

ЦИФРОВОЙ ГЕРБАРИЙ ЦСБС СО РАН В ГЛОБАЛЬНОЙ ИНФОРМАЦИОННОЙ СИСТЕМЕ ПО БИОРАЗНООБРАЗИЮ

Н.К. Ковтонюк, И.В. Хан, Е.А. Гатилова, Ю.А. Пшеничкина

Центральный сибирский ботанический сад СО РАН,
630090, Новосибирск, ул. Золотодолинская, 101, e-mail: nkovtonyuk@csbg.nsc.ru

Цифровой гербарий Центрального сибирского ботанического сада СО РАН начал создаваться в 2017 г. по таксономическому принципу. Полностью оцифрованы гербарные образцы коллекций NS и NSK по родам *Allium*, *Geranium*, *Medicago*, *Rhododendron*, семейству *Primulaceae*, редкие растения флоры Сибири, а также типовые образцы высших сосудистых растений, хранящиеся в ЦСБС. Результаты оцифровки опубликованы в открытом доступе в виде 6 датасетов на портале Global Biodiversity Information Facility (GBIF) с помощью пакета программ Integrated Publishing Toolkit (IPT), установленном на сервере ЦСБС. Оцифровка ведется на двух сканерах ObjectScan 1600 (Microtek) с помощью программ ScanWizard-Botany и MiVapp-Botany по международным стандартам: с баркодом, цветовой шкалой, масштабной линейкой и при оптическом разрешении в 600 dpi. Цифровой гербарий создается как сетевой ресурс по биоразнообразию, материалы которого могут использоваться в GBIF, для таксономических обработок при подготовке “Флоры России”, в различных видах научных исследований и в образовательных целях.

Ключевые слова: биоресурсные коллекции, глобальная информационная система (GBIF) по биоразнообразию (IPT), датасет, редкие растения, флора России, цифровой гербарий, *Allium*, *Geranium*, *Medicago*, *Rhododendron*, *Primulaceae*.

DIGITAL HERBARIUM OF THE CSBG SB RAS IN GLOBAL BIODIVERSITY INFORMATION FACILITIES

N.K. Kovtonyuk, I.V. Han, E.A. Gatilova, Yu.A. Pshenichkina

Central Siberian Botanical Garden, SB RAS,
630090, Novosibirsk, Zolotodolinskaya str., 101, e-mail: nkovtonyuk@csbg.nsc.ru

The Digital Herbarium of the Central Siberian Botanical Garden of SB RAS began to be created on the taxonomic basis in 2017. Completely digitized herbarium specimens of genera *Allium*, *Geranium*, *Medicago*, *Rhododendron*, family *Primulaceae*, rare plants of Siberian flora, as well as type specimens of vascular plants stored in NS and NSK collections at the CSBG SB RAS. The digitization results are published in the open access as 6 datasets on the Global Biodiversity Information Facility (GBIF) portal using the Integrated Publishing Toolkit (IPT) installed on the CSBG server. Digitization is performed on two ObjectScan 1600 scanners (Microtek) using software ScanWizard-Botany and MiVapp-Botany according to the international standards, at 600 dpi, with a barcode, 24-color scale and spatial scale bar. The Digital Herbarium is created as a network resource on biodiversity, the materials of which can be used in GBIF, for taxonomic treatments in the preparation of “Flora of Russia”, in various types of scientific research and educational purposes.

Key words: bioresource collections, Global Biodiversity Information Facility (GBIF), IPT, dataset, rare plants, Flora of Russia, Digital Herbarium, *Allium*, *Geranium*, *Medicago*, *Rhododendron*, *Primulaceae*.

ВВЕДЕНИЕ

Согласно ежегодному отчету директора Гербария Нью-Йоркского ботанического сада (NYBG) и редактора Index Herbariorum В.М. Thiers (2019), на 1 декабря 2018 г. в 178 странах мира активно действуют 3095 гербариев, в которых хранится 387 513 053 образцов. Крупнейшие гербарные коллекции мира уже оцифрованы (Ковтонюк, 2015, 2017; Thiers et al., 2016; Takano et al., 2019). В России зарегистрировано 116 гербарных коллекций с об-

щим фондом около 16 млн образцов. Из них по данным на сентябрь 2019 г. действующие цифровые гербарии, постоянно обновляющие свою информацию, имеют: Московский государственный университет (MW): 1 012 061 образцов, 1 003 480 изображений (<https://plant.depo.msu.ru/>); Центральный сибирский ботанический сад СО РАН (NS, NSK): 680 000 гербарных листов, изображений 33 500, из них в цифровом гербарии 31 225

(<http://herb.csbg.nsc.ru:8081>); Ботанический сад-институт, Владивосток (VBGI): 56 176 образцов, 31 556 изображений: в том числе гербарии ABGI, KAM, SABZ, SAK, SAKH (<http://botsad.ru/herbarium/>); Институт экологии растений и животных УрО РАН, Екатеринбург (SVER): 10 920 образцов и изображений высших сосудистых растений (<https://herbarium.ipae.uran.ru/index.php>) Главный ботанический сад, Москва (МНА) размещает свои образцы на сайте Цифрового гербария МГУ. Недавно Ботанический институт им. В.Л. Комарова (LE) открыл виртуальный доступ к своим коллекциям (<http://herbariumle.ru>).

Гербарии имеют давние традиции общения и сотрудничества. Доступность гербарных коллекций для научных, образовательных целей значительно возросла с появлением цифровых технологий и возможностью сканировать гербарные образцы, создавать цифровые или виртуальные гербарии с доступом по сети интернет. Это современный тренд гербарного дела (Ковтонюк, 2015; Ковтонюк и др., 2018, 2019; Barber et al., 2013; Kislov et al., 2017; James et al., 2018; Kovtonyuk et al., 2018, 2019a; Seregin, 2018).

Цифровой гербарий документирует мировое биоразнообразие (Meineke et al., 2018) и может использоваться в качестве основы для разных типов исследовательских и образовательных проектов, для фенологических исследований (Ellwood et al., 2019), для таксономических обработок при подготовке различных флор, в том числе и “Флоры России”, для изучения ареалов редких и инвазивных видов.

Глобальная информационная система по биоразнообразию (Global Biodiversity Information Facility, GBIF) – это международная сеть и исследовательская инфраструктура, финансируемая правительствами многих стран мира и направленная на предоставление любому человеку, в любом месте, открытого интернет-доступа к данным обо всех типах жизни на Земле через единый портал gbif.org.

Координируемая через свой секретариат в Копенгагене сеть GBIF стран-участниц и организаций, работающих через узлы IPT, обеспечивает учреждения, занимающиеся хранением данных по всему миру, общими стандартами и инструментами с открытым исходным кодом, которые позволяют им обмениваться информацией о том, где и когда были зарегистрированы виды. Эта информация аккумулируется из многих источников, начиная с биоресурсных коллекций, оцифрованных гербарных и музейных образцов, собранных в 17–21 вв., до фотографий с геотегами со смартфонов, которыми поделились любители-натуралисты в последние дни и недели. Сеть GBIF объединяет все эти источники за счет использования стандартов данных в формате Darwin Core, которые формируют основу для сотен миллионов записей о встречаемости видов. Издатели предоставляют открытый доступ к своим наборам данных, что позволяет ученым, исследователям и другим лицам ежегодно применять эти данные в сотнях рецензируемых публикаций и программных документов.

Гербарные коллекции ЦСБС СО РАН (NS и NSK) хранят обширные уникальные коллекции высших сосудистых растений, мхов, лишайников, грибов и водорослей, собранных в различных регионах Сибири, Дальнего Востока, европейской части России, Зарубежной Европы и Зарубежной Азии, Северной Америки и Австралии. Оцифровка типовых образцов началась в 2014 г. специализированным сканером Herbscan (Ковтонюк, 2015; Kovtonyuk et al., 2019h). Сканирование фондовых коллекций, зарегистрированных как USU 440537, и создание Цифрового гербария ведутся с декабря 2017 г. группой УНУ-Гербарий ЦСБС и группой научно-образовательных программ ЦСБС СО РАН. Результаты оцифровки биоресурсных коллекций публикуются в открытом доступе на портале Global Biodiversity Information Facility – GBIF (gbif.org).

МАТЕРИАЛ И МЕТОДЫ

В качестве материала для исследования использовали оцифрованные гербарные образцы высших сосудистых растений из коллекций Гербария им. М.Г. Попова (NSK) и Гербария им. И.М. Красноборова (NS). Оцифровку гербарных листов проводили на двух сканерах ObjectScan 1600 (<http://www.microtek.com/products.php?KindID=12&ID=373>) с использованием специализированных лицензионных программ: ScanWizard-Botany – для автоматизации процесса сканирования и распознавания текста этикеток (OCR) и MiVapp-Botany – для хранения и аутентификации

оцифрованных изображений, составления базы данных с уменьшенным изображением гербарных листов (рис. 1). Обе программы позволяют создать цифровой гербарий. Оцифровку выполняли по международным стандартам: с оптическим разрешением в 600 dpi, на каждом гербарном листе имеется уникальный штрихкод (barcode) с акронимом гербария и 7-значным порядковым номером гербарного листа, сканирование проводится в сопровождении 24-цветной шкалы и масштабной линейки. Цифровое изображение гербарного листа сохраняется в формате *.tif, объемом около

CSBG SBRAS

Verification

Taxonomy

Roles

Language

Keyword

Corydalis

Search

Previous Page

1

2

3

4

5

Next Page

63

	Thumbnail	Specimen ID	Family	Scientific Name	Collector	Collection Date
46		NS0020643	Papaveraceae(Papaveraceae)	Corydalis stricta Steph. ex DC.	Верещагин В.И.	1907-06-00
47		NS0020647	Papaveraceae(Papaveraceae)	Corydalis stricta Steph. ex DC.	Данилов М.П., Фризен Н.В.	1982-06-24
48		NS0020646	Papaveraceae(Papaveraceae)	Corydalis stricta Steph. ex DC.	Данилов М.П., Доронькин В.М.	1982-07-01
49		NS0020645	Papaveraceae(Papaveraceae)	Corydalis stricta Steph. ex DC.	Тимохина С.А.	1964-06-27

Рис. 1. Цифровой гербарий ЦСБС СО РАН (<http://herb.csbg.nsc.ru:8081>).

200 Мбайт, (может высылаться по запросу) и в формате *.jpg, доступном для скачивания пользователями. Информация об оцифрованных гербарных образцах разносится по 25 столбцам в таблице Calc из пакета Open Office и в дальнейшем публикуется в формате Darwin Core на портале GBIF (gbif.org) в открытом доступе. Процесс сканирования и верификации образцов разделен. Вновь отсканированные образцы сначала поступают на верификацию специалистам и лишь потом отправляются в Цифровой гербарий. Пошаговую методику оцифровки на сканерах ObjectScan 1600 разрабатывала к.б.н. Е.А. Гатилова. Координаты точек сбора гербарных образцов выверяет к.б.н. И.В. Хан, используя ресурсы Яндекс-карт (<https://yandex.ru/maps>), а также сайта “Это место” (<http://www.etomesto.ru>), она же ведет таблицу Calc и готовит csv-файл в Darwin Core для публикации

в GBIF. Русские названия таксонов в цифровой гербарий вносит к.б.н. Ю.А. Пшеничкина. Публикация результатов оцифровки в виде массива данных на портале GBIF осуществляется с помощью пакета программ The Integrated Publishing Toolkit (IPT, <https://www.gbif.org/ipt>), установленном на сервере ЦСБС СО РАН. GBIF поддерживает публикацию данных, содержащих описания коллекций или тематических электронных ресурсов, таксономические сводки, данные о находках видов (коллекционных образцах), а также результаты площадных и маршрутных учетов и данные мониторинга. Электронным публикациям в GBIF присваивается DOI, ограничений на объем публикуемых данных нет, один датасет может включать как 10, так и 10 000 записей, или же содержать только описание данных, если они еще не оцифрованы (<http://gbif.ru/datapublish>).

РЕЗУЛЬТАТЫ

Цифровой гербарий ЦСБС СО РАН постоянно пополняется и к октябрю 2019 г. насчитывает более 30 тыс. образцов. Полностью оцифрованы гербарные образцы коллекций NS и NSK по роду *Allium* L., это 6210 гербарных листов (рис. 2), из них 85 % с геолокацией точек сбора, данные опубликованы в GBIF (Kovtonyuk et al., 2019f). Оциф-

рован и полностью внесен в цифровой гербарий род *Medicago* L., 1035 гербарных листов, из них 63 % с координатами, данные опубликованы в GBIF (Kovtonyuk et al., 2019d). Род *Rhododendron* L. – 1355 гербарных листов, из них 71 % с координатами, также полностью оцифрованы и опубликованы в GBIF (Kovtonyuk et al., 2019g). Семей-

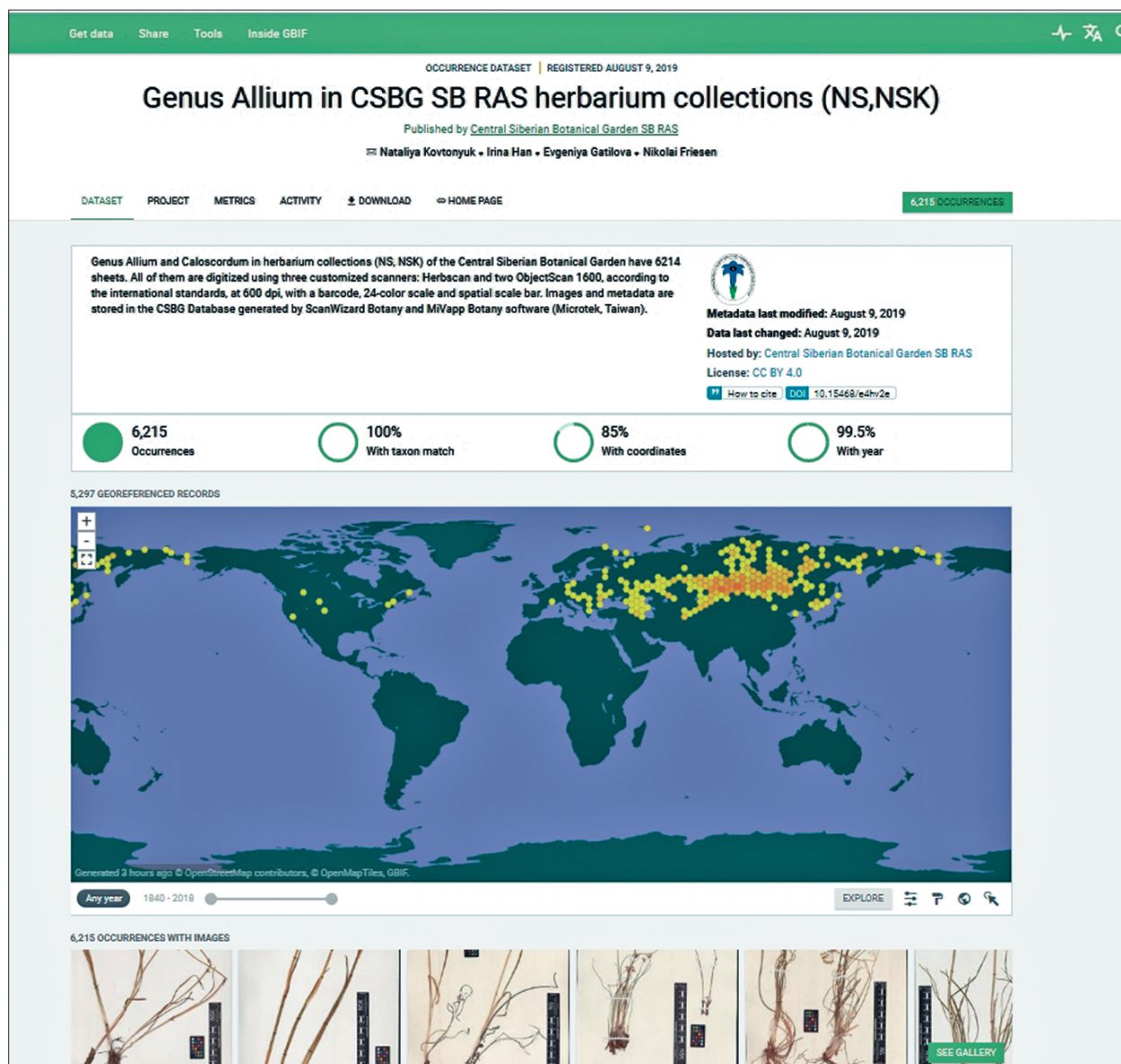


Рис. 2. Публикация результатов оцифровки рода *Allium* из коллекций NS и NSK на портале GBIF.

ство *Primulaceae*, NS – 1926 гербарных листов и 98 % с координатами (Kovtonyuk et al., 2019b), в коллекции NSK – 3209 гербарных листов, из них 98 % с координатами (Kovtonyuk et al., 2019c).

Полностью оцифрованы фондовые коллекции NSK и NS по роду *Geranium* (2686 гербарных листов). Оцифрованные изображения 872 типовых гербарных образцов из коллекции NSK (Kovtonyuk et al., 2019h) и NS (490) также размещены в цифровом гербарии ЦСБС СО РАН. Наибольшее количество гербарных образцов оцифровано из следующих 20 семейств, хранящихся в фондовых коллекциях NS и NSK: *Amaryllidaceae* (6223); *Primulaceae* (5137); *Geraniaceae* (2802); *Fabaceae* (2115); *Ericaceae* (1616); *Poaceae* (1415); *Asteraceae*

(1324); *Orchidaceae* (799); *Cyperaceae* (856); *Ranunculaceae* (649); *Rosaceae* (443); *Amaranthaceae* (459); *Boraginaceae* (604); *Brassicaceae* (351); *Caryophyllaceae* (286); *Liliaceae* (232); *Dryopteridaceae* (219); *Betulaceae* (252); *Lamiaceae* (231); *Apiaceae* (226).

Копии цифровых изображений гербарных образцов в формате *.tif и метаданные к ним передаются в хранилище биоресурсных коллекций в Институт цитологии и генетики СО РАН на ответственное хранение. После публикации данных по оцифровке гербария в виде датасета в GBIF через некоторое время на портале появляется карта с точками сборов, количеством данных с геолокацией и указанием даты сборов (см. рис. 2), а также статистические данные (рис. 3).

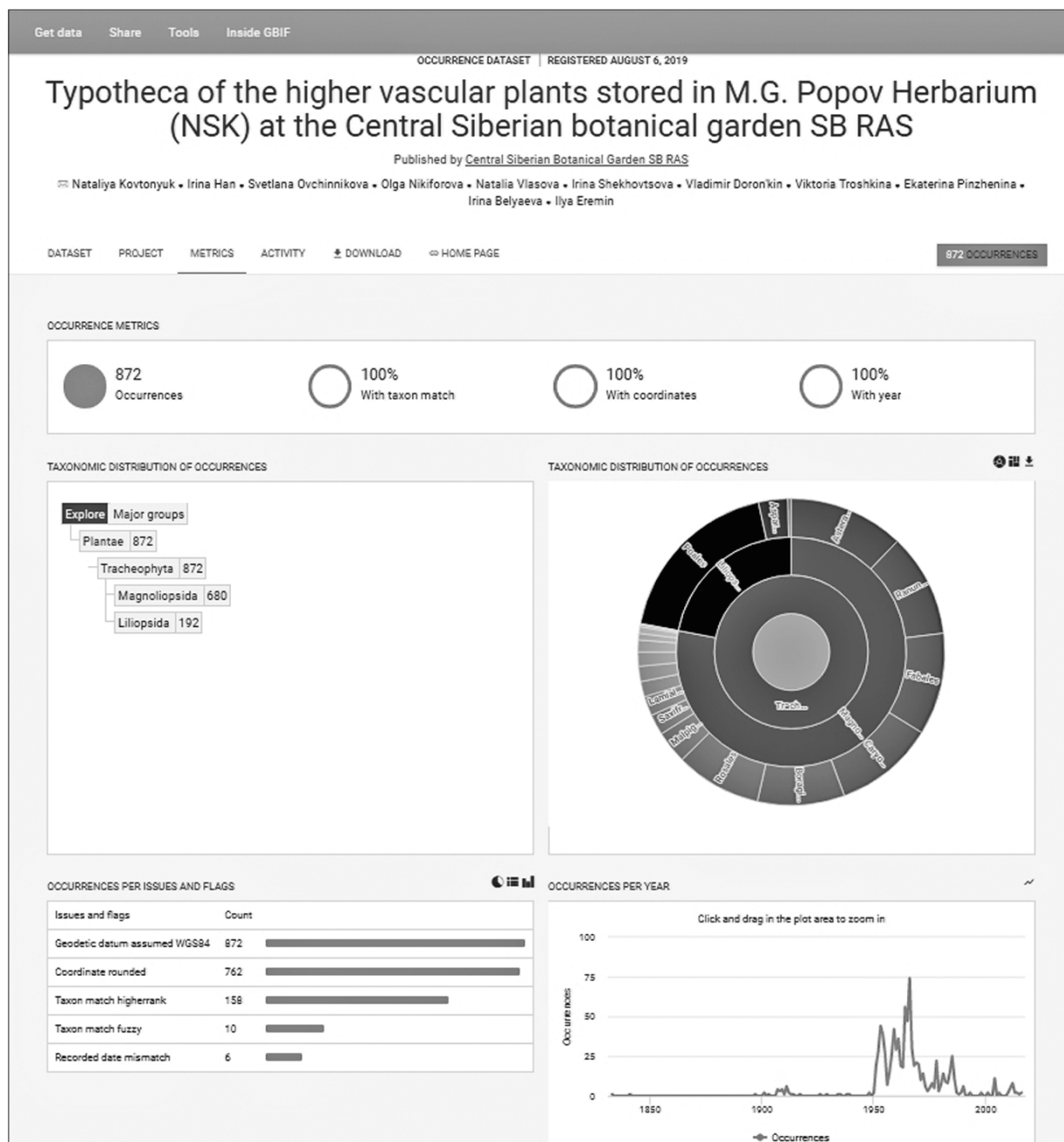


Рис. 3. Статистические показатели публикаций в GBIF: количество таксонов, доля сборов (%) с координатами, количество сборов по годам.

ЗАКЛЮЧЕНИЕ

К настоящему времени ЦСБС СО РАН опубликовал в GBIF 6 датасетов по результатам оцифровки 14 612 гербарных листов из цифрового гербария ЦСБС СО РАН, уделив особое внимание отдельным таксонам. ЦСБС установил на своем сервере пакет IPT-программ для публикации в GBIF и предлагает всем ботаническим учреждениям азиатской части России публиковать данные. Публикация этих сведений на портале GBIF позволяет международному сообществу обмениваться информацией по биоразнообразию, проводить таксономические и флористические исследования,

использовать ресурс GBIF в образовательных целях, проводить статистическую обработку данных по распространению видов, составить ареалы и др.

Благодарности. Выражаем свою признательность за техническую поддержку в публикации данных в GBIF и установку сибирского IPT М. Шашкову, Н. Ивановой. Благодарим сотрудников группы УНУ-Гербарий Л.З. Лукманову, И.М. Деюн, С.А. Красникову и сотрудников группы научно-образовательных программ А.А. Басалаеву, Ю.Ю. Февролину и А.Г. Каулинь за помощь в сканировании гербарных образцов.

ЛИТЕРАТУРА

- Ковтонуик Н.К.** Виртуальная коллекция типовых образцов гербария имени М.Г. Попова // Раст. мир Азиатской России. 2015. № 3 (19). С. 88–93.
- Ковтонуик Н.К.** Виртуальные гербарные коллекции как ресурс для изучения таксономии и биоразнообразия // Раст. Азиатской России. 2017. № 1 (25). С. 98–104. DOI: 10.21782/RMAR1995-2449-2017-1(98-104)
- Ковтонуик Н.К., Хан И.В., Гатилова Е.А.** Семейство *Primulaceae* в виртуальном гербарии Центрального сибирского ботанического сада СО РАН // Раст. мир Азиатской России. 2018. № 4 (32). С. 19–29. DOI: 10.21782/RMAR1995-2449-2018-4(19-29)
- Ковтонуик Н.К., Хан И.В., Гатилова Е.А.** Вклад геоботаников в создание гербарных коллекций Центрального сибирского ботанического сада СО РАН // Итоги и перспективы геоботанических исследований в Сибири: материалы Всерос. науч.-практ. конф., посвящ. 75-летию лаборатории экологии и геоботаники ЦСБС СО РАН (Новосибирск, 13–17 мая 2019 г.). Новосибирск, 2019. С. 49–50. https://elibrary.ru/download/elibrary_38050444_41630685.pdf
- Barber A., Lafferty D., Landrum L.R.** The SALIX Method: A semi-automated workflow for herbarium specimen digitization // Taxon. 2013. V. 62, No. 3. P. 581–590. <https://doi.org/10.12705/623.16>
- Ellwood E.R., Primack R.B., Willis C.G., HilleRis-Lambers J.** Phenology models using herbarium specimens are only slightly improved by using finer-scale stages of reproduction // Appl. Plant Sci. 2019. V. 7, No. 3. e1225. P. 1–9. DOI: 10.1002/aps3.1225
- James S.A., Soltis P.S., Belbin L., Chapman A.D., Nelson G., Paul D.L., Collins M.** Herbarium data: Global biodiversity and societal botanical needs for novel research // Appl. Plant Sci. 2018. V. 6, No. 2. e1024. <https://doi.org/10.1002/aps3.1024>
- Kislov D.E., Bakalin V.A., Pimenova E.A., Verkholat V.P., Krestov P.V.** An electronic management system for a digital herbarium: development and future prospects // Bot. Pacifica. 2017. No. 2. P. 59–68. DOI: 10.17581/bp.2017.06207
- Kovtonyuk N., Han I., Gatilova E.** Digitization of vascular plant herbarium collections at the Central Siberian Botanical Garden, Novosibirsk, Russia // Skvortsovia. 2018. V. 4, No. 3. P. 100–111. <http://skvortsovia.uran.ru/2018/4302.pdf>
- Kovtonyuk N., Han I., Gatilova E.** Digital Herbarium Collections of the Central Siberian Botanical Garden SB RAS, Novosibirsk, Russia // I. Bychkov, V. Voronin (Eds.). Information Technologies in the Research of Biodiversity. Springer Proceedings in Earth and Environmental Sciences. Springer, Cham. 2019a. P. 22–27. DOI: https://doi.org/10.1007/978-3-030-11720-7_4
- Kovtonyuk N., Han I., Gatilova E.** Family *Primulaceae* in Digital Herbaria of CSBG SB RAS (NS). Central Siberian Botanical Garden SB RAS. 2019b. Occurrence dataset <https://doi.org/10.15468/be6owh> accessed via GBIF.org on 2019-08-12.
- Kovtonyuk N., Han I., Gatilova E.** Family *Primulaceae* in Digital Herbaria of CSBG SB RAS (NSK). Central Siberian Botanical Garden SB RAS. 2019c. Occurrence dataset <https://doi.org/10.15468/zw7jnn> accessed via GBIF.org on 2019-08-29.
- Kovtonyuk N., Han I., Gatilova E.** Genus *Medicago* (*Fabaceae*) in CSBG Herbarium collections (NS, NSK). Version 1.3. Central Siberian Botanical Garden SB RAS. 2019d. Occurrence dataset <https://doi.org/10.15468/jvrxeh> accessed via GBIF.org on 2019-08-12.
- Kovtonyuk N., Han I., Gatilova E., Friesen N.** Genus *Allium* in CSBG SB RAS herbarium collections (NS, NSK). Central Siberian Botanical Garden SB RAS. 2019f. Occurrence dataset <https://doi.org/10.15468/e4hv2e> accessed via GBIF.org on 2019-08-12.
- Kovtonyuk N., Han I., Gatilova E., Karakulov A.** Genus *Rhododendron* in Digital Herbarium of CSBG SB RAS (NS, NSK). Central Siberian Botanical Garden SB RAS. 2019g. Occurrence dataset <https://doi.org/10.15468/67ouin> accessed via GBIF.org on 2019-08-12.
- Kovtonyuk N., Han I., Ovchinnikova S., Nikiforova O., Vlasova N., Shekhovtsova I., Doron'kin V., Troshkina V., Pinzhennina E., Belyaeva I., Eremin I.** Typotheca of the higher vascular plants stored in M.G. Popov Herbarium (NSK) at the Central Siberian Botanical Garden SB RAS. Version 1.4. Central Siberian Botanical Garden SB RAS. 2019h. Occurrence dataset <https://doi.org/10.15468/qdfdqj> accessed via GBIF.org on 2019-08-07.
- Meineke E.K., Davies T.J., Daru B.H., Davis C.C.** Biological collections for understanding biodiversity in the Anthropocene // Phil. Trans. Roy. Soc. B: Biological Sciences. 2018. V. 374, No. 20170386. P. 1–9. DOI: 10.1098/rstb.2017.0386
- Seregin A.** The largest Digital Herbarium in Russia is now available online // Taxon. 2018. V. 67, No. 2. P. 463–467.
- Takano A., Horiuchi Y., Fujimoto Y., Aoki K., Mitsuhashi H., Takahashi A.** Simple but long-lasting: A specimen imaging method applicable for small- and medium-sized herbaria // PhytoKeys. 2019. V. 118. P. 1–14. <https://doi.org/10.3897/phytokeys.118.29434>
- Thiers B.M.** The World's Herbaria 2018: A Summary Report Based on Data from Index Herbariorum Issue 3.0, published January 10, 2019. http://sweetgum.nybg.org/science/docs/The_Worlds_Herbaria_2018.pdf
- Thiers B.M., Tulig M.C., Watson K.A.** Digitization of the New York Botanical Garden Herbarium // Brittonia, 2016. V. 68, No. 3. P. 324–333. DOI: 10.1007/s12228-016-9423-7

Поступила в редакцию 16.08.2019 г.,
после доработки – 16.08.2019 г.,
принята к публикации 15.10.2019 г.