

... ГЕНЕТИКА ... GENETICS ...

DOI: 10.26565/2075-5457-2024-43-4
УДК: 575.174+314.92

Прізвища українського населення як інструмент популяційно-генетичних досліджень Л.О. Атраментова, В.В. Шуба, М.Ю. Горпинченко

В статті приводяться результати популяційно-генетичного дослідження, виконаного за допомогою прізвищ. Доведено, що українські прізвища відповідають попереднім вимогам, що висуваються до них, як до інструменту популяційно-генетичного дослідження: вони стабільно успадковуються більше десяти поколінь за патрилійним типом; за семантикою і морфологією є найбільш різноманітними у світі; нерівномірно розподіляються по територіях країни, що відрізняються історією, демографічними процесами, антропологічними типами населення, розподілом біологічних маркерів. За коефіцієнтами кореляції між питомою вагою етнічної групи в популяції і числом етнічних прізвищ виявлено прямий зв'язок. Також прямий зв'язок існує між матрицями генетичних відстаней, розрахованих за групами крові, гаплогрупами Y хромосоми і відстанями за прізвищами. Досить сильний зв'язок прізвищ з генетичними маркерами дає підставу для екстраполяції значень, одержаних за допомогою прізвищ, на генофонд населення. Показано розподіл 20 найчастіших прізвищ в різних регіонах. Розраховані показники, які відбивають структуру популяції і генетичні процеси, що в них протікають: індекс ізонімії I_r , показники різноманіття α , інбридингу F_{st} , міграції v , показник упорядкованості H і надмірності розподілу R . Порівняно ці показники між популяціями у географічному і ієрархічному розташуваннях. Індекси ізонімії демонструють чіткий градієнт, що збільшується у напрямку від південного сходу країни до північного заходу. Інші показники також досить чітко змінюють своє значення з південного сходу у західному і північно-західному напрямках. На нижчих рівнях популяційної структури диференційна здатність показників, розрахованих за прізвищами, збільшується. На рівні районних популяцій виявлений ефект засновника, який є аналогом генетичного дрейфу, що особливо чітко демонструє показник індексу місця I_p . Розподіл популяційних показників знаходить пояснення в історичних подіях і демографічних процесах, що мали місце в історії країни.

Ключові слова: українські популяції, прізвища, показники ізонімії, інбридингу, різноманіття, міграції

Цитування: Атраментова Л.О., Шуба В.В., Горпинченко М.Ю. Прізвища українського населення як інструмент популяційно-генетичних досліджень. Вісник Харківського національного університету імені В.Н.Каразіна. Серія «Біологія», 2024, 43, с. 44-49. <https://doi.org/10.26565/2075-5457-2024-43-4>

Про авторів:

Атраментова Л.О. – Харківський національний університет імені В. Н. Каразіна, майдан Свободи, 4, Харків, Україна, 61022, lubov.atramentova@gmail.com, <https://orcid.org/0000-0002-7143-9411>

Шуба В.В. – Харківський національний університет імені В. Н. Каразіна, майдан Свободи, 4, Харків, Україна, 61022; Харківський науково-дослідний експертно-криміналістичний центр МВС України, вулиця Ковтуна, 34, Харків, Україна, 61022, shubavladislav@gmail.com, <https://orcid.org/0000-0002-4261-5228>

Горпинченко М.Ю. – Харківський національний університет імені В. Н. Каразіна, майдан Свободи, 4, Харків, Україна, 61022, 8.gelios.8@gmail.com, <https://orcid.org/0000-0003-0362-6300>

Подано до редакції: 20.10.2023 / Прорецензовано: 10.07.2024 / Прийнято до друку: 19.08.2024

Вступ

Інформація щодо структури людських популяцій і процесів, що в них протікають, відноситься до фундаментальних знань, які використовуються у медико-генетичному консультуванні, при складанні демографічних та епідеміологічних прогнозів тощо. Від структури популяції залежить здоров'я, працездатність населення, його інтелектуальний та культурний потенціал. До недавнього часу територія України була білою плямою на генетико-географічній мапі світу. Генетичні дослідження українських популяцій, започатковані в Харківському національному університеті, заклали підвалини цього напрямку в країні. Продовження досліджень потребує розширення спектра методів. В світовій антропогенетиці при вивченні популяцій вже більш ніж півстоліття використовуються прізвища. Цей зручний фінансово необтяжливий метод дає змогу працювати з великими масивами даних.

Підстави для використання прізвищ як інструменту популяційних досліджень базуються на декількох їх властивостях. Семантика і морфологія прізвища несе відбиток приналежності особи до етносу з його специфічним генофондом; патрилійне успадкування прізвищ нагадує передачу Y хромосоми, матрилійне — передачу мтДНК. Розподіл прізвищ в населенні аналогічний моделі віртуального локусу *прізвище* з множинною системою алельних варіантів, зчепленого з Y хромосомою у народів з патрилійною традицією, або з мтДНК у народів, де приналежність до роду визначається по жіночій лінії. Рідкісні помилки написання або вимови прізвища, як і генні мутації, стабільно успадковуються, а частоти прізвищ аналогічні частотам алелів. Все перелічене за висловленням Н. Мортон за ефективності прізвища як маркерів порівняне з кодомінантною системою.

Прізвища не кожного народу придатні виконувати роль маркерів, для цього вони мають відповідати декільком умовам: стабільно передаватися не менше десяти поколінь, бути достатньо різноманітними, нерівномірно розподілятися по території, маркуючи окремі групи населення, мати зв'язок з біологічними ознаками, бути здатними диференціювати популяції. З'ясувати, чи відповідають прізвища українського населення вимогам, що висуваються до них як інструменту популяційно-генетичних досліджень було метою цього дослідження.

Матеріали та методи

Матеріалом дослідження слугували історичні та етнографічні джерела, дані демографічної статистики, анкети мешканців різних популяцій Розраховано, прізвища населення України станом на 2001 рік. Використано методичні підходи, розроблені I. Barrai (Barrai, 1992).

Частоти прізвищ розраховані як:

$$q_i = \frac{a_i}{N}$$

де a_i — кількість носіїв i -го прізвища; N — кількість особин у популяції.

Показник випадкової ізонімії I_r розрахований за формулою:

$$I_r = \sum q_i^2$$

Випадкова компонента інбридингу F_{ST} визначалася як:

$$F_{st} = \sum q_i^2 / 4$$

Індекс міграції v розрахований за формулою:

$$v = \frac{1 - I_r}{I_r(N - 1)}$$

Показник різноманітності прізвищ α розрахований за формулою:

$$\alpha = \frac{Nv}{1-v}$$

де N — кількість особин у популяції; v — індекс міграції.

Індекс місця I_p одержано як середнє арифметичне рангів двадцяти найчастіших в популяції прізвищ.

Використано метод багатомірного шкалювання. Генетичні відстані між популяціями розраховані за М. Nei (Nei, 1975).

Результати та обговорення

Перші українські прізвища знаходять в грамотах XVIII століття (Редько, 1966). Це свідчить, що вони успадковуються понад 300 років, тобто протягом не менше 12 поколінь, а за семантикою морфологією українські прізвища є найбільш різноманітними у світі (Редько 1966, 1968). В населенні України було налічено 697 147 прізвищ. Серед вивчених країн більше лише в населенні США — 899 585 прізвищ (Barrai, 2001). Показник їх різноманітності в загальній українській популяції дорівнює $\alpha = 47 \times 10^2$ і є найвищим серед вивчених країн (табл. 1).

Щодо до вимоги нерівномірності розподілу по території, то відомо, що прізвища із суфіксом *-енко* розповсюджені в центрі та на сході країни, прізвища на *-ук*, *-юк*, *-ак*, *-як* поширені на заході (Редько, 1966, 1968). Рис. 1 демонструє розподіл двадцяти найчастіших в Україні прізвищ по п'яти регіонах країни. Ці території мають неоднакову історію (Грушевський, 1990), відрізняються антропологічними типами населення (Сегеда, 2001), розподілом груп крові (Тимошенко, 1978), маркерів мітохондріальної ДНК (Pshenichnov, 2007), гаплогруп Y хромосоми (Utevska, 2015). Наведене свідчить про зв'язок прізвищ з біологічними особливостями населення, з його генофондом, і щоб відобразити цей зв'язок кількісно були обчислені коефіцієнти кореляції між питомою вагою етнічної групи в популяціях і числом відповідних етнічних прізвищ серед ста найчастіших. Між цими

показниками виявлено прямий зв'язок, що описується позитивним коефіцієнтом кореляції Спірмана r_s : для українських прізвищ він дорівнює $r_s = 0,703$, російських $r_s = 0,739$, для «інших» $r_s = 0,393$.

Таблиця 1. Популяційні показники, одержані за прізвищами, в різних країнах
Table 1. Population indicators obtained by surnames in different countries

Популяція	$I_r \times 10^{-4}$	$a \times 10^2$	N	k	Автори
Австрія	2,60	8,54	1 000 000	140 766	Barrai, 2000
Бельгія	2,15	46,46	1 118 004	137 380	Barrai, 2004
Італія	1,71	—	5 100 000	215 623	Barrai, 1999
Німеччина	3,50	28,55	5 150 310	462 526	Rodriguez-Laralde, 1998
Нідерланди	4,20	23,78	2 400 210	126 485	Barrai, 2002
Франція	3,10	42,29	6 000 000	495 104	Scapoli, 2005
Україна	2,13	46,99	46 449 268*	697 147	Горпинченко, 2015
США	3,24	13,66	18 000 000	899 000	Barrai, 2001
Аргентина	11,02	4,22	22 600 000	414 441	Dipierri, 2005
Венесуела	18,00	2,10	3 905 487	68 665	Rodriguez-Laralde, 2000
Болівія	-	1,22	23 244 064	174 922	Rodriguez-Laralde, 2011
Парагвай	66,00	1,51	4 843 868	39 021	Dipierri, 2011
Якутія	23,10	4,33	491 259	44 625	Tarskaia, 2009
Китай	260,00	0,28	1 277 000 000*	7 327	Liu, 2012

Примітка. I_r — індекс ізонімії, a — індекс різноманіття N — чисельність вивчених груп, * — тотальний список, k — кількість прізвищ.

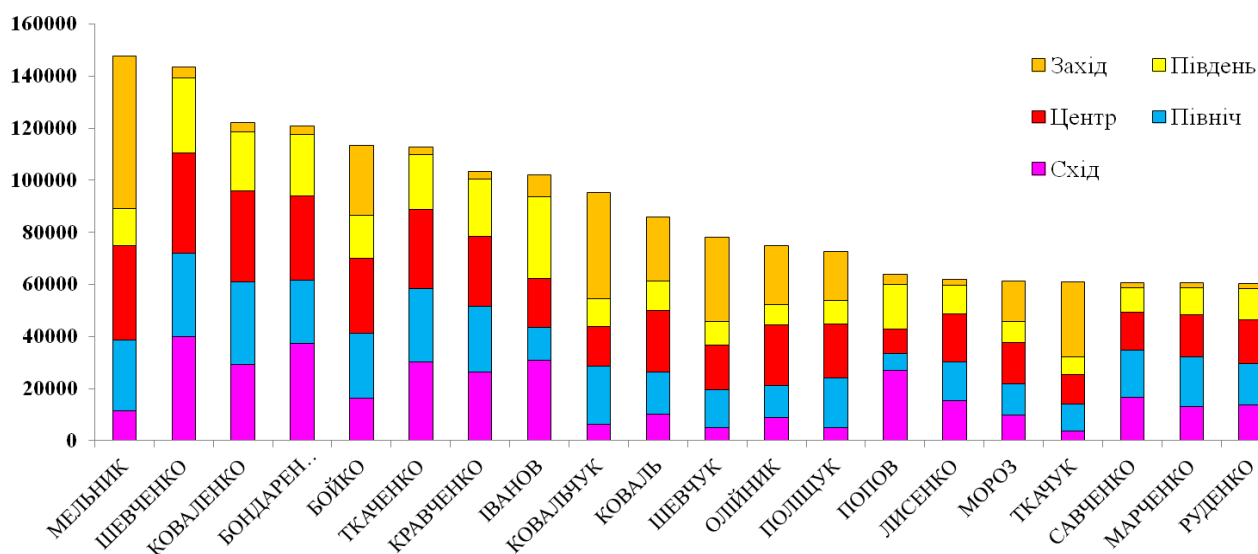


Рис. 1. Розподіл двадцяти найчастіших прізвищ в регіонах України (по вертикалі — кількість носіїв прізвищ)
Fig. 1. Distribution of the twenty most frequent surnames in the regions of Ukraine (vertically - the number of bearers of surnames)

Методом багатовимірного шкалювання одержано найважливіший аргумент для використання українських прізвищ в генетиці — це їх асоціація з біологічними ознаками. Встановлено, що між матрицями генетичних відстаней, розрахованих за групами крові систем АВ0, Rh, гаплогруп Y хромосоми і відстанями, і відстанями, розрахованими за прізвищами, існує зв'язок, що описується коефіцієнтом кореляції (r) становить 0,58–0,60, що співпадає зі значенням такого коефіцієнта в інших

слов'янських популяціях. Можливість українських прізвищ диференціювати популяції була з'ясована при аналізі заздалегідь розрахованих загальноприйнятих показників, за якими порівнювали популяції у горизонтальному (географічно) та вертикальному (ієрархічно) напрямках. Індекс ізонімії — інтегральний показник, що позначає ймовірність випадкового збігу прізвищ у двох членів популяції. Розподіл значень I_r для обласних популяцій демонструють чіткий градієнт, що збільшується у напрямку від південного сходу країни до північного заходу (рис. 2). Інші популяційні показники, розраховані на підставі I_r з використанням чисельності локального населення, також демонструють своєрідні територіально обумовлені градієнти, що заходять логічне пояснення в демографічних процесах, що мали місце в історії. Так показник різноманіття має найбільше значення серед областей України в Криму ($62,3 \times 10^2$) і обумовлений розташуванням місцевості на перетині торговельних шляхів і воєнних конфліктів, що призводили до переміщення великих мас різних народів. Найнижче значення показника різноманіття ($14,0 \times 10^2$) у Волинській області обумовлене особливостями ландшафту (ліси, болота), що перешкоджає міграціям. Саме це відбиває і показник генетичної ефективності міграції, який відбиває стабільність популяції під впливом міграції. Для населення України в цілому він дорівнює $v = 1,0 \times 10^{-4}$. Найбільшого генетичного впливу міграцій зазнало населення південних областей (Крим, Миколаївська та Херсонська області), де значення цього показника варіює в межах $(34,1-38,2) \times 10^{-4}$, а в інших областях в два-п'ять разів нижче. Коефіцієнт інбридингу функціонально пов'язаний з індексом ізонімії ($F_{st} = I_r/4$), тому його розподіл має таких вигляд, як на рис. 2. В цілому для населення України він дорівнює $F_{st} = 5,3 \times 10^{-5}$ з найвищим значенням серед областей в волинській популяції і найнижчим в Криму. Дотичність використання

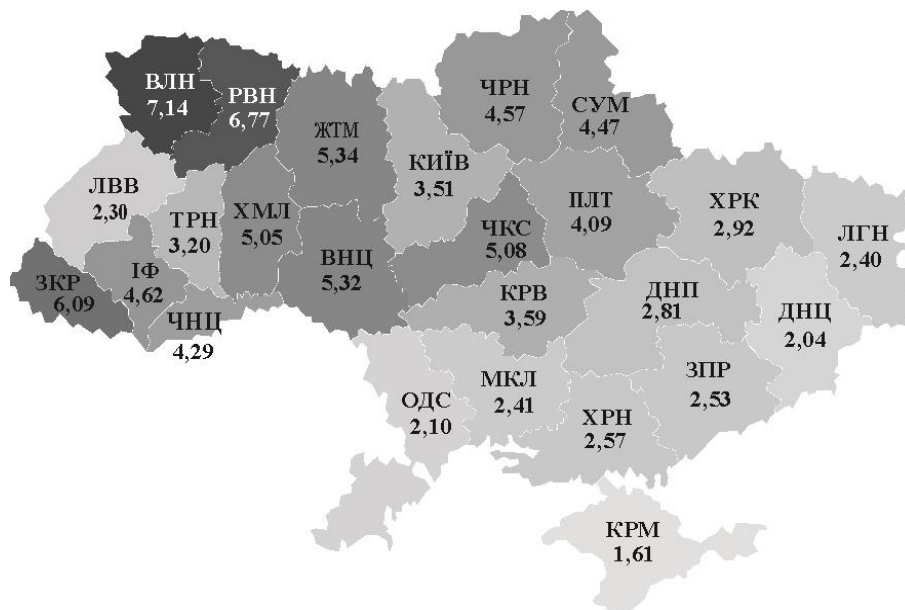


Рис. 2. Розподіл показника ізонімії $I_r (\times 10^2)$ в популяціях обласного рівня

(Позначення областей: ВНЦ — Вінницька, ВЛН — Волинська, ДНП — Дніпропетровська, ДНЦ — Донецька, ЖТМ — Житомирська, ЗКР — Закарпатська, ЗПР — Запорізька, ІФ — Івано-Франківська, КИЇВ — Київська, КРВ — Кіровоградська, КРМ — АР Крим, ЛГН — Луганська, ЛВВ — Львівська, МКЛ — Миколаївська, ОДС — Одеська, ПЛТ — Полтавська, РВН — Рівненська, СУМ — Сумська, ТРН — Тернопільська, ХРК — Харківська, ХРС — Херсонська, ХМЛ — Хмельницька, ЧРС — Черкаська, ЧРЦ — Чернівецька, ЧРН — Чернігівська)

Fig. 2. Distribution of the isonymy indicator $I_r (\times 10^2)$ in populations at the regional level

(Region designations: ВНЦ — Vinnytsia, ВЛН — Volyn, ДНП — Dnipropetrovsk, ДНЦ — Donetsk, ЖТМ — Zhytomyr, ЗКР — Zakarpattia, ЗПР — Zaporizhzhya, ІФ — Ivano-Frankivsk, КИЇВ — Kyiv, КРВ — Kirovohrad, КРМ — AR Crimea, ЛГН — Luhansk, ЛВВ — Lviv, МКЛ — Mykolaiv, ОДС — Odesa, ПЛТ — Poltava, РВН — Rivne, СУМ — Sumy, ТРН — Ternopil, ХРК — Kharkiv, ХРС — Kherson, ХМЛ — Khmelnytsky, ЧРС — Cherkasy, ЧРЦ — Chernivtsi, ЧРН — Chernihiv)

прізвищ для розрахунку інбридингу доведена також тим, що між F_{st} і чисельністю населення в районах існує зворотний зв'язок, відображений від'ємним коефіцієнтом кореляції ($r = -0,38$ в Одеській і $r = -0,80$ в Полтавській області).

Індекс місця I_p , обчислений за прізвищами, є аналогом генетичної відстані між популяціями. Він розраховується як середній ранг найбільш поширених прізвищ в популяціях, що порівнюються.

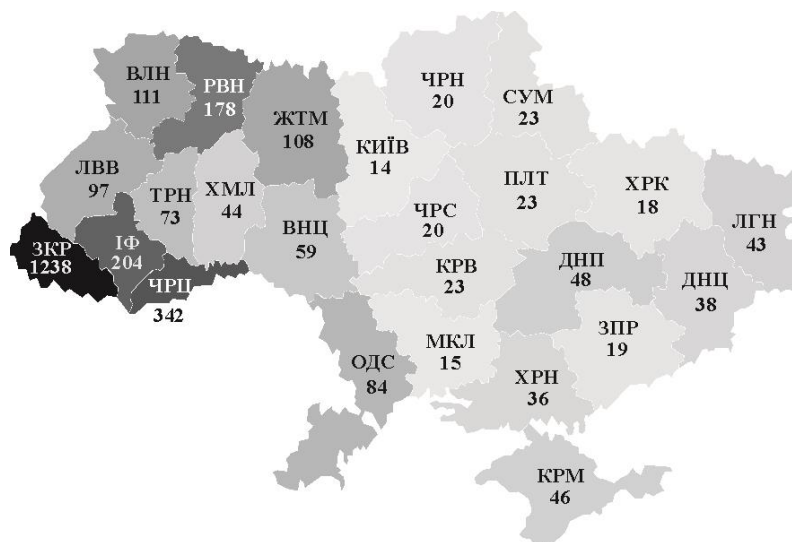


Рис. 3. Розподіл показника I_{20} по областях. Позначення як на рис. 2
Fig. 3. Distribution of the I_{20} indicator by regions. Designation as in Fig. 2

Середній індекс місця, розрахований за двадцятьма прізвищами для всієї України дорівнює $I_{20} = 10,5$. Близьче всіх до цього значення знаходиться київська обласна популяція ($I_{20} = 14$, рис. 3). У напрямку до кордонів значення індексу збільшується, особливо у західному напрямку, досягаючи відстані $I_{20} = 1238$ для закарпатської популяції. На рис. 3 вимальовується картина, що пояснюється подіями, які мали місце в минулому: знаходженню західної і східної частин України у складі різних імперій, західної в Австро-Угорській, східної у Російській. Встановлено, що роздільна здатність популяційних показників підвищується на нижчих ступенях популяційної ієрархії. Так, на обласному рівні чіткіше ніж на регіональному, спостерігається розподіл популяційних показників, який пояснюється глобальними історичними подіями. На районному рівні спостерігається явище, аналогічне дрейфу генів — ефект засновника. Він проявляється у величезній відстані між районними популяціями, що межують. Завдяки цьому мапи областей мають мозаїчний вигляд. Наслідком ефекту засновника служать приклади того, що в деяких областях дуже поширені прізвища, що є надзвичайно рідкісними в країні в цілому.

Досить сильний зв'язок прізвищ з генетичними маркерами дає підставу для екстраполяції значень, одержаних за допомогою прізвищ, на генофонд населення. Зокрема високе різноманіття прізвищ з великою ймовірністю свідчить про існування відповідного генетичного різноманіття, що забезпечує стійкість популяцій в екстремальних умовах, що постійно змінюються. Розподіл прізвищ, що відображується у популяційних показниках дає змогу уточнити історію окремих популяцій, орієнтувати демографів і медичних генетиків на локалітети з загрозливим станом для своєчасного вживання заходів для протидії негативних наслідків, а одержані популяційні показники служать точками у подальших дослідженнях.

Список літератури / References

- Грушевський М. Ілюстрована історія України — Репр. відтворення вид. 1913 р. — Київ, 1990. — 524 с. / Hrushevsky M. Illustrated history of Ukraine — repr. Interpretation of vol. 1913. — Kyiv, 1990. — 524 p.
- Редько Ю. К. Довідник українських прізвищ. / За редакцією кандидата філологічних наук Івана Варченка. Київ, 1968. 256 с. / Redko Y.K. (1968) Directory of Ukrainian surnames. Kiev : Naukova dumka. 256 p.
- Редько Ю. К. Сучасні українські прізвища. Київ, 1966. 216 с. / Redko Y.K. Modern Ukrainian surnames. Kiev : Naukova dumka, 1966 216 p. /

- Середа С. Антропология: Навч. посібник. Київ : Либідь, 2001. 336 с. / *Seheda S. (2001) Anthropology: Education. manual. Kyiv : Lybid. 336 p.*
- Barrai I., Cavalli-Sforza L. L., Moroni A. Isonymy structure of USA population // *Am J Phys Anthropol.* 2001. V. 114, № 2. P. 109–123. [https://doi.org/10.1002/1096-8644\(200102\)114:2<109::AID-AJPA1011>3.0.CO;2-I](https://doi.org/10.1002/1096-8644(200102)114:2<109::AID-AJPA1011>3.0.CO;2-I)
- Barrai I., Formica G., Scapoli C. et al. Microevolution in Ferrara: Isonymy 1890–1990 // *Annals of Human Biology.* 1992. V.19, № 4. P. 371–385. <https://doi.org/10.1080/03014469200002242>
- Nei M. *Molecular Population Genetics and Evolution.* North-Holland American Elsevier. 1975. 288 p.
- Pshenichnov A., Balanovska E., Balanovsky O. et al. Variation of paternal, maternal and autosomal genetic markers on intra-ethnic (Ukrainians) and inter-ethnic (Europe) level supports Y chromosomal marker bias // *International Society for Applied Biological Sciences. Split.* 2007. P.128.
- Timoshenko L. I., Lavrovskaya L. N. (1978) Allocation of erythrocyte antigens and protein factors of blood among the population of some genogeographic zones of the Ukrainian SSR. *Cytology and Genetics.* V. 12, № 6. P. 535–40.
- Utevska O. M., Pshenichnov A. S., Dibirova Kh. D., Rootsi S., Agdzhoyan A. T., Churnosov M. I., Balanovska E. V., Atramentova L. A. Balanovsky O. P. Gene Pool Similarities and Differences between Ukrainians and Russians of Slobozhanshchina Based on Y-Chromosome Data // *Cytology and Genetics.* 2015. V. 49, № 4. P. 245–253. <https://doi.org/10.3103/S0095452715040106>

Surnames of the ukrainian population as a tool of population and genetic research **L.O. Atramentova, V.V. Shuba, M.Yu. Gorpichenko**

The article presents the results of population genetic research, Execution with the help of surnames. It is proved that Ukrainian surnames meet the previous requirements for them as a tool of population genetic research: the stench is stably inherited for more than ten generations by patrilineal type; semantics and morphology are the most diverse in the world; unevenly distributed throughout the country, differing in history, demographic processes, anthropological types of population, the distribution of biological markers. According to the correlation coefficients between the weight of the ethnic group in the population and the number of ethnic surnames, a direct connection was found. There is also a direct link between the matrix of genetic distances calculated by blood groups, haplogroups of the Y chromosome and distances by surnames. A rather strong connection of surnames with genetic markers provides a basis for extrapolating the values obtained with the help of surnames to the gene pool of the population. The distribution of the 20 most frequent surnames in different regions is shown. Calculated indicators that reflect the structure of the population and the genetic processes that take place in them: index of isonymy I_r , diversity α , inbreeding F_{st} , migration v , order H and redundancy distribution R . These indicators are compared between populations in geographical and hierarchical location. The isonymy index shows a clear gradient that increases in the direction from the southeast to the northwest. Other indicators also change their value quite clearly from the South-East to the North-West directions. At lower levels of the population structure, the differential ability of indicators calculated by surnames increases. At the level of district populations, Identifying the founder's effect is analogous to genetic drift, which is particularly clearly demonstrated by the index of the place I_p . The distribution of population indicators is explained by historical events and demographic processes that have taken place in the history of the country.

Key words: *Ukrainian populations, surnames, indicators of isonymy, inbreeding, diversity, migrations*

Cite this article: Atramentova L.O., Shuba V.V., Gorpichenko M.Yu. (2024). *Surnames of the ukrainian population as a tool of population and genetic research. The Journal of V. N. Karazin Kharkiv National University. Series Biology*, 43, p. 42-47. <https://doi.org/10.26565/2075-5457-2024-43-4> (in Ukrainian)

About the authors:

Atramentova L.O. – V. N. Karazin Kharkiv National University, Svobody Maidan, 4, Kharkiv, Ukraine, 61022, lubov.atramentova@gmail.com, <https://orcid.org/0000-0002-7143-9411>

Shuba V.V. – V. N. Karazin Kharkiv National University, Svobody Maidan, 4, Kharkiv, Ukraine, 61022; Kharkiv Research Expert Forensic Center of the Ministry of Internal Affairs of Ukraine, Kovtuna Street, 34, Kharkiv, Ukraine, 61022, shubavladislav@gmail.com, <https://orcid.org/0000-0002-4261-5228>

Gorpichenko M.Yu. – Kharkiv National University named after V. N. Karazin, Svobody Maidan, 4, Kharkiv, Ukraine, 61022, 8.gelios.8@gmail.com, <https://orcid.org/0000-0003-0362-6300>

Received: 20.10.2023 / Revised: 10.07.2024 / Accepted: 19.08.2024