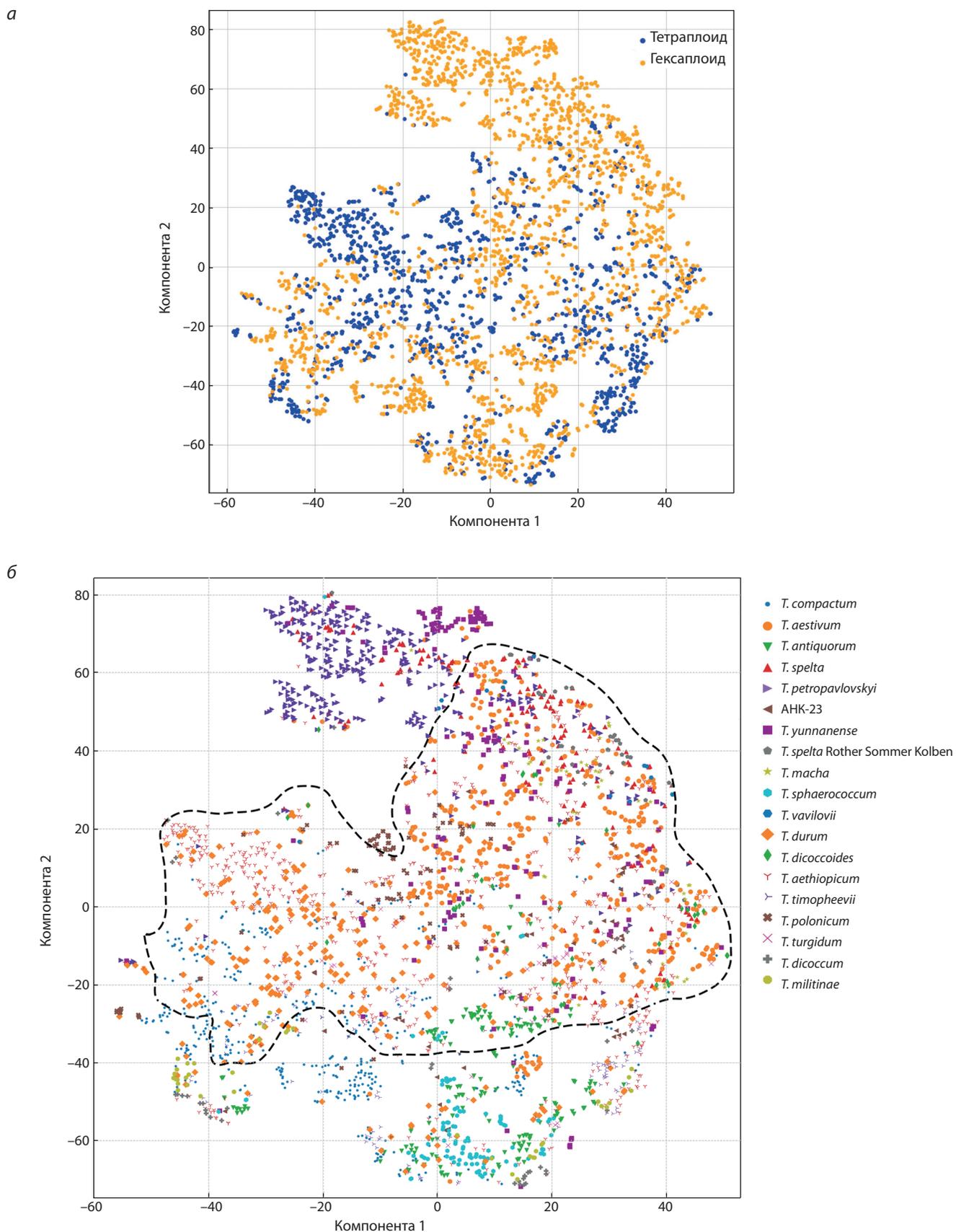
На диаграмме рис. 4, б продемонстрировано, что области, занятые образцами разных видов, в большой степени перекрываются. Например, видам *T. aestivum* и *T. durum* соответствует значительная площадь на графике (см. рис. 4, б, штриховая линия). Однако следует отметить, что

изображения колосьев, принадлежащих к одному виду пшениц, занимают преимущественно компактные области на графике. В то же время существуют виды, для которых образцы колосьев распадаются по своим характеристикам на несколько хорошо заметных кластеров. К таким видам можно отнести *T. compactum* (метки: малый синий круг, компонента 1 от  $-60$  до 0, компонента 2 от  $-60$  до 0) и *T. petropavlovskiyi* (фиолетовый треугольник, компонента 1 от  $-20$  до 0, компонента 2 от 40 до 80).

На рис. 1 гексаплоиды представлены растениями с двумя характерными типами колосьев: длинными и тонкими (*T. aestivum*, *T. spelta*, *T. petropavlovskiyi*, *T. yunnanense*, *T. spelta* cv. Rother Sommer Kolben); короткими и округлыми (*T. compactum*, *T. antiquorum*, i:АНК-23, *T. sphaerococcum*, *T. macha*, *T. vavilovii*). На рис. 4, б группа растений с короткими и округлыми колосьями располагается в диапазоне значений компоненты 2 от  $-80$  до 0 (нижняя часть графика). Растения с длинными и тонкими колосьями имеют значения компоненты 2 в пределах 0 до 80 (верхняя часть графика). На рис. 4, а две этих группы растений грубо соответствуют двум облакам точек у гексаплоидных пшениц в верхней и нижней частях графика, которые слабо перекрываются в его центральной части. Таким образом, продемонстрированные на рис. 4 диаграммы позволяют наглядно показать разнообразие колосьев по своим характеристикам как внутри, так и между видами.

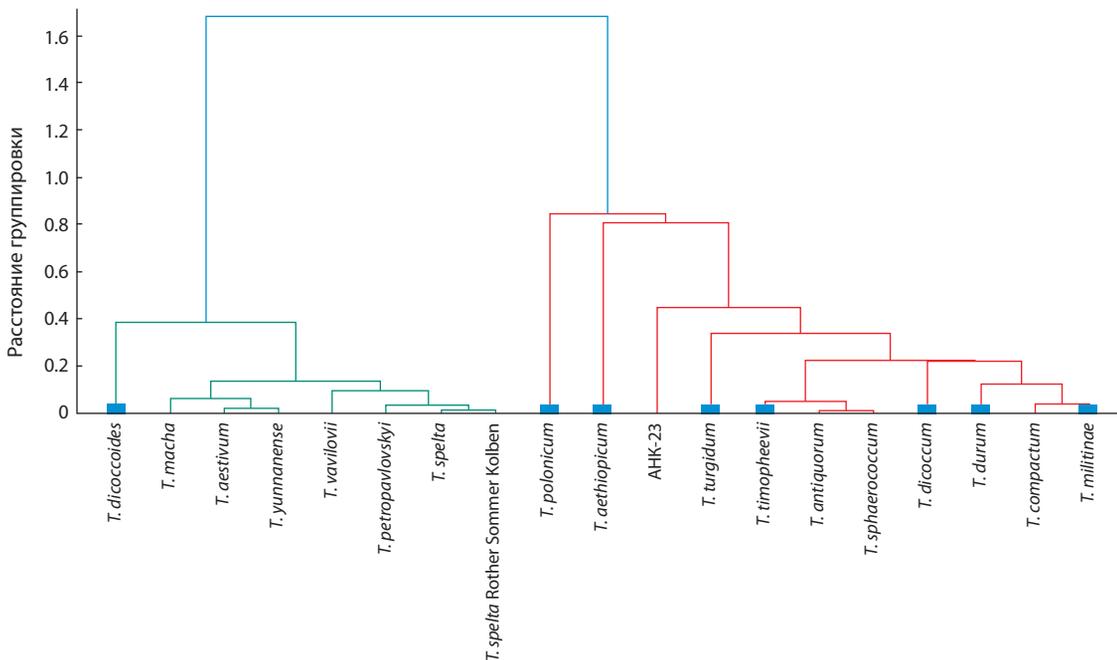
Чтобы охарактеризовать более детально сходство морфометрических параметров колосьев у разных видов пшениц, мы провели иерархический кластерный анализ для них на основе сравнения средних значений изучаемых признаков (рис. 5).

На рис. 5 виды пшениц разделены по характеристикам колосьев на два кластера (они обозначены красным и зеленым цветом). К первому кластеру (красный цвет) преимущественно относятся тетраплоидные виды (показаны синими прямоугольниками у конечных вершин дерева). Однако в него не включен дикий вид тетраплоидных пшениц *T. diccocooides*, а из гексаплоидных видов в него входят *T. compactum*, *T. antiquorum*, *T. sphaerococcum*, отличающиеся от всех остальных видов компактной формой колоса, т. е. из всех изученных видов гексаплоидных пшениц у них самый короткий колос. Отметим, что в работе (Zatybekov et al., 2020) при использовании хозяйственно



**Рис. 4.** Ординация цифровых изображений колосьев отдельных генотипов методом t-SNE, полученная на основании количественных признаков из табл. 2.

*а* – тетраплоидные виды пшеницы – синий цвет, гексаплоидные – оранжевый; *б* – цвет и форма каждой точки соответствуют определенному виду. Штриховой линией отмечена область, занимаемая видами *T. aestivum* и *T. durum*.



**Рис. 5.** Результаты иерархического кластерного анализа для девяти признаков колоса пшеницы. Синие квадраты соответствуют тетраплоидам.

важных признаков образцы шести тетраплоидных видов кластеризовались произвольно, т.е. вне зависимости от их видовой принадлежности. Интересно, что остальные гексаплоидные виды четко разделились на два кластера – среднеколосые (*T. macha*, *T. aestivum*, *T. yunnanense*) и длинноколосые (*T. vavilovii*, *T. petropavlovskiyi*, *T. spelta*).

Включение в один кластер образцов *T. spelta* и *T. spelta* cv. Rother Sommer Kolben (стародавнего немецкого сорта) позволило сделать вывод, что «видовая» форма колоса в процессе многолетней селекции этого вида пшениц не менялась продолжительное время (в данном случае более 50 лет) и может быть успешно использована при классификации видов.

Следует отметить, что к гексаплоидам попал единственный в роде дикий тетраплоидный рыхлоколосый вид *T. dicoccoides*. В то время как в тетраплоидные виды попали гексаплоидные пшеницы с компактным типом колоса, *T. compactum*, *T. antiquorum*, *T. sphaerococcum*, и рукотворная изогенная линия АНК-23, созданная на сорте яровой мягкой пшеницы Новосибирская 67 (Коваль, 1997). Последнее допускает сделать заключение, что, несмотря на то, что изогенные линии создаются на определенном (конкретном) виде, тем не менее к их видовой принадлежности следует относиться с осторожностью.

Чуть подробнее остановимся на *T. petropavlovskiyi*. Вид приручен к китайскому Припамирию – маршруту Великого шелкового пути. По результатам изучения запасных белков, все образцы этого вида были очень похожи на потомков гибридной комбинации, полученной от скрещивания мягкой пшеницы с *T. polonicum* (Watanabe et al., 2004). Авторы монографии «Культурная флора СССР» также считали возможным гибридогенное происхождение вида (Дорофеев и др., 1979). Кроме того, по ряду таксономических признаков *T. petropavlovskiyi* также похожа

на мягкую пшеницу (Goncharov, 2005). К *T. petropavlovskiyi* ssp. *mexicana* Bogusl. P.Л. Богуславским (1982) были отнесены межвидовые гибриды, полученные селекционерами СИММУТ. Исходя из вышеизложенного, мы посчитали целесообразным перевод *T. petropavlovskiyi* в подвид *T. aestivum*:

***Triticum aestivum* ssp. *petropavlovskiyi* comb. et stat. nov. (Udacz. et Migusch.) N.P. Gontsch.** – *T. turanicum* Jakubz. convar. *montanostepposum* Jakubz. f. *aristiforme* Jakubz. *Бот. журн.* 1959;10:1428, nom. illig. – *T. petropavlovskiyi* Udacz. et Migusch. *Вестник с.-х. науки.* 1970;9:20. – **II. Петропавловского.**

**Тип:** описан по образцу из Китая «Происхождение: Китай, провинция Синцзян, сел. Курля, К-48376, 1957. Эксп. А.М. Горского. Репродукция Средняя Азия, Ташкент, САС ВИР. 8 VII 1969. Собрал/определил: Р.А. Удачин и Э.Ф. Мигушова» в С.-Петербурге (WIR!) (Гербарные экземпляры типа и паратипа приведены в Приложениях 3 и 4).

***Triticum aestivum* ssp. *petropavlovskiyi* comb. et stat. nov. (Udacz. et Migusch.) N.P. Gontsch.** – *T. turanicum* Jakubz. convar. *montanostepposum* Jakubz. f. *aristiforme* Jakubz. 1959. *Bot. Zhur.* 10:1428, nom. illig. – *T. petropavlovskiyi* Udacz. et Migusch. 1970. *Vestn. Sel'skokhoz. Nauki.* 9:20.

**Типус:** described by an accession from China “Origin: China, Xinjiang Province, village Kurlia, K-48376, 1957. A.M. Gorsky exp[edition]. Reproduction of Central Asia, Tashkent, Central Asian Station of VIR. 08. VII. 1969, Collected/defined: R.A. Udachin & E.F. Migushova” in St. Petersburg (WIR!).

Отметим, что результаты, представленные на рис. 3 и 4, а, свидетельствуют, что у гексаплоидных видов характеристики формы, размеров и остистости колосьев оказываются более разнообразными. Поэтому мы пред-

**Таблица 3.** Результаты применения *F*-статистики для подтверждения гипотезы о значимом различии дисперсии двух распределений

| Название признака | <i>F</i> -статистика | <i>p</i> -value | Дисперсия    |              | Среднее      |              |
|-------------------|----------------------|-----------------|--------------|--------------|--------------|--------------|
|                   |                      |                 | гексаплоидов | тетраплоидов | гексаплоидов | тетраплоидов |
| Awns area         | 0.376                | 1.000           | 1.415        | 3.763        | 84.875       | 160.643      |
| Circularity index | 1.188                | 0.065           | 0.959        | 0.807        | 0.178        | 0.181        |
| Roundness         | 1.828                | 1.110e-07       | <b>1.312</b> | 0.718        | 0.141        | 0.172        |
| Perimeter         | 1.570                | 4.710e-05       | <b>1.080</b> | 0.688        | 218.124      | 185.015      |
| Rachis length     | 3.500                | < 1e-15         | <b>1.320</b> | 0.377        | 74.136       | 59.280       |
| $X_{u2}$          | 3.928                | < 1e-15         | <b>1.336</b> | 0.340        | 53.837       | 36.853       |
| $X_{b2}$          | 4.437                | < 1e-15         | <b>1.331</b> | 0.300        | 54.004       | 36.726       |
| $Y_{u2}$          | 4.275                | < 1e-15         | <b>2.491</b> | 0.583        | 3.844        | 4.171        |
| $Y_{bm}$          | 1.081                | 0.248           | 0.695        | 0.643        | 0.225        | 0.246        |

Примечание. Значимые различия дисперсии выделены полужирным шрифтом.

положили, что характеристики колосьев у гексаплоидных видов могут иметь более высокий разброс значений, по сравнению с тетраплоидными видами. Для проверки этого предположения мы провели сравнение дисперсий по оцененным параметрам при помощи *F*-распределения (табл. 3).

Из результатов, приведенных в табл. 3, следует, что дисперсия большинства характеристик для гекса- и тетраплоидов имеет значимые различия ( $p < 0.05$ ). В то же время значимые различия дисперсий не были нами обнаружены для таких признаков, как  $y_{bm}$  (параметр модели четырехугольников), Awns area и Circularity index. Интересно, что для всех значимых различий мы наблюдаем значение дисперсии у гексаплоидов выше, чем у тетраплоидов. Таким образом, проведенный анализ показал, что гексаплоидные виды демонстрируют более высокое разнообразие по морфометрическим характеристикам колоса, по сравнению с тетраплоидными.

В данных представлены растения 17 видов: 9 гексаплоидных: *T. compactum*, *T. aestivum* (в том числе изогенная линия сорта Новосибирская 67 АНК-23), *T. antiquorum*, *T. spelta* (включая стародавний сорт *T. spelta* Rother Sommer Kolben), *T. petropavlovskiyi*, *T. yunnanense*, *T. macha*, *T. sphaerococcum*, *T. vavilovii* и 8 тетраплоидных: *T. aethiopicum*, *T. dicoccoides*, *T. dicoccum*, *T. durum*, *T. militinae*, *T. polonicum*, *T. timopheevii*, *T. turgidum*. Результаты их кластеризации даны в таком варианте, чтобы цвет и форма каждой точки соответствовали определенному виду (см. рис. 5).

Хорошо известно, что удвоение генома в результате дупликаций (автополиплоидия) или гибридизации и последующей полиплоидизации (аллополиплоидия) приводит к заметным изменениям фенотипа растений (Finigan et al., 2012; Романов, Пимонов, 2018; Родионов и др., 2019). Эти изменения у растений происходят как на клеточном уровне (Liu et al., 2018), так и на уровне органов (Robinson et al., 2018). Во многих случаях у растений увеличение плоидности ведет к увеличению размеров клеток и органов (Comai, 2005; Williams Oliveira, 2020), повышению устойчивости к стрессам (Tan et al., 2015). В настоящее

время исследователи предполагают, что существует четыре типа молекулярных механизмов такой изменчивости: 1) увеличение дозы гена/аллеля, 2) увеличение генетического разнообразия, 3) изменение генетической регуляции и 4) эпигенетические перестройки генома (Chen, 2007; Finigan et al., 2012).

В нашей работе на примере анализа морфологических характеристик колосьев пшеницы гекса- ( $2n = 6x = 42$ ) и тетраплоидных ( $2n = 4x = 28$ ) видов мы показали, что для большинства признаков колоса их вариации у пшениц с большей плоидностью значимо выше. Полученные нами результаты находятся в согласии с представлениями о влиянии плоидности на изменчивость фенотипа растений.

### Закключение

В работе проведен массовый анализ цифровых изображений колосьев для 595 растений восьми тетра- и девяти гексаплоидных видов пшениц. Рассмотрено девять количественных признаков, описывающих форму, размер и остистость колоса. Изучена изменчивость генотипов по указанным выше признакам и показано, что в пространстве характеристик колоса формируются два кластера. К первому относятся преимущественно гексаплоидные виды (за исключением дикого тетраплоидного вида *T. dicoccoides*). Ко второму – тетраплоидные (за исключением трех гексаплоидных с компактной формой колоса видов, *T. antiquorum*, *T. sphaerococcum* и изогенной линии АНК-23). Анализ дисперсий этих признаков у гекса- и тетраплоидных растений показал значимое увеличение дисперсии для шести из девяти признаков в выборке гексаплоидных растений, т. е. большая плоидность дает более варибельные значения количественных признаков морфологии колоса.

Таким образом, морфологические характеристики колосьев для гекса- и тетраплоидных видов, полученные на основе компьютерного анализа изображений, демонстрируют различия, которые в дальнейшем могут быть использованы для разработки методики классификации растений по плоидности и их видовой принадлежности в автоматическом режиме.

## Список литературы / References

- Афонников Д.А., Генаев М.А., Дорошков А.В., Комышев Е.Г., Пшеничникова Т.А. Методы высокопроизводительного фенотипирования растений для массовых селекционно-генетических экспериментов. *Генетика*. 2016;52(7):788-803. DOI 10.7868/S001667581607002X.  
[Afionnikov D.A., Genaev M.A., Doroshkov A.V., Komyshev E.G., Pshenichnikova T.A. Methods of high-throughput plant phenotyping for large-scale breeding and genetic experiments. *Russ. J. Genet.* 2016;52(7):688-701. DOI 10.1134/S1022795416070024.]
- Богуславский Р.Л. Новая ботаническая форма гексаплоидной пшеницы. *Науч.-техн. бюл. ВИР*. 1982;119:73-74.  
[Boguslavsky R.L. A new botanical form of hexaploid wheat. *Nauchno-Tekhnicheskiy Byulleten VIR = Scientific and Technological Bulletin of the Vavilov Institute of Plant Industry*. 1982;119:73-74. (in Russian)]
- Генаев М.А., Комышев Е.Г., Фу Хао, Коваль В.С., Гончаров Н.П., Афонников Д.А. SpikeDroidDB – информационная система для аннотации морфометрических характеристик колоса пшеницы. *Вавиловский журнал генетики и селекции*. 2018;22(1):132-140. DOI 10.18699/VJ18.340.  
[Genaev M.A., Komyshev E.G., Fu Hao, Koval V.S., Goncharov N.P., Afionnikov D.A. SpikeDroidDB: an information system for annotation of morphometric characteristics of wheat spike. *Vavilovskii Zhurnal Genetiki i Selekcii = Vavilov Journal of Genetics and Breeding*. 2018;22(1):132-140. DOI 10.18699/VJ18.340. (in Russian)]
- Головнина К.А., Кондратенко Е.Я., Блинов А.Г., Гончаров Н.П. Филология А геномов диких и возделываемых видов пшениц. *Генетика*. 2009;45(11):1540-1547.  
[Golovnina K.A., Kondratenko E.Ya., Blinov A.G., Goncharov N.P. Phylogeny of the A genome of wild and cultivated wheat species. *Russ. J. Genet.* 2009;45(11):1360-1367. DOI 10.1134/S1022795409110106.]
- Гончаров Н.П. Определитель разновидностей мягкой и твердой пшениц. Новосибирск: Изд-во СО РАН, 2009.  
[Goncharov N.P. Manual Book of Common and Hard Wheat Varieties. Novosibirsk: SO RAN Publ., 2009. (in Russian)]
- Гончаров Н.П., Кондратенко Е.Я. Происхождение, доместикация и эволюция пшениц. *Информ. вестник ВОГИС*. 2008;12(1/2):159-179.  
[Goncharov N.P., Kondratenko E.Ya. Wheat origin, domestication and evolution. *Informatsionnyy vestnik VOGiS = The Herald of Vavilov Society for Geneticists and Breeders*. 2008;12(1/2):159-179. (in Russian)]
- Гончаров Н.П., Шумный В.К. От сохранения генетических коллекций к созданию национальной системы хранения генофондов растений в вечной мерзлоте. *Информ. вестник ВОГИС*. 2008;12(4):509-523.  
[Goncharov N.P., Shumny V.K. From preservation of genetic collections to organization of National project of plant gene pools conservation in permafrost. *Informatsionnyy vestnik VOGiS = The Herald of Vavilov Society for Geneticists and Breeders*. 2008;12(4):509-523. (in Russian)]
- Демидчик В.В., Шашко А.Ю., Бондаренко В.Ю., Смоликова Г.Н., Пржевальская Д.А., Черныш М.А., Пожванов Г.А., Барковский А.В., Смолич И.И., Соколик А.И., Ю М., Медведев С.С. Феномика растений: фундаментальные основы, программно-аппаратные платформы и методы машинного обучения. *Физиол. растений*. 2020;67(3):227-245. DOI 10.31857/S0015330320030069.  
[Demidchik V.V., Shashko A.Y., Bandarenka V.Y., Smolikova G.N., Przhevalskaya D.A., Charnysh M.A., Pozhvanov G.A., Barkovskiy A.V., Smolich I.I., Sokolik A.I., Yu M., Medvedev S.S., Plant phenomics: fundamental bases, software and hardware platforms, and machine learning. *Russ. J. Plant Physiol.* 2020;67:397-412. DOI 10.1134/S1021443720030061.]
- Дорофеев В.Ф. Внутривидовая классификация пшеницы. *Докл. ВАСХНИЛ*. 1985;9:3-5.  
[Dorofeev V.F. Intraspecific taxonomy of wheat. *Doklady VASKhNIL = Reports of the Academy of Agricultural Sciences*. 1985;9:1-4. (in Russian)]
- Дорофеев В.Ф., Филатенко А.А., Мигушова Э.Ф., Удачин Р.А., Якубцинер М.М. Культурная флора СССР. Т. 1. Пшеница. Л.: Колос, 1979.  
[Dorofeev V.F., Filatenko A.A., Migushova E.F., Udachin R.A., Yakubtsiner M.M. Flora of Cultivated Plants of USSR. Vol. 1. Wheat. Leningrad: Kolos Publ., 1979. (in Russian)]
- Зуев Е.В., Амри А., Брыкова А.Н., Пюккенен В.П., Митрофанова О.П. Атлас разнообразия мягкой пшеницы (*Triticum aestivum* L.) по признакам колоса и зерновки. СПб.: Копи-Р, 2019.  
[Zuev E.V., Amri A., Brykova A.N., Pyukkenen V.P., Mitrofanova O.P. Atlas of the Diversity of Soft Wheat (*Triticum aestivum* L.) by Ear and Grain Characteristics. St. Petersburg: Kopi-R Publ., 2019. (in Russian)]
- Коваль С.Ф. Каталог изогенных линий яровой мягкой пшеницы Новосибирская 67 и принципы их использования в эксперименте. *Генетика*. 1997;33(8):1168-1173.  
[Koval S.F. The catalog of near-isogenic lines of Novosibirskaya-67 common wheat and principles of their use in experiments. *Russ. J. Genet.* 1997;33(8):995-1000.]
- Международный классификатор СЭВ рода *Triticum* L./ Сост. Дорофеев В.Ф., Руденко М.И., Филатенко А.А., Бареш И., Сегналова Я., Леманн Х. Л.: ВИР, 1984.  
[Dorofeev V.F., Rudenko M.I., Filatenko A.A., Baras J., Segnalova J., Lemann H. (Compilers). The International Comecon List of Descriptors for the Genus *Triticum* L. Leningrad: VIR Publ., 1984. (in Russian)]
- Пальмова Е.Ф. Введение в экологию пшениц. М.; Л.: Сельхозгиз, 1935.  
[Palmova E.F. Introduction to Wheat Ecology. Moscow; Leningrad: Selkhozgiz Publ., 1935. (in Russian)]
- Родионов А.В., Амосова А.В., Беляков Е.А., Журбенко П.М., Михайлова Ю.В., Пунина Е.О., Шнеер В.С., Лоскутов И.Г., Муравенко О.В. Генетические последствия межвидовой гибридизации, ее роль в видообразовании и фенотипическом разнообразии растений. *Генетика*. 2019;55(3):255-272. DOI 10.1134/S0016675819030159.  
[Rodionov A.V., Amosova A.V., Belyakov E.A., Zhurbenko P.M., Mikhailova Y.V., Punina E.O., Shneyer V.S., Loskutov I.G., Muravenko O.V. Genetic consequences of interspecific hybridization, its role in speciation and phenotypic diversity of plants. *Russ. J. Genet.* 2019;55(3):278-294. DOI 10.1134/S1022795419030141.]
- Родионов А.В., Шнеер В.С., Гнутиков А.А., Носов Н.Н., Пунина Е.О., Журбенко П.М., Лоскутов И.Г., Муравенко О.В. Диалектика видов: от исходного единообразия, через максимально возможное разнообразие к конечному единообразию. *Бот. журн.* 2020;105(9):835-853. DOI 10.31857/S0006813620070091.  
[Rodionov A.V., Shneyer V.S., Gnutikov A.A., Nosov N.N., Punina E.O., Zhurbenko P.M., Loskutov I.G., Muravenko O.V. Species dialectics: from initial uniformity, through the greatest possible diversity to ultimate uniformity. *Botanicheskiy Zhurnal = Botanical Journal*. 2020;105(9):835-853. DOI 10.31857/S0006813620070091. (in Russian)]
- Романов Б.В., Пимонов К.И. Феномономика продукционных признаков видов пшеницы. пос. Персиановский: Донской ГАУ, 2018.  
[Romanov B.V., Pimonov K.I. Phenomogenomics of Production Traits of Wheat Species. Persianovsky: Donskoy GAU Publ., 2018. (in Russian)]
- Синская Е.Н. Историческая география культурной флоры (На заре земледелия). Ленинград: Колос, 1969.  
[Sinskaya E.N. Historical Geography of Cultural Flora (At the Dawn of Agriculture). Leningrad: Kolos Publ., 1969. (in Russian)]

- Удачин Р.А., Мигушова Э.Ф. Новое в познании рода *Triticum*. *Вестн. с.-х. науки*. 1970(9):20-24.  
[Udachin R.A., Migushova E.F. New in the knowledge of the genus *Triticum*. *Vestnik Selskokhozyaystvennoy Nauki = Herald of Agricultural Sciences*. 1970;9:20-24. (in Russian)]
- Якубцинер М.М. К познанию пшениц Китая. *Бот. журн.* 1959; 44(10):1425-1436.  
[Yakubtsiner M.M. More on Chinese wheats. *Botanicheskiy Zhurnal = Botanical Journal*. 1959;44(10):1425-1436. (in Russian)]
- Chen Z.J. Genetic and epigenetic mechanisms for gene expression and phenotypic variation in plant polyploids. *Annu. Rev. Plant Biol.* 2007;58:377-406. DOI 10.1146/annurev.arplant.58.032806.103835.
- Comai L. The advantages and disadvantages of being polyploid. *Nat. Rev. Genet.* 2005;6(11):836-846. DOI 10.1038/nrg1711.
- Finigan P., Tanurdzic M., Martienssen R.A. Origins of novel phenotypic variation in polyploids. In: *Polyploidy and Genome Evolution*. Berlin; Heidelberg: Springer Press, 2012;57-76. DOI 10.1007/978-3-642-31442-1\_4.
- Genяев М.А., Комышев Е.Г., Smirnov N.V., Kruchinina Y.V., Goncharov N.P., Afonnikov D.A. Morphometry of the wheat spike by analyzing 2D images. *Agronomy*. 2019;9(7):390. DOI 10.3390/agronomy9070390.
- Goncharov N.P. Comparative-genetic analysis – a base for wheat taxonomy revision. *Czech J. Genet. Plant Breed.* 2005;41:52-55.
- Goncharov N.P. Genus *Triticum* L. taxonomy: the present and the future. *Plant Syst. Evol.* 2011;295(1-4):1-11. DOI 10.1007/s00606-011-0480-9.
- Hammer K., Filatenko A.A., Pistrick K. Taxonomic remarks on *Triticum* L. and  $\times$ *Triticosecale* Wittm. *Genet. Resour. Crop Evol.* 2011; 58(1):3-10. DOI 10.1007/s10722-010-9590-4.
- Johnson S.C. Hierarchical clustering schemes. *Psychometrika*. 1967; 32(3):241-254.
- Liu W., Zheng Y., Song S., Huo B., Li D., Wang J. *In vitro* induction of allohexaploid and resulting phenotypic variation in *Populus*. *Plant Cell Tiss. Organ Cult.* 2018;134(2):183-192. DOI 10.1007/s11240-018-1411-z.
- Müllner D. Modern hierarchical, agglomerative clustering algorithms. *arXiv*. 2011;1109.2378.
- Pedregosa F., Varoquaux G., Gramfort A., Michel V., Thirion B., Grisel O., Blondel M., Müller A., Nothman J., Louppe G., Prettenhofer P., Weiss R., Dubourg V., Vanderplas J., Passos A., Cournapeau D., Brucher M., Perrot M., Duchesnay E. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* 2011;12:2825-2830.
- Robinson D.O., Coate J.E., Singh A., Hong L., Bush M., Doyle J.J., Roeder A.H. Ploidy and size at multiple scales in the *Arabidopsis* sepal. *Plant Cell*. 2018;30(10):2308-2329. DOI 10.1105/tpc.18.00344.
- Snedecor G.W., Cochran W.G. *Statistical Methods*. Ames, Iowa: Iowa State University Press, 1989.
- Tan F., Tu H., Liang W., Long J.M., Wu X.M., Zhang H.Y., Guo W.W. Comparative metabolic and transcriptional analysis of a doubled diploid and its diploid citrus rootstock (*C. junos* cv. Ziyang xiangcheng) suggests its potential value for stress resistance improvement. *BMC Plant Biol.* 2015;15:89. DOI 10.1186/s12870-015-0450-4.
- van der Maaten L., Hinton G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* 2008;9:2579-2605.
- van Slageren M., Payne T. Concepts and nomenclature of the Farro wheats, with special reference to Emmer, *Triticum turgidum* subsp. *dicoccum* (Poaceae). *Kew Bull.* 2013;68:477-494. DOI 10.1007/S12225-013-9459-8.
- Virtanen P., Gommers R., Oliphant T.E., Haberland M., Reddy T., Cournapeau D., Burovski E., Peterson P., Weckesser W., Bright J., van der Walt S.J., Brett M., Jones E., Kern R., Larson E., Carey C.J., Polat I., Feng Yu, Moore E.W., VanderPlas J., Laxalde D., Perktold J., Cimrman R., Henriksen I., Quintero E.A., Harris C.R., Archibald A.M., Riberio A.H., Pedregosa F., van Mulbregt P. SciPy 1.0 Contributors. SciPy 1.0 – fundamental algorithms for scientific computing in Python. *Nat. Meth.* 2020;17(3):261-272. DOI 10.1038/s41592-019-0686-2.
- Watanabe N., Bannikova S.V., Goncharov N.P. Inheritance and chromosomal location of the genes for long glume phenotype found in Portuguese landraces of hexaploid wheat, 'Arrancada'. *J. Genet. Breed.* 2004;58:273-278.
- Williams J.H., Oliveira P.E. For things to stay the same, things must change: polyploidy and pollen tube growth rates. *Ann. Bot.* 2020; 125(6):925-935. DOI 10.1093/aob/mcaa007.
- Yang W., Feng H., Zhang X., Zhang J., Doonan J.H., Batchelor W.D., Xiong L., Yan J. Crop phenomics and high-throughput phenotyping: past decades, current challenges, and future perspectives. *Mol. Plant*. 2020;13(2):187-214. DOI 10.1016/j.molp.2020.01.008.
- Zatybekov A., Anuarbek S., Abugaliev S., Turuspekov Y. Phenotypic and genetic variability of a tetraploid wheat collection grown in Kazakhstan. *Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding*. 2020;24(6):605-612. DOI 10.18699/VJ20.654.
- Zhang Y., Zhao C., Du J., Guo X., Wen W., Gu S., Wang J., Fan J. Crop phenomics: current status and perspectives. *Front. Plant Sci.* 2019; 10:714. DOI 10.3389/fpls.2019.00714.

#### ORCID ID

A.Yu. Pronozin orcid.org/0000-0002-3011-6288  
Y.V. Kruchinina orcid.org/0000-0002-1084-9521

**Благодарности.** Подготовка образцов колосьев, фенотипирование, разработка алгоритмов анализа формы и классификации выполнены за счет финансирования Курчатковского геномного центра Федерального исследовательского центра ИЦиГ СО РАН, соглашение с Министерством образования и науки РФ № 075-15-2019-1662. Выращивание экспериментальных растений и определение *de visu* их видовой принадлежности по признакам, определяющим архитектуру колоса, поддержано грантом РНФ 16-16-10021. Обработка данных проведена с использованием ресурсов ЦКП «Биоинформатика» в рамках бюджетного проекта № 0259-2021-0009. Работа авторов А.А.П., Е.А.З., А.Ю.П., Н.М.П. осуществлена при поддержке Математического центра в Академгородке, соглашение с Министерством науки и высшего образования Российской Федерации № 075-15-2019-1675. Авторы благодарны Д.А. Афонникову (ИЦиГ СО РАН, г. Новосибирск) за замечания и рекомендации в процессе работы и И.Г. Чухиной (ВИР, г. Санкт-Петербург) за предоставление фотографий гербарных экземпляров пшеницы Петропавловского.

**Прозрачность финансовой деятельности.** Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

**Конфликт интересов.** Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию 27.10.2020. После доработки 31.12.2020. Принята к публикации 02.01.2021.