

IRSTI 34.01.11

DOI: <https://doi.org/10.62687/NSJ.2.2.2026.27>**DIGITALIZATION OF THE COLLECTIONS IN THE HERBARIUM FUND OF THE DEPARTMENT OF BIOLOGY AND ITS PLACEMENT ON THE GBIF PLATFORM**¹**G.S. Tleubergenova**^{ID}, ²**M.Zh. Sharipova**^{ID}*, ¹**E.V. Galaktionova**^{ID}¹Kozybayev North Kazakhstan University, Petropavl, Republic of Kazakhstan²Astana International University, Astana, Kazakhstan

*e-mail: mzhakypzhan@mail.ru

G.S. Tleubergenova – candidate of Biological Sciences, Associate Professor, Department of Biology, Kozybayev North Kazakhstan University, Petropavl, Republic of Kazakhstan, e-mail: gbattalova@ku.edu.kz, <https://orcid.org/0000-0002-1001-9528>

M.Zh. Sharipova – PhD, Associate Professor, Astana International University, Astana, Kazakhstan, e-mail: mzhakypzhan@mail.ru, <https://orcid.org/0000-0002-9467-1483>

E.V. Galaktionova – Senior Lecturer, Master of Biology, Kozybayev North Kazakhstan University, Petropavl, Republic of Kazakhstan, e-mail: evgalaktionova@ku.edu.kz, <https://orcid.org/0009-0007-9110-9747>

Abstract. Herbarium collections held by regional universities in Kazakhstan largely remain outside the international scientific circulation due to the lack of digital data formats. This study addresses this gap with respect to the herbarium fund of the Department of Biology at M. Kozybayev North Kazakhstan University. The research material comprised 925 herbarium specimens from nine families – *Boraginaceae*, *Labiatae*, *Solanaceae*, *Scrophulariaceae*, *Orchidaceae*, *Liliaceae*, *Ericaceae*, *Convallariaceae*, *Fabaceae*, and *Asteraceae* – collected between 1957 and 2025. The collection includes both historical specimens from the mid-twentieth century and recent field material. Digitization was carried out in accordance with the international Darwin Core (DwC) standard. Data integrity was verified using the GBIF validator, after which the dataset was published via the IPT platform and made available on the global GBIF portal, ensuring its openness and integration into international biodiversity data infrastructure. Taxonomic analysis revealed a predominance of species characteristic of forest, meadow, and meadow-steppe communities, consistent with the vegetation patterns typical of the region. The temporal coverage of the collection spans more than 70 years, allowing the dynamics of floristic collecting to be traced over an extended period. Geographic analysis showed an uneven distribution of records across administrative districts, with the largest number of specimens originating from the Kyzylzhar district of the North Kazakhstan region. The resulting digital dataset is of value for floristic, ecological-geographic, and taxonomic research, contributes to the preservation of the herbarium fund, and expands its potential use in education and nature conservation.

Keywords: flora, herbarium collections; digital databases; digitization; GBIF; Darwin Core; IPT; geolocation.

ЦИФРОВИЗАЦИЯ КОЛЛЕКЦИЙ ГЕРБАРНОГО ФОНДА КАФЕДРЫ БИОЛОГИИ И ИХ РАЗМЕЩЕНИЕ НА ПЛАТФОРМЕ GBIF¹**Г.С. Тлеубергенова**, ²**М.Ж. Шарипова***, ¹**Е.В. Галактионова**¹НАО "Северо-Казахстанский университет имени М. Козыбаева", Петропавловск, Казахстан²Международный университет Астана, Астана, Казахстан

*e-mail: mzhakypzhan@mail.ru

Г.С. Тлеубергенова – кандидат биологических наук, доцент кафедры биологии, Северо-Казахстанский университет имени М. Козыбаева, Петропавловск, Казахстан, e-mail: gbattalova@ku.edu.kz, <https://orcid.org/0000-0002-1001-9528>

М.Ж. Шарипова – PhD, ассоциированный профессор Международного университета Астана, Астана, Казахстан, e-mail: mzhakypzhan@mail.ru, <https://orcid.org/0000-0002-9467-1483>

Е.В. Галактионова – старший преподаватель, магистр биологии, Северо-Казахстанский университет имени М. Козыбаева, Петропавловск, Казахстан, e-mail: evgalaktionova@ku.edu.kz, <https://orcid.org/0009-0007-9110-9747>

Аннотация. Гербарные коллекции региональных вузов Казахстана в большинстве своём остаются вне международного научного оборота из-за отсутствия цифровых форматов представления данных. Настоящая работа направлена на восполнение этого пробела применительно к фондам кафедры "Биология" Северо-Казахстанского университета имени М. Козыбаева. Материалом исследования послужили 925 гербарных образцов девяти семейств – *Boraginaceae*, *Labiatae*, *Solanaceae*, *Scrophulariaceae*, *Orchidaceae*, *Liliaceae*, *Ericaceae*, *Convallariaceae*, *Fabaceae* и *Asteraceae*, собранных в период с 1957 по 2025 год. Коллекция включает как исторические сборы середины прошлого века, так и материалы недавних полевых исследований. Оцифровка выполнена в соответствии с международным стандартом Darwin Core (DwC). Корректность структуры данных подтверждена валидатором GBIF, после чего набор был размещён на платформе IPT и опубликован на глобальном портале GBIF, обеспечив тем самым его открытость и включение в международную биоинформационную инфраструктуру. Анализ таксономического состава показал преобладание видов, характерных для лесных, луговых и лугово-степных сообществ, что соответствует особенностям растительного покрова региона. Временной охват коллекции превышает 70 лет, что позволяет проследить динамику флористических сборов на протяжении длительного периода. Географический анализ выявил неравномерность распределения находок по административным районам: наибольшее число образцов происходит из Кызылжарского района Северо-Казахстанской области. Сформированный цифровой ресурс представляет ценность для флористических, эколого-географических и таксономических исследований, способствует сохранности гербарного фонда и расширяет перспективы его использования в образовательной и природоохранной деятельности.

Ключевые слова: флора, гербарные коллекции; цифровые базы данных; оцифровка; GBIF; Darwin Core; IPT; геопривязка.

БИОЛОГИЯ КАФЕДРАСЫНЫҢ ГЕРБАРИЙ ҚОРЫ КОЛЛЕКЦИЯЛАРЫН ЦИФРЛАНДЫРУ ЖӘНЕ ОЛАРДЫ GBIF ПЛАТФОРМАСЫНДА ОРНАЛАСТЫРУ

¹Г.С. Тілеубергенова, ²М.Ж. Шарипова*, ¹Е.В. Галактионова

¹М. Қозыбаев атындағы Солтүстік Қазақстан университеті, Петропавл, Қазақстан

²Астана халықаралық университеті, Астана, Қазақстан

*e-mail: mzhakypzhan@mail.ru

Г.С. Тілеубергенова – биология ғылымдарының кандидаты, биология кафедрасының доценті, М. Қозыбаев атындағы Солтүстік Қазақстан университеті, Петропавл, Қазақстан, e-mail: gbattalova@ku.edu.kz, <https://orcid.org/0000-0002-1001-9528>

М.Ж. Шарипова – PhD, Астана халықаралық университетінің қауымдастырылған профессоры, Астана, Қазақстан, e-mail: mzhakypzhan@mail.ru, <https://orcid.org/0000-0002-9467-1483>

Е.В. Галактионова – аға оқытушы, биология магистрі, М. Қозыбаев атындағы Солтүстік Қазақстан университеті, Петропавл, Қазақстан, e-mail: evgalaktionova@ku.edu.kz, <https://orcid.org/0009-0007-9110-9747>

Аңдатпа. Қазақстанның өңірлік университеттеріндегі гербарий коллекциялары деректердің цифрлық форматының болмауына байланысты халықаралық ғылыми айналымнан тыс қалып келеді. Бұл жұмыс осы олқылықты М. Қозыбаев атындағы Солтүстік Қазақстан университетінің "Биология" кафедрасының гербарий қорына қатысты жоюға бағытталған. Зерттеу материалы ретінде 1957–2025 жылдар аралығында жиналған тоғыз тұқымдасқа – *Boraginaceae*, *Labiatae*, *Solanaceae*, *Scrophulariaceae*, *Orchidaceae*, *Liliaceae*, *Ericaceae*, *Convallariaceae*, *Fabaceae* және *Asteraceae* – жататын 925 гербарий үлгісі пайдаланылды. Коллекцияға өткен ғасырдың ортасындағы тарихи үлгілер де, соңғы далалық зерттеулердің материалдары да кіреді. Цифрландыру жұмысы Darwin Core (DwC) халықаралық стандартына сәйкес жүргізілді. Деректердің дұрыстығы GBIF валидаторы арқылы тексерілді, содан кейін

жиынтық IPT платформасында орналастырылып, GBIF жаһандық порталында жарияланды, бұл оның ашықтығын және халықаралық биоалуантүрлілік деректері инфрақұрылымына кірігуін қамтамасыз етті. Таксономиялық талдау нәтижесінде орман, шалғын және шалғын-дала қауымдастықтарына тән түрлердің басым екені анықталды, бұл өңірдің өсімдік жамылғысының ерекшеліктеріне сәйкес келеді. Коллекцияның уақыттық қамтуы 70 жылдан астам кезеңді қамтиды, бұл флористикалық жинау динамикасын ұзақ мерзім бойы қадағалауға мүмкіндік береді. Географиялық талдау әкімшілік аудандар бойынша табылымдардың біркелкі бөлінбегенін көрсетті, ең көп үлгі Солтүстік Қазақстан облысының Қызылжар ауданынан алынған. Алынған цифрлық деректер жиынтығы флористикалық, экологиялық-географиялық және таксономиялық зерттеулер үшін құнды болып табылады, гербарий қорын сақтауға ықпал етеді және оны білім беру мен табиғат қорғау саласында пайдалану мүмкіндіктерін кеңейтеді.

Кілт сөздер: флора, гербарий жинақтары; сандық мәліметтер базасы; цифрландыру; GBIF; Darwin Core; IPT; геореференция.

Introduction. Herbarium collections are far more than static archives of pressed plants — they represent a tangible record of biological diversity across time. A herbarium sheet collected decades ago can reveal how vegetation looked in a given locality, allow researchers to trace shifts in species distribution, document taxa that have since become rare or vanished altogether, and provide the raw material on which systematic and monitoring studies depend. Yet for a long time, many regional collections have remained largely invisible to the wider scientific community simply because they existed only on paper.

Over the past decade, Kazakhstan has seen substantial progress in digitizing its herbarium holdings. The country's leading collection belongs to the Institute of Botany and Phytointroduction, registered internationally under the index AA. It is not only the largest herbarium in the Republic but ranks among the largest in the former USSR, having been established in 1946 and now holding over 250,000 specimens. Since 2022, the Institute has been building an electronic database of its material in Excel, organizing specimens hierarchically — from family down to species, and from floristic district down to individual collection point (Osmonali B. B. et al., 2022: 521–528).

A younger but rapidly growing collection is that of the Astana Botanical Garden, founded in 2019 as part of the laboratory of flora and plant resources. By 2022, it already comprised 13,492 sheets — 10,733 of them duplicates from other herbaria and private collections, and 2,759 gathered during expeditions across Central and Northern Kazakhstan between 2019 and 2022. The collection was entered into the international Index Herbariorum database in 2021, receiving the acronym NUR (Kubentaev S. A., Mukhtubaeva S. K., 2022: 354–357).

Researchers at Astana International University have also been active in this area, producing three taxonomic checklists of Kazakhstan's vascular flora in 2026; one of these lists alone covers 5,833 species across 978 genera and 128 families (Kubentaev S. A. et al., 2026a, 2026b; Alibekov et al., 2026).

The herbarium of North Kazakhstan University has a somewhat longer history, dating back to 1960. Early records indicate that the regional flora included 881 plant species — mosses, higher vascular plants, and flowering plants alike (Galiyev Zh. M., Kalkamanova A. B., 2006: 58 p). More recent surveys have revised this figure upward (Tleubergenova G. S., Kuznetsova M. A., 2017), and the species list continues to grow: among the notable additions, *Pleurospermum uralense* Hoffm. was recorded as a new genus and species for the Tobol-Ishim botanical-geographical district (Tleubergenova, 2023: 34–40). The collection is estimated to hold around 8,000 specimens. Several targeted studies have already been carried out on its material, including an inventory and taxonomic analysis of the subclass Lamiidae (Romanchuk V. V., Tleubergenova G. S., 2023: 24–34; Romanchuk V. V., 2024: 1–7) and a separate analysis of the family Asteraceae (Shakeeva Zh. E., 2018: 1–7).

Against this backdrop, the digitization of herbarium collections has taken on new significance in an era of rapidly advancing digital technology. The process typically involves three stages: systematizing the data, standardizing it according to international formats, and publishing it through open global platforms such as GBIF (Global Biodiversity Information Facility). Moving from

paper-based archives to digital resources does more than simply modernize storage — it broadens the scientific reach of the material, improves accessibility, and links regional biodiversity data into global analytical frameworks.

In 2024, the M. Kozybayev NKU Herbarium reached a milestone of its own: for the first time in its history, it was registered in Index Herbariorum, becoming the tenth Kazakhstani organization listed with GBIF and the fourth in the country to publish a dataset there — in this case, 400 digitized specimens. The herbarium now carries the international index SQU (Romanchuk V. V., 2024: 17–27).

Digitizing the herbarium held by the Department of Biology at SKU carries particular weight, since it captures the floristic diversity of the North Kazakhstan region across a 70-year span, from 1957 to 2026. Building a digital dataset, verifying and standardizing it, and publishing it on GBIF together serve three purposes at once: preserving the underlying information, ensuring scientific transparency, and creating lasting value for biological, ecological, and conservation research.

This view is consistent with the broader literature. Herbarium collections are now widely regarded as among the most valuable sources for understanding biodiversity at both global and regional scales — they offer a historical lens on species distribution and make it possible to study range shifts, phenological change, and the effects of human impact over time (Heberling J. M. et al., 2019: 381–387). Digitization amplifies this value considerably, turning static local collections into resources that are accessible worldwide and drawing them into the broader international scientific community (Meineke E. K. et al., 2018: 505–525). Publishing herbarium data through GBIF has additional benefits: it tends to increase citation of the underlying collections, draws in a wider circle of researchers, and allows regional information to feed into large-scale biodiversity conservation projects (Nelson, Ellis, 2019). Indeed, recent reviews identify data mobilization as one of the central priorities in the ongoing development of digital herbaria [16] — a priority that matters especially for university collections, which have historically been confined to local, in-house use.

Taken together, these developments place the digitization and publication of the SKU Department of Biology's herbarium material squarely within current global trends in natural history collections, while giving the flora of the North Kazakhstan region a place in the global information space it previously lacked.

This paper presents that work — the digitization, systematization, and publication of the collection — along with an assessment of its taxonomic composition and its temporal and geographic structure.

Materials and Methods. Materials and Methods. The study drew on 950 herbarium specimens collected across the North Kazakhstan Region. This material combines two sources: specimens already held in the herbarium collection of the Department of Biology, and fresh field collections carried out by the authors in 2025.

Recent fieldwork took place in and around the city of Petropavlovsk and the surrounding territories of the North Kazakhstan Region. Specimens were gathered during flowering and fruiting, since these stages allow the most reliable identification. For every specimen, we recorded the family, genus, species, exact location, habitat, collection date, and the names of the collectors. Once collected, plants were pressed and dried in a standard herbarium press, then mounted on sheets and labeled.

Digitization followed the established guidelines for dataset preparation prior to publication (Buivolov Yu. A., 2019). Label data were transferred into an electronic spreadsheet formatted according to Darwin Core (DwC) standards (Figure 1). The spreadsheet included fields such as occurrenceID, family, scientificName, organismQuantity, organismQuantityType, lifeStage, verbatimEventDate, eventDate, year, month, day, countryCode, country, stateProvince, county, locality, verbatimLocality, habitat, language, recordedBy, identifiedBy, dateIdentified, decimalLatitude, decimalLongitude, coordinateUncertaintyInMeters, coordinatePrecision, geodeticDatum, georeferencedBy, georeferencedDate, vernacularName, taxonRank, kingdom, disposition, occurrenceStatus, collectionCode, institutionID, basisOfRecord, and continent (Buivolov Yu. A., 2019).

Once the table was complete, we ran it through the built-in GBIF data validator, which checks a range of parameters automatically — the correctness of Latin names, the uniqueness of occurrenceID

values, whether coordinates fall within plausible bounds for the region, date and text formatting, and any structural inconsistencies in the file. The validator flagged a handful of issues, mostly involving date formats and a few coordinate entries, which we corrected by hand. After these fixes, the validator returned a clean "no errors" result, clearing the way for the next stage (Meineke E. K. et al., 2018: 505–525).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	occurrenceID	family	scientificName	organismQuantity	organismQuantityType	lifeStage	verbatimEventDate	eventDate	year	month	day	countryCode	country	stateProvince	county	locality
2	SQU00401	Asteraceae	<i>Senecio jacobaea</i> L.		Druid scale	flowering	08.07.2003r	2003-07-08	2003	7	8	KZ	Kazakhstan	North Kazakhstan Region	Kyzyltar district	Petrozavodsk village
3	SQU00402	Asteraceae	<i>Senecio jacobaea</i> L.		Druid scale	flowering	08.07.95r	1995-07-08	1995	7	8	KZ	Kazakhstan	North Kazakhstan Region	Kyzyltar district	Petrozavodsk village
4	SQU00403	Asteraceae	<i>Senecio jacobaea</i> L.		Druid scale	flowering	12.07.72r	1972-07-12	1972	7	12	KZ	Kazakhstan	North Kazakhstan Region	Kyzyltar district	Krasnaya village
5	SQU00404	Asteraceae	<i>Lepidosiphon crassipes</i> (L.) Dub.		Druid scale	flowering	08.07.95r	1995-07-08	1995	8	8	KZ	Kazakhstan	North Kazakhstan Region	Kyzyltar district	Petrozavodsk village
6	SQU00405	Asteraceae	<i>Antennaria dioica</i> L.		Druid scale	flowering	20.06.71r	1971-06-20	1971	6	20	KZ	Kazakhstan	North Kazakhstan Region	Kyzyltar district	Krasnaya village
7	SQU00406	Asteraceae	<i>Antennaria dioica</i> L.		Druid scale	flowering	1974r	1974-06	1974	6		KZ	Kazakhstan	North Kazakhstan Region	Kyzyltar district	Krasnaya village
8	SQU00407	Asteraceae	<i>Antennaria dioica</i> L.		Druid scale	flowering	1971r	1971-06	1971	6		KZ	Kazakhstan	North Kazakhstan Region	Kyzyltar district	Serolova village
9	SQU00408	Asteraceae	<i>Antennaria dioica</i> L.		Druid scale	flowering	20.07.70r	1970-07-20	1970	7	20	KZ	Kazakhstan	North Kazakhstan Region	Kyzyltar district	Serolova village
10	SQU00409	Asteraceae	<i>Edmonia rufa</i> L.		Druid scale	flowering	04.07.05r	2005-07-04	2005	7	4	KZ	Kazakhstan	North Kazakhstan Region	Kyzyltar district	Barovka village

Figure 1. Fragment of a completed Darwin Core (DwC) format table

Digitizing natural history collections is rarely just about data entry — it typically involves producing high-quality images of the specimens, transcribing label information, georeferencing localities, standardizing metadata, and ultimately publishing everything in an internationally recognized format, most commonly Darwin Core (Groom Q. et al., 2017). Standardized data of this kind stay usable over time and can be folded into global repositories such as GBIF without further conversion (Nelson, Ellis, 2019).

Metadata were entered through the IPT interface, covering a description of the dataset itself, the goals of the research, and why digitizing this particular herbarium matters. Well-constructed metadata do more than tidy up the record — they establish clear provenance, add to the scientific credibility of the published dataset, and make the material more likely to be cited by other researchers (Borsch T. et al., 2020: 819–832).

Once both the data table and the metadata had been reviewed, we published the dataset through the GBIF platform. IPT generated a Darwin Core Archive automatically and assigned the dataset a unique DOI, after which the records were indexed and made accessible to researchers worldwide (Figure 2).

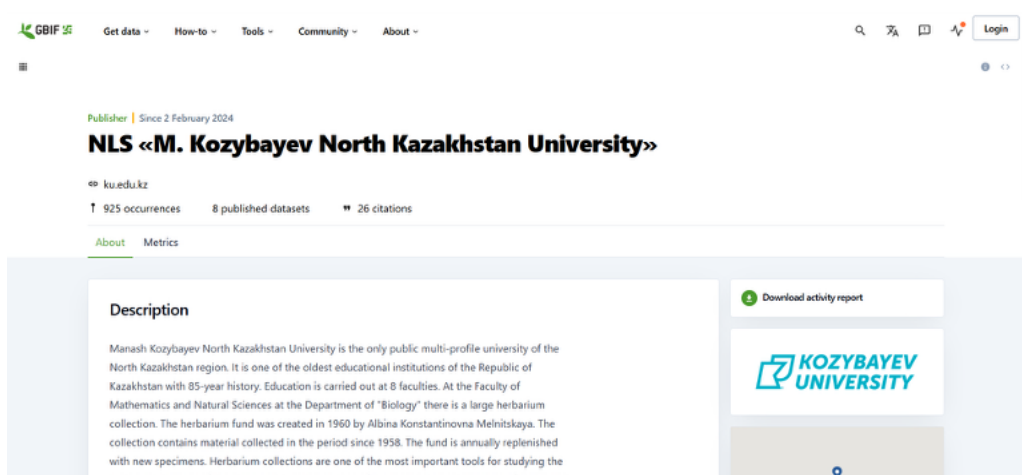


Figure 2. Main page of M.Kozybaev North Kazakhstan University on the GBIF platform

All records now display correctly in the GBIF occurrence catalog, where users can filter results by species, coordinates, collection year, and habitat type. The dataset is freely available to researchers, ecologists, botanists, and anyone working in biodiversity science.

The digitization effort itself was carried out by staff of the Department of Biology together with master's students, as part of their broader involvement in the department's research projects.

Research results. Once the digitized collections from several families of the department's herbarium were published on GBIF, we turned to the "Metrics" section the platform generates automatically and analyzed it in detail.

At present, eight datasets are hosted on GBIF, together covering metadata for 925 herbarium specimens spanning a range of taxonomic groups of flowering plants from the North Kazakhstan Region. Each dataset records the taxonomic identity of the specimens, where and when they were collected, who collected them, and their broader geographic origin. Most of the underlying material comes from the North Kazakhstan Region itself, though a smaller portion originates from other parts of Kazakhstan.

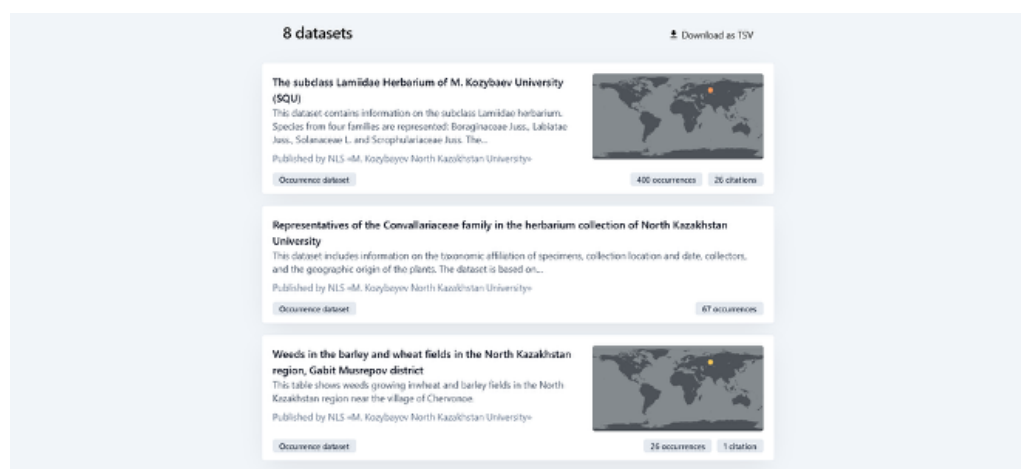


Figure 3. Page 8 of the published datasets from M. Kozybaev North Kazakhstan University on the GBIF platform

The datasets were published over a period stretching from May 1, 2024, to March 15, 2026. In roughly two years of digitization work, this yielded eight published datasets — seven of them covering herbaria across ten plant families, and one dedicated specifically to weed species (Table 1).

Table 1. Herbarium collections of North Kazakhstan University on the GBIF platform

No.	Collection name	Number of samples	Date posted	Number of citations	Identifier
1	Subclass <i>Lamiidae</i> Herbarium of M. Kozybayev University (<i>Boraginaceae</i> Juss., <i>Labiatae</i> Juss., <i>Solanaceae</i> Juss., <i>Scrophulariaceae</i> Juss.)	400	01.05.2024	26	DOI 10.15468/ubzs5t
2	Herbarium of the <i>Fabaceae</i> Family of Manash Kozybayev North Kazakhstan University	200	20.10.2025	1	DOI 10.15468/j95r2t
3	Herbarium Collection of <i>Asteraceae</i> (<i>Compositae</i>) of Manash Kozybayev North Kazakhstan University	200	20.10.2025	1	DOI 10.15468/hrb39k
4	Herbarium Collection of Orchids (Family <i>Orchidaceae</i>) of Manash Kozybayev North Kazakhstan University	17	05.11.2025	1	DOI 10.15468/quwnan
5	Herbarium Collection of <i>Liliaceae</i> (Family <i>Liliaceae</i>) of Manash Kozybayev North Kazakhstan University (NKU)	5	05.11.2025	—	DOI 10.15468/gq23d8
6	Herbarium Collection of <i>Ericaceae</i> from Manash Kozybayev North Kazakhstan University	10	07.11.2025	1	DOI 10.15468/ctmk23
7	Weeds in Barley and Wheat Fields of North Kazakhstan Region, Gabit Musrepov District	26	13.11.2025	1	DOI 10.15468/peshvt

No.	Collection name	Number of samples	Date posted	Number of citations	Identifier
8	Representatives of the <i>Convallariaceae</i> Family in the Herbarium Collection of North Kazakhstan University	67	15.03.2026	1	DOI 10.15468/f9r8qq

Of the full set, 924 specimens were identified down to species level. Coordinates are available for 829 records, collection years for 914, and collector names for 913 — though a number of older labels lack one or both of the latter two details (Table 2). Taken together, the 925 records span collection dates from 1957 through 2025.

Table 2. Frequency indicators

Data richness	Number of cases
Identification to species	924
With coordinates	829
With year	914
With collector	913
With media	--
Sequenced	--

Roughly 5.8% of the records correspond to species listed on the IUCN Global Red List, drawn mainly from the families Orchidaceae, Liliaceae, and Ericaceae.

For the taxonomic composition, we opted to list every genus within each family alongside its specimen count, rather than relying on charts alone — this gives a more precise picture of how specimens are distributed across taxa than a graphic summary would.

The temporal spread of specimens tells its own story: it reflects not just how the collection grew over time, but how field activity varied from year to year. Looking at the yearly breakdown makes it possible to see how consistently collecting was carried out, to spot years of particularly active fieldwork, and conversely, to identify stretches where almost nothing was added.

Geographically, the specimens are distributed quite unevenly — a pattern shaped both by how accessible different habitats actually were and by the routes field researchers happened to take. In total, the collection draws from seven administrative districts of the North Kazakhstan Region, though collection points cluster far more densely in some areas than others (Figure 4).

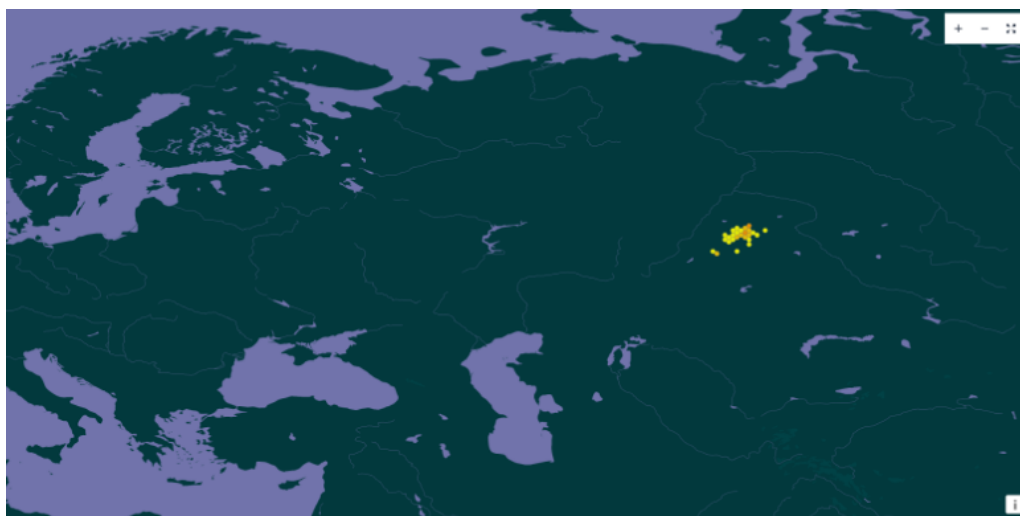


Figure 4. Map of georeferenced herbarium records (collection points)

The dataset's main GBIF page also tracks citation activity. So far, it has been cited in 26 sources: journal articles account for the largest share at 61.5%, followed by preprints at 19.2%, conference materials at 7.7%, reports at 3.8%, and abstracts at another 3.8% (Figure 5).

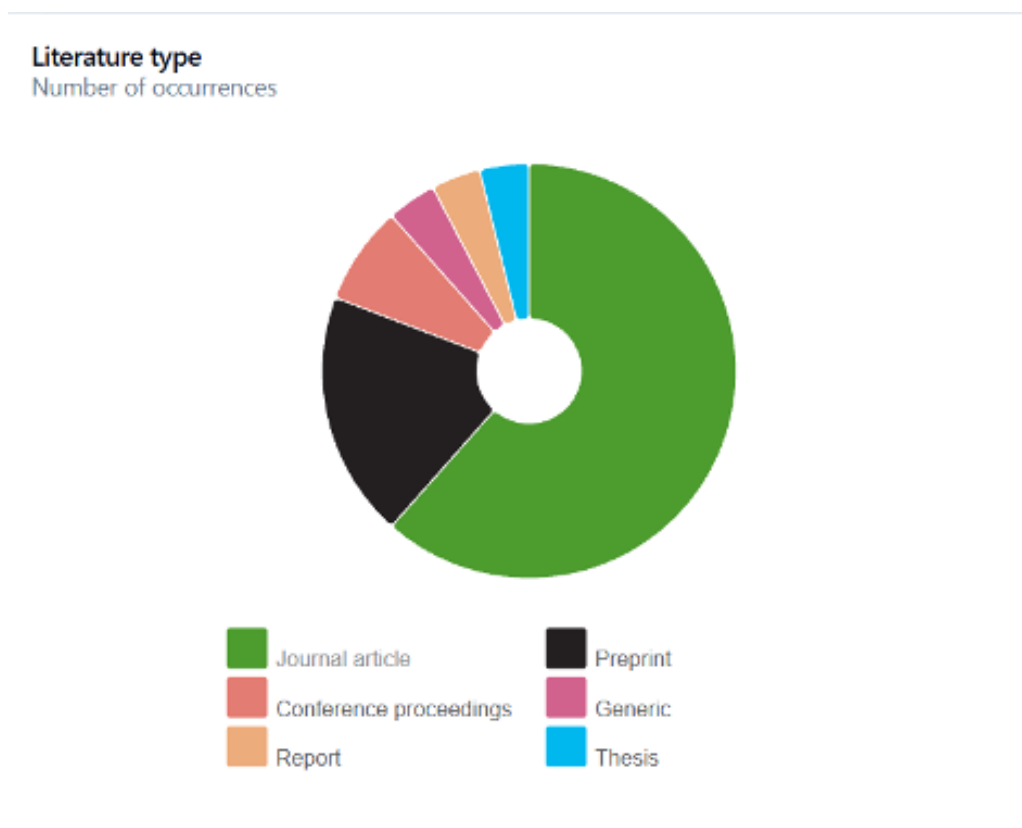


Figure 5. Citation metrics of the published dataset

Conclusion. This study covered the full cycle of work with herbarium material from ten families — Boraginaceae, Labiatae, Solanaceae, Scrophulariaceae, Orchidaceae, Liliaceae, Ericaceae, Convallariaceae, Fabaceae, and Asteraceae — held by the Department of Biology at M. Kozybayev North Kazakhstan University: collecting, organizing, digitizing, and finally publishing the material. The result is a dataset of 925 carefully prepared specimens, now feeding into both regional and global biodiversity information systems.

Publishing this material on GBIF changed its status considerably. Data that had previously sat in local storage became openly accessible, entered the international scientific community, and can now be reused by researchers anywhere in the world. Beyond simple accessibility, digitization plays a protective role as well — it helps safeguard specimens that are otherwise vulnerable to physical loss or degradation over time, while at the same time opening up new avenues for biodiversity analysis, flora monitoring, and conservation planning grounded in real data.

Taken together, this work adds a meaningful piece to the broader picture of North Kazakhstan's flora. It also lays groundwork for what comes next — further botanical research, an expanded collection, and the continued development of digital herbaria at the local level.

Литература

Alibekov D., Kubentayev S., Nobis M., Ishmuratova M., Chepinoga V., Sennikov A. (2026). Checklist of Alien Vascular Plants of Kazakhstan. Astana International University (AIU). Checklist dataset .[https://doi.org/ 10.15468/tz64jk](https://doi.org/10.15468/tz64jk) accessed via GBIF.org on 2026-06-24.

Буйволов Ю. А., Иванова Н. В., Шашков М. П. Оцифровка данных Летописей природы и научных биологических коллекций особо охраняемых природных территорий: учебное пособие. – ФГБУ «Приокско-Террасный государственный природный биосферный заповедник», 2019.

Borsch T., et al. The herbarium of the future: advancing digitization and data integration // Trends in Ecology & Evolution. – 2020. – Vol. 35(10). – P. 819–832. DOI: 10.1016/j.tree.2020.06.010

- Галиев Ж. М., Калкаманова А. Б. Гербарный фонд кафедры общей биологии Северо-Казахстанского государственного университета имени М. Козыбаева: учебно-методическое пособие для студентов по ботаническим дисциплинам. – Петропавловск: СКГУ им. М. Козыбаева, 2006. – 58 с.
- Heberling J.M., Prather L.A., Tonsor S.J. Herbarium specimens as exaptations: New uses for old collections // *American Journal of Botany*. – 2019. – Vol. 106(3). – P. 381–387. DOI: 10.1002/ajb2.1225
- Groom Q., et al. The importance of publishing data papers for biodiversity datasets // *Biodiversity Data Journal*. – 2017. – Vol. 5. – e20519. DOI: 10.3897/BDJ.5.e20519
- Hardisty A., Roberts D. A decadal view of biodiversity informatics: challenges and priorities // *BMC Ecology*. – 2013. – Vol. 13. – 16. DOI: 10.1186/1472-6785-13-16
- Краткое руководство по подготовке наборов данных для публикации через Глобальный информационный фонд по биоразнообразию (GBIF) / сост. О. И. Бородин. Минск: Экоперспектива, 2022. – 64 с. ISBN 978-985-469-896-0.
- Kuandiykova M. (2025). Herbarium of the Fabaceae family North Kazakhstan University named after Manash Kozybayev (SQU). NLS «M. Kozybayev North Kazakhstan University». Occurrence dataset <https://doi.org/10.15468/j95r2t> accessed via GBIF.org on 2026-06-24.
- Кубентаев С. А., Мухтубаева С. К. Формирование гербарного фонда Астанинского ботанического сада // *Материалы международной научно-практической конференции «Изучение, сохранение и рациональное использование растительного мира Евразии»*. – Алматы, 2022. – С. 354–357.
- Kubentayev S, Alibekov D, Baasanmunkh S, Nobis M, Myrzagaliyeva A, Tustubayeva S, Kubentayeva B, Idrissova Z, Munkhtulga D, Tsegmed Z, Choi H J, Sennikov A (2026). Checklist of Native Vascular Plants of Kazakhstan. Astana International University (AIU). Checklist dataset <https://doi.org/10.15468/m75hmd> accessed via GBIF.org on 2026-06-24.
- Lyubchic V (2025). Weeds in the barley and wheat fields in the North Kazakhstan region, Gabit Musrepov district. NLS «M. Kozybayev North Kazakhstan University». Occurrence dataset <https://doi.org/10.15468/peshvt> accessed via GBIF.org on 2026-06-24.
- Meineke E.K., Davis C.C., Davies T.J. The unrealized potential of herbaria for global change biology // *Ecological Monographs*. – 2018. – Vol. 88(4). – P. 505–525. DOI: 10.1002/ecm.1307
- Nelson G., Ellis S. The history and impact of digitization and digital data mobilization on biodiversity research // *Philosophical Transactions of the Royal Society B*. – 2019. – Vol. 374. – 20170391. DOI: 10.1098/rstb.2017.0391
- Осмонали Б. Б., Веселова П. В., Кудабаева Г. М., Данилов М. П. Формирование онлайн электронной базы гербарных образцов гербарного фонда (АА) Института ботаники и фитоинтродукции // *Материалы международной научно-практической конференции «Изучение, сохранение и рациональное использование растительного мира Евразии»*. – Алматы, 2022. – С. 521–528.
- Robertson T., et al. The GBIF Integrated Publishing Toolkit: facilitating the efficient publishing of biodiversity data on the internet // *PLoS ONE*. – 2014. – Vol. 9(8). – e102623. DOI: 10.1371/journal.pone.0102623
- Романчук В. В. Оцифровка и индексация гербария растений подкласса Lamiidae гербарного фонда кафедры «Биология» СКУ им. М. Козыбаева: автореф. магистерской диссертации. 2024. – 7 с.
- Romanchuk V. (2024). The subclass Lamiidae Herbarium of M. Kozybaev University (SQU). NLS «M. Kozybayev North Kazakhstan University». Occurrence dataset <https://doi.org/10.15468/ubzs5t> accessed via GBIF.org on 2026-06-24.
- Романчук В. В., Мадиева А. Н., Тлеубергенова Г. С. Состояние вопроса создания электронного гербария в Казахстане // *Вестник Северо-Казахстанского университета им. М. Козыбаева*. – 2024. – № 1 (61). – С. 17–27.
- Романчук В. В., Тлеубергенова Г. С. Таксономический состав подкласса Lamiidae в гербарной коллекции кафедры «Биология» СКУ им. М. Козыбаева // *Вестник Северо-Казахстанского университета им. М. Козыбаева*. – 2023. – № 4 (60). – С. 24–34. <https://doi.org/10.54596/2958-0048-2023-4-24-34>
- Skakunova P. Herbarium Collection of Asteraceae (Compositae) from North Kazakhstan University named after Manash Kozybayev (SQU). Version 1.2. NLS «M. Kozybayev North Kazakhstan University». Occurrence dataset. <https://doi.org/10.15468/hrb39k> (accessed 17.11.2025).
- Tleubergenova G. (2026). Representatives of the Convallariaceae family in the herbarium collection of North Kazakhstan University. NLS «M. Kozybayev North Kazakhstan University». Occurrence dataset: <https://doi.org/10.15468/f9r8qq> accessed via GBIF.org on 2026-06-24.
- Tleubergenova G, Galaktionova E, Skakunova P (2025). Herbarium Collection of Orchidaceae (Orchid family) from North Kazakhstan University named after Manash Kozybayev (SQU). Version 1.2. NLS «M. Kozybayev North Kazakhstan University». Occurrence dataset <https://doi.org/10.15468/quwnan> accessed via GBIF.org on 2026-06-24.
- Tleubergenova G, Galaktionova E, Skakunova P (2025). Herbarium Collection of Liliaceae (Lily family) from North Kazakhstan University named after Manash Kozybayev (SQU). NLS «M. Kozybayev North Kazakhstan University». Occurrence dataset <https://doi.org/10.15468/gq23d8> accessed via GBIF.org on 2026-06-24.
- Tleubergenova G, Galaktionova E, Skakunova P (2025). Herbarium Collection of Ericaceae from North Kazakhstan University named after Manash Kozybayev (SQU). Version 1.3. NLS «M. Kozybayev North Kazakhstan University». Occurrence dataset <https://doi.org/10.15468/ctmk23> accessed via GBIF.org on 2026-06-24.
- Тлеубергенова Г. С., Кузнецова М. А. Флора и растительность Северо-Казахстанской области. – Петропавловск: СКГУ им. М. Козыбаева. – 2017. – 150 с.
- Tleubergenova G.S., Safonov A.I., Galaktionova E.V. Floristic finds in the Northern Priishimye, Kazakhstan // *Вестник Северо-Казахстанского университета им. М. Козыбаева*. – 2024. – № 3 (63). – С. 34–40.
- Шакеева Ж. Е. Биоразнообразие представителей семейства Астровых во флоре СКО: автореф. магистерской диссертации. – Петропавловск: СКГУ им. М. Козыбаева. – 2018. – 7 с.

References

- Alibekov, D., Kubentayev, S., Nobis, M., Ishmuratova, M., Chepinoga, V., Sennikov, A. (2026). Checklist of Alien Vascular Plants of Kazakhstan. Astana International University (AIU). Checklist dataset. <https://doi.org/10.15468/tz64jk> accessed via GBIF.org on 2026-06-24. [in Eng.]
- Bujvolov, Yu. A., Ivanova, N. V., Shashkov, M. P. (2019). *Ocifrovka dannyh Letopisej prirody i nauchnyh biologicheskikh kolekcij osobno ohranyaemykh prirodnykh territorij: uchebnoe posobie*. – FGBU «Prioksko-Terrasnyj gosudarstvennyj prirodnyj biosfernyj zapovednik», 2019. [in Russ.]
- Borsch, T., et al. (2020). The herbarium of the future: advancing digitization and data integration // *Trends in Ecology & Evolution*. – 2020. – Vol. 35(10). – P. 819–832. DOI: 10.1016/j.tree.2020.06.010. [in Eng.]
- Galiev, Zh. M., Kalkamanova, A. B. (2006). *Gerbarnyj fond kafedry obshchej biologii Severo-Kazhastanskogo gosudarstvennogo*

- universiteta imeni M. Kozybaeva: uchebno-metodicheskoe posobie dlya studentov po botanicheskim disciplinam. – Petropavlovsk: SKGU im. M. Kozybaeva, 2006. – 58 s. [in Russ.]
- Heberling, J.M., Prather, L.A., Tonsor, S.J. (2019). Herbarium specimens as exaptations: New uses for old collections // *American Journal of Botany*. – 2019. – Vol. 106(3). – P. 381–387. DOI: 10.1002/ajb2.1225. [in Eng.]
- Groom, Q., et al. (2017). The importance of publishing data papers for biodiversity datasets // *Biodiversity Data Journal*. – 2017. – Vol. 5. – e20519. DOI: 10.3897/BDJ.5.e20519. [in Eng.]
- Hardisty, A., Roberts D. (2013). A decadal view of biodiversity informatics: challenges and priorities // *BMC Ecology*. – 2013. – Vol. 13. – 16. DOI: 10.1186/1472-6785-13-16. [in Eng.]
- Borodin, O. I. (2022). Kratkoe rukovodstvo po podgotovke naborov dannyh dlya publikacii cherez Global'nyj informacionnyj fond po bioraznoobraziyu (GBIF) / sost. O. I. Borodin. Minsk: Ekoperspektiva, 2022. – 64 s. ISBN 978-985-469-896-0. [in Russ.]
- Kuandykova, M. (2025). Herbarium of the Fabaceae family North Kazakhstan University named after Manash Kozybayev (SQU). NLS «M. Kozybayev North Kazakhstan University». Occurrence dataset <https://doi.org/10.15468/j95r2t> accessed via GBIF.org on 2026-06-24. [in Eng.]
- Kubentaev, S. A., Muhtubaeva, S. K. (2022). Formirovanie gerbarnogo fonda Astaninskogo botanicheskogo sada // *Materialy mezhdunarodnoj nauchno-prakticheskoy konferencii «Izuchenie, sohranenie i racional'noe ispol'zovanie rastitel'nogo mira Evrazii»*. – Almaty, 2022. – S. 354–357. [in Russ.]
- Kubentayev, S., Alibekov, D., Baasanmunkh, S., Nobis, M., Myrzagaliyeva, A., Tustubayeva, S., Kubentayeva, B., Idrissova, Z., Munkhtulga, D., Tsegmed, Z., Choi, H. J., Sennikov, A. (2026). Checklist of Native Vascular Plants of Kazakhstan. Astana International University (AIU). Checklist dataset <https://doi.org/10.15468/m75hmd> accessed via GBIF.org on 2026-06-24. [in Eng.]
- Lyubchic, V. (2025). Weeds in the barley and wheat fields in the North Kazakhstan region, Gabit Musrepov district. NLS «M. Kozybayev North Kazakhstan University». Occurrence dataset <https://doi.org/10.15468/peshvt> accessed via GBIF.org on 2026-06-24. [in Eng.]
- Meineke, E.K., Davis, C.C., Davies, T.J. (2018). The unrealized potential of herbaria for global change biology // *Ecological Monographs*. – 2018. – Vol. 88(4). – P. 505–525. DOI: 10.1002/ecm.1307. [in Eng.]
- Nelson, G., Ellis, S. (2019). The history and impact of digitization and digital data mobilization on biodiversity research // *Philosophical Transactions of the Royal Society B*. – 2019. – Vol. 374. – 20170391. DOI: 10.1098/rstb.2017.0391. [in Eng.]
- Osmonali, B. B., Veselova, P. V., Kudabaeva, G. M., Danilov, M. P. (2022). Formirovanie onlajn elektronnoy bazy gerbarnykh obrazcov gerbarnogo fonda (AA) Instituta botaniki i fitointrodukcii // *Materialy mezhdunarodnoj nauchno-prakticheskoy konferencii «Izuchenie, sohranenie i racional'noe ispol'zovanie rastitel'nogo mira Evrazii»*. – Almaty, 2022. – S. 521–528. [in Russ.]
- Robertson, T., et al. (2014). The GBIF Integrated Publishing Toolkit: facilitating the efficient publishing of biodiversity data on the internet // *PLoS ONE*. – 2014. – Vol. 9(8). – e102623. DOI: 10.1371/journal.pone.0102623. [in Eng.]
- Romanchuk, V. V. (2024). Ocifrovka i indeksaciya gerbariya rastenij podklasa Lamiidae gerbarnogo fonda kafedry «Biologiya» SKU im. M. Kozybaeva: avtoref. masterskoj dissertacii. 2024. – 7 s. [in Russ.]
- Romanchuk, V. (2024). The subclass Lamiidae Herbarium of M. Kozybaev University (SQU). NLS «M. Kozybayev North Kazakhstan University». Occurrence dataset <https://doi.org/10.15468/ubzs5t> accessed via GBIF.org on 2026-06-24. [in Eng.]
- Romanchuk, V. V., Madieva, A. N., Tleubergenova, G. S. (2024). Sostoyanie voprosa sozdaniya elektronnoy gerbariya v Kazahstane // *Vestnik Severo-Kazhastanskogo universiteta im. M. Kozybaeva*. – 2024. – № 1 (61). – S. 17–27. [in Russ.]
- Romanchuk, V. V., Tleubergenova, G. S. (2023). Taksonomicheskij sostav podklasa Lamiidae v gerbarnoj kollekcii kafedry «Biologiya» SKU im. M. Kozybaeva // *Vestnik Severo-Kazhastanskogo universiteta im. M. Kozybaeva*. – 2023. – № 4 (60). – S. 24–34. <https://doi.org/10.54596/2958-0048-2023-4-24-34>. [in Russ.]
- Skakunova, P. (2025). Herbarium Collection of Asteraceae (Compositae) from North Kazakhstan University named after Manash Kozybayev (SQU). Version 1.2. NLS «M. Kozybayev North Kazakhstan University». Occurrence dataset. <https://doi.org/10.15468/hrb39k> (accessed 17.11.2025). [in Eng.]
- Tleubergenova, G. (2026). Representatives of the Convallariaceae family in the herbarium collection of North Kazakhstan University. NLS «M. Kozybayev North Kazakhstan University». Occurrence dataset: <https://doi.org/10.15468/f9r8qq> accessed via GBIF.org on 2026-06-24. [in Eng.]
- Tleubergenova, G., Galaktionova, E., Skakunova, P. (2025). Herbarium Collection of Orchidaceae (Orchid family) from North Kazakhstan University named after Manash Kozybayev (SQU). Version 1.2. NLS «M. Kozybayev North Kazakhstan University». Occurrence dataset <https://doi.org/10.15468/quwnan> accessed via GBIF.org on 2026-06-24. [in Eng.]
- Tleubergenova, G., Galaktionova, E., Skakunova, P. (2025). Herbarium Collection of Liliaceae (Lily family) from North Kazakhstan University named after Manash Kozybayev (SQU). NLS «M. Kozybayev North Kazakhstan University». Occurrence dataset <https://doi.org/10.15468/gq23d8> accessed via GBIF.org on 2026-06-24. [in Eng.]
- Tleubergenova, G., Galaktionova, E., Skakunova, P. (2025). Herbarium Collection of Ericaceae from North Kazakhstan University named after Manash Kozybayev (SQU). Version 1.3. NLS «M. Kozybayev North Kazakhstan University». Occurrence dataset <https://doi.org/10.15468/ctmk23> accessed via GBIF.org on 2026-06-24. [in Eng.]
- Tleubergenova, G. S., Kuznecova, M. A. (2017). Flora i rastitel'nost' Severo-Kazhastanskoy oblasti. – Petropavlovsk: SKGU im. M. Kozybaeva. – 2017. – 150 s. [in Russ.]
- Tleubergenova, G.S., Safonov, A.I., Galaktionova, E.V. (2024). Floristic finds in the Northern Priishimye, Kazakhstan // *Vestnik Severo-Kazhastanskogo universiteta im. M. Kozybaeva*. – 2024. – № 3 (63). – C. 34–40. [in Eng.]
- Shakeeva, Zh. E. (2018). Bioraznoobraziye predstavitelej semejstva Astrovykh vo flore SKO: avtoref. masterskoj dissertacii. – Petropavlovsk: SKGU im. M. Kozybaeva. – 2018. – 7 s. [in Russ.]