

Міністерство освіти і науки України
Харківський національний університет імені В. Н. Каразіна

Кваліфікаційна наукова
праця на правах рукопису

ШУБА ВЛАДИСЛАВ ВОЛОДИМИРОВИЧ

УДК 575.174+314.92

ДИСЕРТАЦІЯ
СТРУКТУРА УКРАЇНСЬКИХ ПОПУЛЯЦІЙ РІЗНОГО
ІЄРАРХІЧНОГО РІВНЯ ЗА ДАНИМИ ПРІЗВИЩ

Спеціальність 091 Біологія
Галузь знань 09 Біологія

Подається на здобуття наукового ступеня
доктора філософії

Дисертація містить результати власних досліджень. Використання ідей,
результатів і текстів інших авторів має посилання на відповідні джерела

_____ В. В. Шуба

Науковий керівник: Атраментова Любов Олексіївна, доктор біологічних наук,
професор

Харків — 2026

АНОТАЦІЯ

Шуба В. В. Структура українських популяцій різного ієрархічного рівня за даними прізвищ. — Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття ступеня доктора філософії за спеціальністю 091 Біологія. Галузь знань 09 Біологія — Харківський національний університет імені В. Н. Каразіна, Міністерства освіти і науки України, Харків, 2026.

Дисертаційне дослідження присвячене аналізу популяційно-генетичної структури України на основі розподілу прізвищ як непрямих генетичних маркерів. Робота виконана на масиві даних, що охоплює 46 431 365 носіїв 697 147 прізвищ з усіх областей України станом на 2001 рік, що дає підстави вважати її одним із найбільших популяційно-генетичних досліджень українського населення.

Теоретичною основою роботи є метод ізонімії, що базується на фундаментальній аналогії між патрилійною передачею прізвищ і успадкуванням Y-хромосоми. Наукова новизна дослідження полягає у: (1) реалізації популяційно-генетичного аналізу на основі бази даних загальнонаціонального масштабу для аналізу на районному рівні; (2) адаптації та практичному застосуванні показника CRSED для аналізу адміністративно-територіальних одиниць України; (3) використанні підходу до кластерного аналізу із залученням трьох незалежних метрик міжпопуляційної дистанції; (4) а також у впровадженні ієрархічної моделі аналізу «область–район», що забезпечує системне відображення просторової та демографічної структури населення України.

Застосовано багаторівневий підхід до аналізу, що передбачає обчислення популяційно-генетичних показників (індексу ізонімії, коефіцієнта інбридингу, індексу різноманітності Фішера, індексу міграції Карліна — МакГрегора) та інноваційних метрик (коефіцієнта покриття розширеного експоненціального

розподілу CRSED). Аналіз проведено на двох ієрархічних рівнях: обласному (25 областей) та районному (490 адміністративних одиниць).

Аналіз 25 областей і 490 адміністративних районів України виявив складну багатопшарову популяційно-генетичну структуру з виразною просторовою диференціацією. На районному рівні індекс ізонімії варіює від $1,61 \times 10^{-4}$ (м. Сімферополь, АР Крим) до $98,48 \times 10^{-4}$ (Зарічненський район Рівненської області), коефіцієнт інбридингу — від $3,35 \times 10^{-5}$ (Джанкойський район АР Крим) до $246,19 \times 10^{-5}$ (Зарічненський район), індекс різноманітності Фішера — від $2,49 \times 10^2$ (Шацький район) до $63,18 \times 10^2$ (м. Сімферополь). Така варіабельність відображає історично сформовану демографічну гетерогенність, інтенсивні міграційні процеси та контрастні моделі заселення — від урбанізованих індустріальних центрів до ізольованих сільських територій.

Застосування методу CRSED дало змогу класифікувати області України на три категорії залежно від їхньої демографічної історії. До районів імміграції (клас I, $CRSED \leq 0,01$) віднесено переважно східні та південні області (АР Крим, Донецьку, Луганську, Харківську, Запорізьку, Дніпропетровську), що характеризуються високою мобільністю населення та інтенсивним припливом мігрантів у XIX–XX століттях. Рекультивовані області (клас II, $0,01 < CRSED \leq 0,05$) представлені центральними областями (Полтавською, Черкаською, Хмельницькою, Вінницькою, Тернопільською, Сумською, Житомирською), які зазнали масштабних демографічних потрясінь у 1930–1940-х роках. Области еміграції (клас III, $CRSED > 0,05$) локалізовані на заході та півночі України (Львівській, Волинській, Рівненській, Чернівецькій, Чернігівській областях), де історично переважав відтік населення.

Детальний аналіз на районному рівні із застосуванням методу Уорда та трьох типів генетичних дистанцій (евклідової, Ласкера, Нея) виявив 10–11 основних популяційних кластерів. У результаті аналізу ідентифіковано шість стабільних філогенетичних одиниць, які зберігають свою цілісність незалежно від обраної метрики: АР Крим, Закарпатську, Волинську, Вінницьку, Львівську

області та Східну Україну. Виявлено чотири основні філогенетичні осі популяційної структури України: Схід–Захід, Північ–Південь, рівнина–гори та вісь індустріалізації.

Аналіз внутрішньообласної гетерогенності дав змогу встановити ієрархію областей за складністю демографічної історії. Одеська область є найгетерогеннішим регіоном (розподіл між 3–4 кластерами), що відображає її багатокомпонентну історію заселення. Київська (4–5 кластерів) та Кіровоградська (4 кластери) області також характеризуються високою внутрішньою диференціацією. Застосування тесту Мантеля виявило статистично значущу кореляцію між генетичними та географічними відстанями, що свідчить про модель ізоляції за відстанню в українських популяціях. Порівняльний аналіз українських популяцій з іншими країнами світу засвідчив проміжне положення України між західноєвропейськими та більш ізольованими популяціями.

Практичне значення роботи полягає в отриманні похідного набору референтних характеристик популяційно-генетичної структури України станом на початок ХХІ століття, що дасть змогу оцінити зміни, які відбуваються нині та матимуть значення для медичної генетики, антропології, соціології та демографії.

Основні результати дисертації опубліковано у семи наукових працях, повний перелік яких наведено в додатку А.

Ключові слова: Україна, Європа, історична демографія, філогенетика, біоінформатика, міграція, популяція, географічне різноманіття, біогеографія прізвищ.

SUMMARY

Shuba V. V. Structure of Ukrainian Populations at Different Hierarchical Levels Based on Surname Data. — Qualifying scientific work presented as a manuscript.

Dissertation for the degree of Doctor of Philosophy, Specialty 091 Biology. Field of Knowledge 09 Biology — V. N. Karazin Kharkiv National University, Ministry of Education and Science of Ukraine, Kharkiv, 2026.

The thesis analyzes the population-genetic structure of Ukraine using surname distributions as indirect genetic markers. The study was conducted on a nationwide dataset comprising 46,431,365 individuals bearing 697,147 distinct surnames across all regions of Ukraine as of 2001. This represents one of the largest population-genetic studies of the Ukrainian population to date.

The theoretical foundation of the research is the isonymy method, based on the fundamental analogy between the patrilineal transmission of surnames and the inheritance of the Y chromosome. The scientific novelty of the study lies in the following: (1) a district-level population-genetic analysis using a nationwide database; (2) the adaptation and application of the CRSED (Coverage Ratio of the Stretched Exponential Distribution) coefficient for classifying administrative units; (3) a multi-metric cluster analysis employing three independent interpopulation distance measures; and (4) the development of a hierarchical “region–district” analytical framework that provides a systematic representation of the spatial and demographic structure of the Ukrainian population.

A multilevel approach was applied. It incorporated both classical population-genetic indices (isonymy index, inbreeding coefficient, Fisher’s alpha diversity index, and Karlin–McGregor migration index) and the novel CRSED metric. This analysis was conducted at two hierarchical levels: regional (25 regions) and district (490 administrative units).

The results reveal a complex, multilayered population-genetic structure with pronounced spatial differentiation. At the district level, the isonymy index ranged

from 1.61×10^{-4} (Simferopol, AR Crimea) to 98.48×10^{-4} (Zarichne District, Rivne Region); the inbreeding coefficient ranged from 3.35×10^{-5} (Dzhankoi District, AR Crimea) to 246.19×10^{-5} (Zarichne District); and Fisher's alpha diversity index ranged from 2.49×10^2 (Shatsk District) to 63.18×10^2 (Simferopol). This variability reflects historically shaped demographic heterogeneity, intensive migration processes, and contrasting settlement patterns — from highly urbanized industrial centers to isolated rural areas.

Application of the CRSED method enabled the classification of Ukrainian regions into three categories based on their demographic history: Immigration regions (Class I, $CRSED \leq 0.01$) — predominantly eastern and southern regions (AR Crimea, Donetsk, Luhansk, Kharkiv, Zaporizhzhia, and Dnipropetrovsk), characterized by high population mobility and intensive in-migration in the 19th–20th centuries; Recovered regions (Class II, $0.01 < CRSED \leq 0.05$) — central regions (Poltava, Cherkasy, Khmelnytskyi, Vinnytsia, Ternopil, Sumy, and Zhytomyr), which suffered large-scale demographic upheavals during the 1930s–1940s; Emigration regions (Class III, $CRSED > 0.05$) — located in western and northern Ukraine (Lviv, Volyn, Rivne, Chernivtsi, and Chernihiv regions), where historical population outflow prevailed.

Detailed district-level cluster analysis using Ward's method and three genetic distance metrics (Euclidean, Lasker, and Nei) identified 10–11 major population clusters. Six stable phylogenetic units were consistently recovered regardless of the distance metric used: AR Crimea, Zakarpattia, Volyn, Vinnytsia, Lviv regions, and Eastern Ukraine. Four principal axes of population structure were distinguished: East–West, North–South, plain–mountain, and an industrialization axis.

Analysis of intra-regional heterogeneity established a hierarchy of regions according to the complexity of their demographic histories. Odesa Region was the most heterogeneous (distributed across 3–4 clusters), reflecting its multicomponent settlement history. Kyiv and Kirovohrad regions also demonstrated high internal differentiation. The Mantel test confirmed a statistically significant correlation between genetic and geographic distances, supporting an isolation-by-distance

model in Ukrainian populations. Comparative analysis placed Ukraine in an intermediate position between highly mobile Western European populations and more isolated ones.

The practical significance of the study lies in the establishment of a reference dataset of population-genetic characteristics of Ukraine at the beginning of the 21st century. These results provide a baseline for monitoring current demographic changes and are relevant for medical genetics, anthropology, sociology, and demography.

The main results of the dissertation have been published in seven scientific papers, the full list of which is provided in Appendix A.

Keywords: Ukraine, Europe, historical demography, phylogenetics, bioinformatics, migration, population, geographic variation, biogeography of surnames.

СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА ЗА ТЕМОЮ ДИСЕРТАЦІЇ

Наукові праці, у яких опубліковані основні наукові результати дисертації (публікації у фахових виданнях України):

1. **Шуба В. В.**, Горпинченко М. Ю., Атраментова Л. О. (2023). Вплив господарської діяльності на структуру популяцій на прикладі Луганської області. *Фактори експериментальної еволюції організмів*, (33), 114–117. [DOI: 10.7124/FEEO.v33.1577](https://doi.org/10.7124/FEEO.v33.1577) (Особистий внесок здобувача: проведено пошук наукової літератури, виконано розрахунки, здійснено їх аналіз та узагальнення, написано текст; особистий внесок Атраментової Л. О.: концепція роботи, написання тексту; особистий внесок Горпинченка М. Ю.: підготовка наукової ілюстрації, написання та форматування тексту).
2. **Шуба В. В.**, Горпинченко М. Ю., Атраментова Л. О. (2024). Популяційно-генетичний аналіз населення Львівської області за даними прізвищ. *Фактори експериментальної еволюції організмів*, (34), 44–48. [DOI: 10.7124/FEEO.v34.1615](https://doi.org/10.7124/FEEO.v34.1615) (Особистий внесок здобувача: проведено пошук наукової літератури, виконано розрахунки, здійснено їх аналіз та узагальнення, написано текст; особистий внесок Атраментової Л. О.: концепція роботи, написання тексту; особистий внесок Горпинченка М. Ю.: підготовка наукової ілюстрації, написання та форматування тексту).
3. Атраментова Л. О., **Шуба В. В.**, Горпинченко М. Ю. (2024). Прізвища українського населення як інструмент популяційно-генетичних досліджень *Вісник Харківського національного університету імені В. Н. Каразіна. Серія «Біологія»*, (43), 44–49. [DOI: 10.26565/2075-5457-2024-43-4](https://doi.org/10.26565/2075-5457-2024-43-4) (Особистий внесок здобувача: проведено пошук наукової літератури, виконано частину розрахунків, написано текст; особистий внесок Атраментової Л. О.: концепція роботи, написання тексту; особистий внесок Горпинченка М. Ю.: виконано частину розрахунків, підготовлено наукової ілюстрації, написано текст).

4. **Шуба В. В.**, Атраментова Л. О. (2024). Популяційно-генетичний аналіз населення Запорізької області за даними прізвищ. *Вісник Львівського університету. Серія біологічна*, (93), 3–9. [DOI: 10.30970/vlubs.2024.93.01](https://doi.org/10.30970/vlubs.2024.93.01) (Особистий внесок здобувача: проведено пошук наукової літератури, виконано розрахунки, здійснено їх аналіз та узагальнення, написано текст; особистий внесок Атраментової Л. О.: концепція роботи, написання тексту).
5. **Шуба В. В.**, Горпинченко М. Ю., Атраментова Л. О. (2025). Популяційний аналіз населення Волинської області за даними прізвищ. *Вісник Львівського університету. Серія біологічна*, (94), 37–43. [DOI: 10.30970/VLUBS.2025.94.04](https://doi.org/10.30970/VLUBS.2025.94.04) (Особистий внесок здобувача: проведено пошук наукової літератури, виконано розрахунки, здійснено їх аналіз та узагальнення, написано текст; особистий внесок Атраментової Л. О.: концепція роботи, написання тексту; особистий внесок Горпинченка М. Ю.: підготовка наукової ілюстрації, написання та форматування тексту).
6. **Шуба В. В.**, Горпинченко М. Ю., Атраментова Л. О. (2024). Характеристика популяцій Київської області за даними прізвищ. *Фактори експериментальної еволюції організмів*, (36), 28–32. [DOI: 10.7124/FEEO.v36.1708](https://doi.org/10.7124/FEEO.v36.1708) (Особистий внесок здобувача: проведено пошук наукової літератури, виконано розрахунки, здійснено їх аналіз та узагальнення, написано текст; особистий внесок Атраментової Л. О.: концепція роботи, написання тексту; особистий внесок Горпинченка М. Ю.: підготовка наукової ілюстрації, написання і форматування тексту).

Наукові праці, в яких опубліковані основні наукові результати дисертації у періодичних наукових виданнях що входять до міжнародних наукометричних баз Scopus або Web of Science:

1. **Shuba V.**, Yasinskyi Y., Atramentova L. (2025). The Structure of the Ukrainian Population Based on Surname Distribution. *American Journal of Human Biology*. (37,7), e70097. [DOI: 10.1002/ajhb.70097](https://doi.org/10.1002/ajhb.70097) (Особистий внесок

здобувача: проведено пошук та аналіз наукової літератури, виконано розрахунки, здійснено їх аналіз та узагальнення, написано текст; особистий внесок Атраментової Л. О.: концепція роботи, написання тексту; особистий внесок Ясінського Я.С.: аналіз наукової літератури, виконано розрахунки, формальний аналіз).

Наукові праці, які засвідчують апробацію матеріалів дисертації

1. **Шуба В. В.**, Горпинченко М. Ю., Атраментова Л. О. (2023). Вплив господарської діяльності на структуру популяцій на прикладі Луганської області. XVIII-а Міжнародна наукова конференція «Фактори експериментальної еволюції організмів», 26–29 вересня 2023 р., м. Київ, 2023.
2. **Шуба В. В.**, Горпинченко М. Ю., Атраментова Л. О. (2024). Популяційно-генетичний аналіз населення Львівської області за даними прізвищ. XIX-а Міжнародна наукова конференція «Фактори експериментальної еволюції організмів», 23–26 вересня 2024 р., м. Тернопіль, 2024.
3. **Шуба В. В.**, Горпинченко М. Ю., & Атраментова Л. О. (2025). Характеристика популяцій Київської області за даними прізвищ. // XX-а Міжнародна наукова конференція «Фактори експериментальної еволюції організмів», 22–26 вересня 2025 р., м. Київ, 2025.

ЗМІСТ

ВСТУП	14
РОЗДІЛ 1 ОГЛЯД І АНАЛІЗ НАУКОВОЇ ЛІТЕРАТУРИ.....	20
1.1. Теоретичні основи популяційно-генетичного аналізу прізвищ.....	20
1.2. Основні популяційно-генетичні показники, отримані з використанням прізвищ.....	22
1.2.1. Показники внутрішньопопуляційної структури	22
1.2.2. Показники різноманітності розподілу	23
1.2.3. Міжпопуляційні відстані	24
1.2.4. Просторовий аналіз.....	25
1.3. Популяційно-генетичні дослідження прізвищ у різних країнах світу.....	25
1.3.1. Європейські популяції	25
1.3.2. Острівні та ізольовані популяції	31
1.3.3. Азійські популяції	31
1.3.4. Північноамериканські популяції	35
1.3.5. Латиноамериканські популяції	36
1.4. Вплив історичних подій, географічних умов і культурних особливостей на структуру популяцій	42
1.4.1. Історичні чинники.....	42
1.4.2. Географічні чинники	44
1.4.3. Культурні чинники	46
1.5. Популяційно-генетичні дослідження прізвищ в Україні.....	46
1.5.1. Молекулярно-генетичні дослідження.....	46
1.5.2. Перші дослідження прізвищ	48
1.5.3. Основні результати попередніх досліджень.....	48
1.5.4. Валідація методу ізонімії через молекулярно-генетичні дані	49
1.6. Порівняльний аналіз популяційно-генетичних показників	49
1.7. Методологічні обмеження та перспективи досліджень	51
УЗАГАЛЬНЕННЯ ТА ВИСНОВКИ ДО РОЗДІЛУ 1	54
РОЗДІЛ 2 МАТЕРІАЛИ ТА МЕТОДИ	56
2.1. Джерело даних.....	56
2.2. Статистичний аналіз	58

2.2.1. Частоти прізвищ та показники ізонімії.....	60
2.2.2. Показники різноманітності	61
2.2.3. Показники популяційної структури.....	66
2.2.4. Міжпопуляційні відстані	67
2.2.5. Просторовий аналіз та регіональна специфіка	69
2.2.6. Кластеризація та багатовимірні візуалізації.....	71
2.2.7. Контроль якості та числова стабільність	71
2.3. Програмне забезпечення генетико-статистичного аналізу	72
УЗАГАЛЬНЕННЯ ТА ВИСНОВКИ ДО РОЗДІЛУ 2	74
РОЗДІЛ 3 РЕЗУЛЬТАТИ ТА ОБГОВОРЕННЯ	76
3.1. Загальноукраїнські значення показників.....	76
3.2. Аналіз розподілу прізвищ за методом CRSED	78
3.2.1. Теоретичні основи методу	78
3.2.2. Регіональна варіація параметрів CRSED	79
3.2.3. Класифікація областей за CRSED	81
3.2.4. Порівняльний аналіз і методологічні висновки	84
3.3. Районні значення показників: детальний просторовий аналіз	87
3.3.1. Схід України	89
3.3.2. Північ України	93
3.3.3. Центр України	98
3.3.4. Захід України	103
3.3.5. Південь України.....	110
3.4. Міжпопуляційні генетичні дистанції та кластерний аналіз.....	115
3.4.1. Кластерна структура за евклідовою дистанцією	115
3.4.2. Кластерна структура за дистанцією Ласкера	121
3.4.3. Кластерна структура за дистанцією Нея.....	127
3.4.4. Порівняльний аналіз та інтерпретація кластерної структури районного рівня	133
3.4.4.1. Топологічна стабільність кластерної архітектури	133
3.4.4.2. Консенсусні філогенетичні одиниці.....	135
3.4.4.3. Метрикоспецифічні ефекти та їхнє біологічне значення.....	137
3.4.4.4. Області з екстремальною внутрішньообласною гетерогенністю	140

3.4.4.5. Ізоляція за відстанню	142
3.4.4.6. Історико-демографічна інтерпретація макрорегіональної структури	143
3.4.4.7. Чотири філогенетичні осі популяційної структури України	146
3.4.4.8. Методологічне значення порівняння трьох метрик дистанцій	149
3.4.4.9. Внутрішньообласна гетерогенність як індикатор демографічної складності	151
3.4.4.10. Висновки.....	156
УЗАГАЛЬНЕННЯ ТА ВИСНОВКИ ДО РОЗДІЛУ 3	157
НАУКОВЕ ТА ПРАКТИЧНЕ ЗНАЧЕННЯ РОБОТИ	158
ВИСНОВКИ	161
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	163
Додаток А.....	176
Додаток Б	179

ВСТУП

Обґрунтування вибору теми дослідження. Популяційно-генетичний аналіз є фундаментальним інструментом вивчення демографічної структури людських популяцій, реконструкції міграційних процесів та оцінювання рівнів спорідненості між регіональними групами населення. Розуміння генетичної структури популяцій має не лише теоретичне значення, але й безпосередні практичні застосування — від медичної генетики до антропології та соціальних наук.

Україна являє собою унікальний об'єкт для популяційно-генетичних досліджень. Розташована на перехресті міграційних шляхів між Сходом і Заходом, вона упродовж тисячоліть зазнавала впливу різноманітних демографічних процесів: хвиль заселення з різних напрямків, завоювань, колонізації, масових депортацій та індустріальних міграцій ХХ століття. Наслідком цих процесів є надзвичайна популяційно-генетична гетерогенність, яка дотепер не отримала повноцінного кількісного опису на районному рівні.

Метод ізонімії — аналіз розподілу прізвищ як непрямих генетичних маркерів — є оптимальним підходом для повнопопуляційних досліджень у великих часових і просторових масштабах. На відміну від молекулярно-генетичних методів, він дозволяє охопити все населення без формування вибірки, а також ретроспективно аналізувати демографічну структуру на основі адміністративних реєстрів. Аналогія між патрилійною передачею прізвищ та успадкуванням Y-хромосоми робить прізвища надійними маркерами для реконструкції патрилійної компоненти генетичної структури.

Незважаючи на значний масив попередніх досліджень, виконаних на матеріалах окремих областей України, загальнонаціональний популяційно-генетичний аналіз на районному рівні з охопленням усіх 490 адміністративних одиниць досі не проводився. Відсутній також систематичний порівняльний аналіз України у контексті світових популяцій з використанням єдиного

методологічного апарату. Заповнення цієї прогалини є актуальним науковим завданням, що зумовлює доцільність цього дослідження.

Зв'язок роботи з науковими програмами, планами, темами.

Дисертацію виконано на кафедрі генетики і цитології Харківського національного університету імені В. Н. Каразіна в межах науково-дослідної роботи «Генетико-демографічні процеси в українських популяціях» (номер держреєстрації 0224U032384, 2021–2024 рр.).

Мета дослідження: здійснити популяційно-генетичний аналіз структури населення України на обласному та районному ієрархічних рівнях на основі розподілу прізвищ, встановити просторові закономірності демографічної диференціації та реконструювати основні чинники, що сформували сучасний генетичний ландшафт країни.

Для досягнення мети було поставлено такі ***завдання:***

1. Провести систематичний аналіз наукової літератури з метою визначення місця українських популяцій за популяційно-генетичними показниками у глобальному контексті.

2. Розрахувати популяційно-генетичні показники (індекс ізонімії, коефіцієнт інбридингу, індекс різноманітності Фішера, індекс міграції Карліна — МакГрегора, ентропію Шеннона, показник надмірності розподілу) для всіх 490 адміністративних районів України.

3. Проаналізувати показник коефіцієнта покриття розширеного експоненціального розподілу (CRSED) для класифікації областей України за типом демографічної динаміки.

4. Провести кластерний аналіз 490 районів методом Уорда з використанням трьох незалежних метрик міжпопуляційної генетичної дистанції (евклідової, Ласкера, Нея) та ідентифікувати стабільні популяційно-генетичні одиниці.

5. Визначити, чи існує кореляція між генетичними і географічними дистанціями в українських популяціях, за тестом Мантеля.

Об'єкт дослідження: українські популяції.

Предмет дослідження: генетична структура популяцій людини.

Методи дослідження: популяційно-генетичні — розрахунок індексу ізонімії, коефіцієнта інбридингу, показників різноманітності та міграції; статистичний — нелінійна регресія методом Левенберга — Марквардта для підгонки моделі CRSED; кластерний аналіз — ієрархічна агломерація методом Уорда за трьома метриками генетичних дистанцій; просторово-статистичний — тест Мантеля для оцінювання кореляції між матрицями генетичних і географічних відстаней; порівняльний — зіставлення отриманих показників з опублікованими даними для понад двадцяти країн світу. Усі обчислення виконано засобами мови програмування R з використанням пакетів *minpack.lm*, *vegan*, *dplyr* та ін.

Наукова новизна одержаних результатів.

Здійснено повнонаціональний популяційно-генетичний аналіз населення України на районному рівні з охопленням усіх 490 адміністративних одиниць на основі даних про 46 431 365 носіїв 697 147 прізвищ.

Для аналізу українських популяцій застосовано метод CRSED (коефіцієнт покриття розширеного експоненціального розподілу), що дозволило класифікувати регіони на три категорії за типом демографічної динаміки: регіони імміграції, рекультивовані регіони та регіони еміграції.

Запропоновано та реалізовано багатометричний підхід до кластерного аналізу з паралельним застосуванням трьох незалежних метрик міжпопуляційної дистанції (евклідової, Ласкера, Нея), що забезпечує методологічну валідацію стабільних популяційно-генетичних одиниць через їх незалежну ідентифікацію.

Впроваджено ієрархічну аналітичну модель «область — район», яка забезпечує системне відображення просторової та демографічної структури населення на двох взаємопов'язаних ієрархічних рівнях.

Ідентифіковано шість популяційно-генетичних одиниць (АР Крим, Закарпатську, Волинську, Вінницьку, Львівську області та Східну Україну), стабільних незалежно від обраної метрики дистанції, та чотири основні

філогенетичні осі структури населення України: Схід–Захід, Північ–Південь, рівнина–гори та вісь індустріалізації.

Теоретичне значення одержаних результатів. На методологічному рівні робота обґрунтовує переваги багатометричного підходу до кластерного аналізу популяцій: паралельне застосування трьох незалежних метрик генетичної дистанції (евклідової, Ласкера, Нея) дозволяє розмежувати стабільні популяційно-генетичні одиниці, що відтворюються незалежно від обраного методу виміру, від артефактів конкретної метрики. Цей підхід є загальним і може застосовуватися в популяційно-генетичних дослідженнях будь-яких національних популяцій.

На рівні популяційно-генетичної теорії результати емпірично підтверджують дієвість моделі ізоляції за відстанню (isolation by distance) для великої континентальної популяції з неоднорідною та катастрофічною демографічною історією. Тест Мантеля виявив статистично значущий зв'язок між генетичною та географічною відстанню ($p < 0,001$): коефіцієнти кореляції становлять 0,581 (за Евклідом), 0,691 (за Ласкером) і 0,246 (за Неєм). Це свідчить про те, що географічна структура генофонду збереглася навіть після масштабних міграцій ХХ століття. Це є теоретично нетривіальним результатом, що ставить під сумнів уявлення про повну гомогенізацію популяцій унаслідок радянської індустріалізації та урбанізації.

Дослідження виявило чотири ключові осі популяційної структури: Схід–Захід, Північ–Південь, рівнина–гори та вісь індустріалізації. Це дозволяє краще зрозуміти, як формується генетичне різноманіття великих національних популяцій: поряд із географічними бар'єрами та відстанню важливу роль відіграє характер господарського освоєння території — чинник, який для України досі не був кількісно підтверджений.

У контексті історичної демографії робота демонструє, що демографічні катастрофи залишають вимірний і просторово диференційований слід у розподілі прізвищ навіть через шість–сім десятиліть. Виявлена типологія регіонів за показником CRSED (регіони імміграції, рекультивовані регіони,

регіони еміграції) є кількісним відображенням цих процесів і становить внесок у теорію демографічних наслідків масових примусових міграцій.

Практичне значення одержаних результатів. Сформовано базову референтну характеристику популяційно-генетичної структури України станом на початок ХХІ століття, яка може слугувати відправною точкою для моніторингу демографічних змін, спричинених подіями 2014 та 2022 років. Виявлені регіональні відмінності коефіцієнта інбридингу (від $4,0 \times 10^{-5}$ у Криму до $17,9 \times 10^{-5}$ на Волині) можуть використовуватися для оптимізації медико-генетичного консультування та скринінгових програм у регіонах із підвищеним рівнем спорідненості. Результати мають значення для антропології, соціології, демографії та можуть застосовуватися як допоміжний статистичний інструмент у судово-медичній практиці для оцінювання ймовірного географічного походження.

Особистий внесок здобувача. Здобувач взяв участь у розробці концепції дослідження, опрацював наукову літературу, виконав розрахунки та статистичний аналіз, здійснив інтерпретацію результатів і написав текст дисертації. У спільних публікаціях з М. Ю. Горпинченком останній забезпечував підготовку наукових ілюстрацій і форматування текстів; у публікації з Я. С. Ясінським — участь у розрахунках і формальному аналізі. Ідеї та матеріали співавторів у дисертації не використовуються.

Апробація результатів дисертації. Основні положення та результати дисертації оприлюднено на ХVІІІ-тій Міжнародній науковій конференції «Фактори експериментальної еволюції організмів», 26–29 вересня 2023 р., м. Київ, 2023; ХІХ-тій Міжнародній науковій конференції «Фактори експериментальної еволюції організмів», 23–26 вересня 2024 р., м. Тернопіль, 2024; ХХ-тій Міжнародній науковій конференції «Фактори експериментальної еволюції організмів», 22–26 вересня 2025 р., м. Київ, 2025.

Публікації. Основні результати дисертації викладено у семи наукових працях: шість статей у фахових виданнях України та одна стаття у виданні, що входить до міжнародної наукометричної бази Web of Science / Scopus

(*American Journal of Human Biology*, 2025). Повний перелік публікацій наведено в Додатку А.

Структура та обсяг дисертації. Дисертація складається зі вступу, трьох розділів, узагальнень і висновків до кожного розділу, розділу «Наукове та практичне значення роботи», висновків, списку використаних джерел і двох додатків. Загальний обсяг роботи становить 203 сторінки. Основний текст викладено на 148 сторінках. Робота містить 28 таблиць, 33 рисунки. Список використаних джерел налічує 126 найменувань, з яких 91 іноземними мовами.

Подяки. Автор висловлює щиру подяку Михайлу Юрійовичу Горпинченку за багаторічну плідну співпрацю в написанні спільних наукових публікацій, форматуванні текстів, а також за цінні методичні поради та обговорення результатів аналізу.

Окрему подяку автор висловлює Ярославу Сергійовичу Ясінському за участь у розрахунках і формальному аналізі під час підготовки статті у виданні *American Journal of Human Biology* (2025).

РОЗДІЛ 1 ОГЛЯД І АНАЛІЗ НАУКОВОЇ ЛІТЕРАТУРИ

1.1. Теоретичні основи популяційно-генетичного аналізу прізвищ

Прізвища як культурно передавані ідентифікатори містять значну інформацію про родинне походження, соціальну організацію та складну взаємодію історичних і географічних факторів. Метод ізонімії — аналіз розподілу прізвищ для оцінки генетичної структури популяцій — базується на фундаментальній аналогії між передачею прізвищ по батьківській лінії та успадкуванням Y-хромосоми. Ця аналогія дає змогу використовувати прізвища як непрямі генетичні маркери для реконструкції демографічних процесів, міграційних потоків і рівня інбридингу в популяціях.

Застосування прізвищ у популяційній генетиці має глибоке теоретичне обґрунтування. Подібно до Y-хромосомних маркерів, прізвища передаються по чоловічій лінії без рекомбінації, що робить їх придатними для відстеження патрилійної спадковості [101]. Однак необхідно враховувати суттєві обмеження цього підходу.

По-перше, прізвища можуть змінюватися внаслідок соціальних процесів (русифікації, полонізації, адміністративних реформ), тоді як генетичні маркери залишаються стабільними. Для України це особливо актуально через складну історію перебування у складі різних держав: процеси русифікації в XVIII–XX століттях призвели до масової зміни українських прізвищ на російські форми (*Василенко* → *Васильєв*, *Гриценко* → *Гриценков*), полонізація в XVI–XVIII століттях трансформувала прізвища за польськими моделями (*Коваль* → *Ковальський*) [46, 82, 93]. Ці процеси могли зменшити інформативність прізвищ як маркерів патрилійного походження, особливо у східних та центральних областях.

По-друге, ідентичні прізвища можуть виникати незалежно внаслідок конвергентних процесів називництва, особливо якщо йдеться про прізвища, похідні від поширених професій (*Коваль*, *Ковальчук*, *Мельник*) або географічних назв (*Дніпровський*, *Київський*). Проте систематична похибка

такого типу є найбільш імовірною для дуже поширених прізвищ, тоді як рідкісні прізвища з більшою ймовірністю відображають реальну генетичну спорідненість.

По-третє, передача прізвищ не є виключно патрилійною. В Україні застосовується патрилійна модель передачі прізвищ, а винятки з цього правила є настільки рідкісними, що їх можна розглядати як своєрідні «мутації» в контексті аналогії з Y-хромосомою. Проте в глобальному масштабі спостерігаються різноманітні системи. Існують культури з матрилійною передачею прізвищ або родових імен, у яких ідентифікація відбувається по материнській лінії — наприклад, у народів мінангкабау (Індонезія), мосо/на (Китай), ашанті (Гана), а також у деяких корінних народів Північної Америки (навахо, хопі, ірокези) [17, 72, 75].

Також трапляються випадки успадкування материнського прізвища в патрилійних суспільствах, особливо в ситуаціях позашлюбного народження. У багатьох іспаномовних і португаломовних країнах застосовується система подвійних прізвищ, за якої нащадки успадковують прізвища від обох батьків, що не узгоджується з виключно патрилійною моделлю передачі Y-хромосоми.

До прізвищ як інструменту популяційно-генетичних досліджень висувається низка вимог, серед яких: стабільне успадкування протягом не менше ніж десяти поколінь, суттєве різноманіття, нерівномірний територіальний розподіл, кореляція з біологічними маркерами та здатність диференціювати окремі групи населення. Прізвищам українського народу притаманні всі зазначені характеристики, що робить їх придатними для використання в популяційно-генетичних дослідженнях [101].

Методологічна основа аналізу прізвищ сформувалася в другій половині XX століття завдяки працям Crow & Mange (1965) [21], які запропонували використовувати коефіцієнт інбридингу для оцінки генетичної структури популяцій за допомогою прізвищ. Подальший розвиток методу пов'язаний із роботами Lasker (1977) [43], який розробив показник генетичної відстані

на основі розподілу прізвищ, та Rodriguez-Larralde et al. (1998) [65, 66], які адаптували популяційно-генетичні моделі до аналізу ономастичних даних. Karlin & McGregor (1967) [36] запропонували індекс міграції, що кількісно характеризує інтенсивність потоку генів між популяціями.

1.2. Основні популяційно-генетичні показники, отримані з використанням прізвищ

1.2.1. Показники внутрішньопопуляційної структури

Показник внутрішньопопуляційної ізонімії (I^{jj}) кількісно визначає ймовірність того, що дві випадково обрані особи з популяції мають однакове прізвище [29]. Цей показник, що виражається в одиницях 10^{-4} , є фундаментальною характеристикою популяційної структури. Високі значення I^{jj} вказують на домінування обмеженої кількості прізвищ у популяції, що може бути наслідком ендогамії, ефекту засновника або географічної ізоляції. Натомість низькі значення свідчать про високе різноманіття прізвищ, типове для популяцій з інтенсивною міграцією та культурним змішуванням [29].

Коефіцієнт інбридингу (F) — це показник, який кількісно характеризує ймовірність того, що дві випадково вибрані особи, які є носіями одного й того самого прізвища, генетично споріднені [21]. Коефіцієнт інбридингу розраховується за евристичною формулою $F \approx I^{jj}/4$ [16, 21, 22, 43]. Ймовірність такого випадкового споріднення особливо висока, якщо популяція була заснована невеликою групою людей. Терміном «інбридинг» у людини умовно позначають спорідненість між особами. Оцінка цього показника має важливе прогностичне значення в медичній генетиці. Відомо, що кровна спорідненість (навіть випадкова) суттєво підвищує ризик появи у нащадків спадкових хвороб. Одні й ті самі рідкісні варіанти генів із набагато більшою ймовірністю присутні в осіб, які мають спільного предка, ніж у людей, які не перебувають у родинних стосунках [101]. Слід зазначити, що інтерпретація F , отриманого на основі прізвищ, має враховувати специфіку соціокультурних процесів, які можуть не повною мірою відображати генетичну реальність.

Показник різноманітності прізвищ Фішера (α), виражений в одиницях 10^2 , забезпечує оцінку ефективної кількості прізвищ у популяції, аналогічно до мір різноманітності в екології [27]. Високі значення α вказують на широке різноманіття прізвищ, що свідчить про значну історичну міграцію та змішування. Цей показник є менш чутливим до розміру вибірки порівняно з простим підрахунком кількості унікальних прізвищ і точніше відображає ефективне різноманіття популяції.

Індекс міграції Карліна — МакГрегора (ν), виражений в одиницях 10^{-4} , кількісно характеризує зміни в популяції під впливом міграційних процесів [36, 60, 83]. Цей показник ґрунтується на теорії нейтральної еволюції в скінченних популяціях сталого розміру. Оскільки фактичні мутації прізвищ трапляються рідко, високий рівень ν свідчить про значний приплив нових прізвищ унаслідок міграції.

1.2.2. Показники різноманітності розподілу

Ентропія Шеннона (H) [76] характеризує ступінь невизначеності або різноманіття в системі прізвищ. Домінування окремих прізвищ зменшує H , оскільки знижується невизначеність вибору. Показник надмірності (R), також відомий як дефіцит ентропії, вимірює відхилення реального розподілу від рівномірного. Високі значення R свідчать про те, що більшість населення носить порівняно невелику кількість прізвищ (асиметричний розподіл).

Ефективне число прізвищ може бути виражене через різні метрики: N_2 (зворотний індекс Сімпсона) [33, 78] відображає кількість однаково частих прізвищ, еквівалентну спостережуваному розподілу за ймовірністю збігу, тоді як N_1 (ентропійне ефективне число) ґрунтується на ентропії Шеннона. Ці показники є менш чутливими до рідкісних прізвищ порівняно з простим підрахунком унікальних варіантів і дають змогу коректно порівнювати популяції різного розміру.

Коефіцієнт покриття розширеного експоненціального розподілу ($CRSED$) є новітнім показником, що визначає пропорцію прізвищ у частині

розподілу, яка підпорядковується експоненціальному закону, порівняно зі степеневим [20]. Модель розтягнутого експоненціального розподілу охоплює як найпоширеніші прізвища (степеневий закон), так і рідкісні (експоненціальний зріз). Вищі значення *CRSED* вказують на більшу соціальну та географічну мобільність населення. На відміну від показників, які фокусуються на поширених прізвищах, *CRSED* ураховує рідкісні варіанти, що дає змогу виявити тонкі демографічні процеси.

Використання *CRSED* має кілька переваг:

1. Урахування рідкісних прізвищ: традиційна ізонімія фокусується на поширених прізвищах, тоді як *CRSED* виявляє внесок рідкісних, що є важливим для виявлення недавньої міграції;
2. Незалежність від розміру популяції: на відміну від α , *CRSED* є менш чутливим до N завдяки нормалізації на загальну кількість прізвищ.

1.2.3. Міжпопуляційні відстані

Відстань Ласкера (L) вимірює несхожість між популяціями на основі розподілу прізвищ шляхом логарифмічного перетворення міжпопуляційної ізонімії [43, 66]. Більше значення L означає меншу кількість спільних прізвищ і більші відмінності між популяціями. Оскільки рідкісні прізвища часто відображають історичні або сімейні зв'язки, цей показник дає змогу виявити малопомітні міграційні сигнали.

Евклідова відстань (D_E^{ij}) оцінює відмінності в розподілі прізвищ у геометричному просторі частот [19]. Значення, що дорівнює нулю, вказує на ідентичні розподіли, тоді як значення, близькі до одиниці, свідчать про значні відмінності.

Генетична відстань Нея (D_{NEI}^{ij}) [55, 57, 66] оцінює еволюційну та генетичну розбіжність, ураховуючи як міжпопуляційну, так і внутрішньо-популяційну ізонімію.

Хордова відстань Кавалі–Сфорца — Едвардса (D_{CSE}^{ij}) [19], вимірює генетичну віддаленість у багатовимірному просторі, де кожна популяція представлена точкою на поверхні сфери.

1.2.4. Просторовий аналіз

Модель ізоляції відстанню базується на припущенні, що чим далі одна від одної розташовані популяції, тим менше вони генетично схожі [37, 47]. Параметр B у моделі $\Phi_{ij}(x) = K \exp(-Bx)$ пов'язує інтенсивність змін прізвищ із дисперсію їхнього географічного переміщення. Тест Мантеля використовується для перевірки кореляції між матрицями генетичних і географічних відстаней шляхом послідовних перестановок даних [48, 80].

Індекс місця (IP_k) є кількісною мірою подібності розподілу прізвищ у конкретному регіоні із загальнонаціональним шаблоном [4, 99, 101, 102]. Кожному прізвищу присвоюється національний ранг за позицією у зведеному списку частот. Якщо значення (IP_k) є близьким до загальнонаціонального стандарту, це свідчить про низьку регіональну специфіку [4, 99, 101, 102]. Зростання відхилення від еталону характеризує регіон як більш своєрідний, із прізвищами, що на національному рівні займають нижчі позиції.

1.3. Популяційно-генетичні дослідження прізвищ у різних країнах світу

Для розуміння позиції України в загальному контексті світових популяцій необхідно проаналізувати накопичений досвід популяційно-генетичних досліджень прізвищ у різних регіонах світу. Нижче подано систематичний огляд основних закономірностей, виявлених у понад 20 країнах на п'яти континентах (табл. 1.1, рис. 1.1–1.5).

1.3.1. Європейські популяції

Австрія демонструє високий показник внутрішньогрупової ізонімії ($I^{jj} = 11,71 \times 10^{-4}$) та коефіцієнт інбридингу ($F = 29,27 \times 10^{-5}$) [7], що пояснюється географічною ізоляцією, зумовленою Альпами. Дослідження, проведене на основі даних 1 081 002 телефонних користувачів і 140 766 різних прізвищ,

виявило значну регіональну гетерогенність. Найнижче значення α було зафіксовано в Матрей (36), найвище — у Відні (3850), що відображає контраст між ізольованими гірськими громадами та космополітичними міськими центрами. Міські центри демонструють високий індекс міграції ($\nu = 7,89 \times 10^{-4}$) і більше різноманіття прізвищ ($\alpha = 8,54 \times 10^2$) [7].

Албанія, із 3 068 447 зареєстрованими виборцями та 37 184 унікальними прізвищами, демонструє показник ізонімії ($7,54 \times 10^{-4}$) та F ($18,84 \times 10^{-5}$) [51]. Ці показники відображають помірний рівень ізоляції серед європейських країн і свідчать про тривалу історію ізоляції, зумовлену гірською місцевістю та посилену стійкими клановими структурами (фісами). Низький показник різноманіття ($\alpha = 13,27 \times 10^2$) та індекс міграції ($\nu = 4,32 \times 10^{-4}$) вказують на обмежений приплив нових прізвищ і тривалу ендегамію [51].

Бельгія, досліджена на основі даних 1 118 004 осіб і 137 380 прізвищ, показує помірний показник I^{jj} ($2,15 \times 10^{-4}$) та F ($5,38 \times 10^{-5}$), високе різноманіття ($\alpha = 46,64 \times 10^2$) і значний індекс міграції ($\nu = 4,20 \times 10^{-4}$) [8]. На фонд прізвищ істотно вплинули три основні лінгвістичні групи — фламандська, французька та німецька. Помірне значення F ($5,38 \times 10^{-5}$) вказує на відповідний рівень генетичної диференціації, що відображає історію співіснування та конфліктів між мовними спільнотами [8].

Нідерланди, з даними 2 400 210 осіб і 126 485 прізвищ, демонструють відносно високий показник ізонімії ($4,21 \times 10^{-4}$) поряд зі значним індексом міграції ($\nu = 1,00 \times 10^{-4}$) [10]. Раннє запровадження обов'язкової реєстрації прізвищ під час французького правління створило стандартизовану систему називництва, що в окремих випадках призвело до високого рівня повторюваності прізвищ. Проте щільність населення та розгалужена торговельна мережа сприяли значній внутрішній міграції, що відображається в помірних значеннях α ($23,78 \times 10^2$) і F ($10,51 \times 10^{-5}$) [10].

Італія має один із найбагатших і найскладніших фондів прізвищ у Європі. На основі даних 5 043 580 осіб і 215 623 унікальних прізвищ отримано показники: $I^{jj} = 1,71 \times 10^{-4}$, $F = 4,27 \times 10^{-5}$, $\alpha = 58,55 \times 10^2$,

$\nu = 11,60 \times 10^{-4}$ [15]. Поліфілетична природа італійських прізвищ сягає корінням у фрагментарну середньовічну історію країни. Міські центри (Мілан, Рим, Турин) зазнали значної міграції та промислового зростання, тоді як багато сільських районів зберегли традиційні сімейні структури з обмеженим зовнішнім впливом [15].

Німеччина, досліджена на основі даних 5 150 310 осіб і 462 526 прізвищ, демонструє низький рівень ізонімії ($I^{jj} = 3,50 \times 10^{-4}$), велике різноманіття ($\alpha = 28,55 \times 10^2$) та помірний індекс міграції ($\nu = 5,54 \times 10^{-4}$) [65]. Міські центри — Гамбург, Берлін, Мюнхен — є осередками економічної та культурної активності, що приваблюють різноманітні групи населення. Значення F ($8,76 \times 10^{-5}$) вказує на помірну генетичну диференціацію [65]. У невеликих містах і сільських регіонах зберігається вищий рівень повторюваності прізвищ.

Франція, досліджена без урахування Корсики на основі 6 442 846 осіб і 495 104 прізвищ, демонструє помірні показники: $I^{jj} = 1,27 \times 10^{-4}$, $F = 3,17 \times 10^{-5}$, $\alpha = 78,77 \times 10^2$, $\nu = 0,02 \times 10^{-4}$ [73], що відображають вплив революційних змін і глобальної міграції. Мегаполіси (Париж, Ліон, Марсель) характеризуються високим рівнем внутрішньої та міжнародної міграції. У деяких сільських регіонах (Бретань, північно-східна Франція) зберігається підвищений рівень повторюваності прізвищ унаслідок локальних шлюбних практик [73].

Іспанія із системою подвійних прізвищ відображає складну історію, сформовану століттями мавританського впливу, Реконквісти та європейської імміграції. У базі даних батьківських прізвищ (90 150 прізвищ) α становить $2,16 \times 10^2$, F ($115,74 \times 10^{-5}$), ν ($0,59 \times 10^{-4}$). У наборі материнських прізвищ (106 723 прізвища) α становить $1,99 \times 10^2$, F ($125,63 \times 10^{-5}$), ν ($0,55 \times 10^{-4}$) [69]. Система подвійного прізвища функціонує як буфер проти екстремального інбридингу, подвоюючи кількість прізвищ, що передаються з покоління в покоління, і тим самим уповільнюючи втрату різноманітності та запобігаючи надмірній концентрації генетичного матеріалу в окремих лініях [69].

Швейцарія, досліджена на основі даних 1 702 287 телефонних абонентів і 166 116 прізвищ, демонструє показники: $I^{jj} = 4,18 \times 10^{-4}$, $F = 10,43 \times 10^{-5}$, $\alpha = 23,96 \times 10^2$, $\nu = 1,41 \times 10^{-4}$ [11, 66]. Багатомовний склад населення (німецька, французька, італійська, ретороманська мови) зумовив значні відмінності в розподілі прізвищ між кантонами. Міські центри (Женева, Цюріх) демонструють високе різноманіття, тоді як віддалені альпійські громади виявляють вищий рівень ізонімії внаслідок географічної ізоляції та традиційної ендогамії [11, 66].

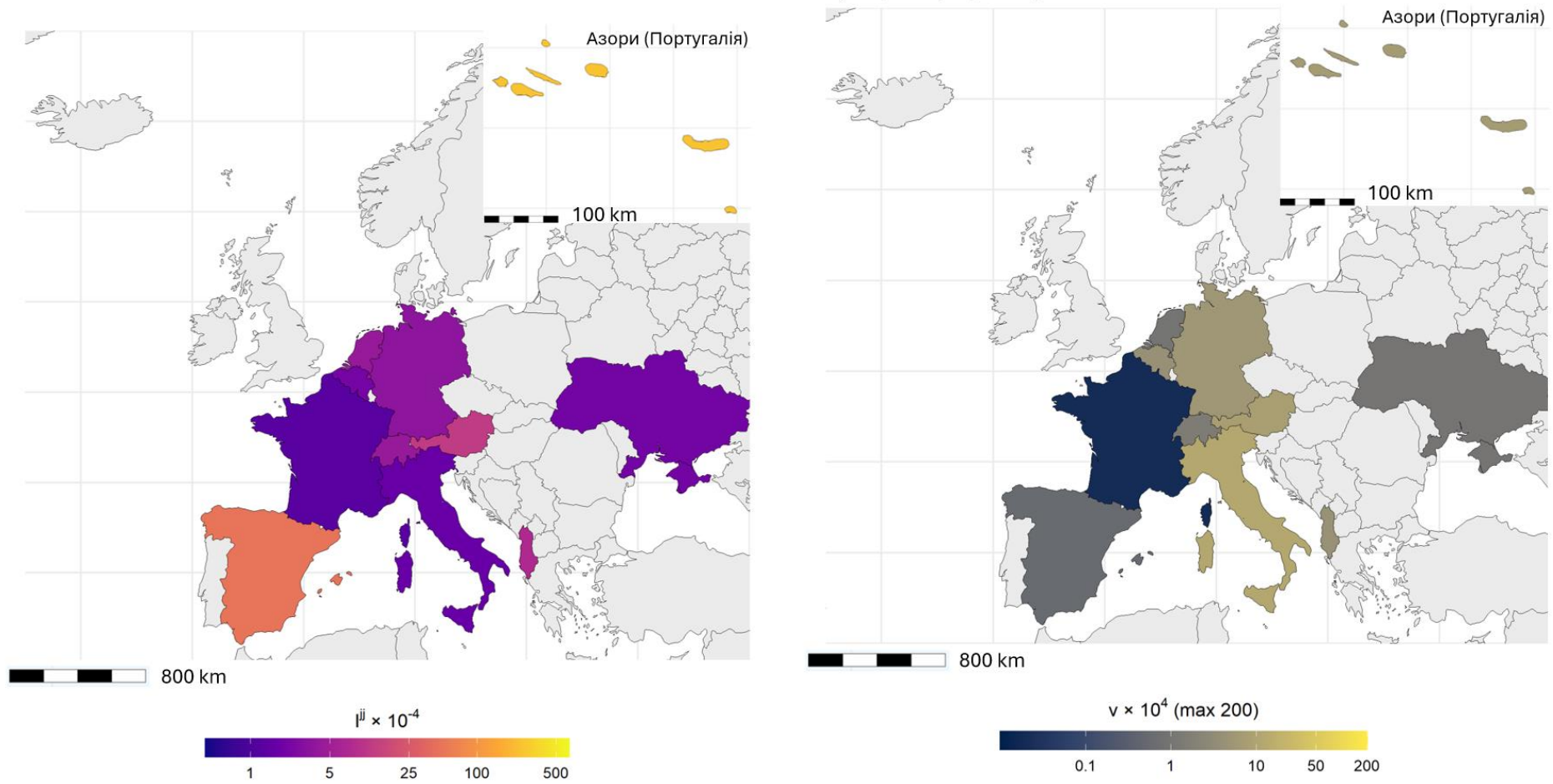


Рис. 1.1. Показники внутрішньопопуляційної ізонімії (I^{ij}) та ефективної міграції Карліна — МакҐрегора (v) у європейських популяціях. Дані наведено в табл. 1.1.

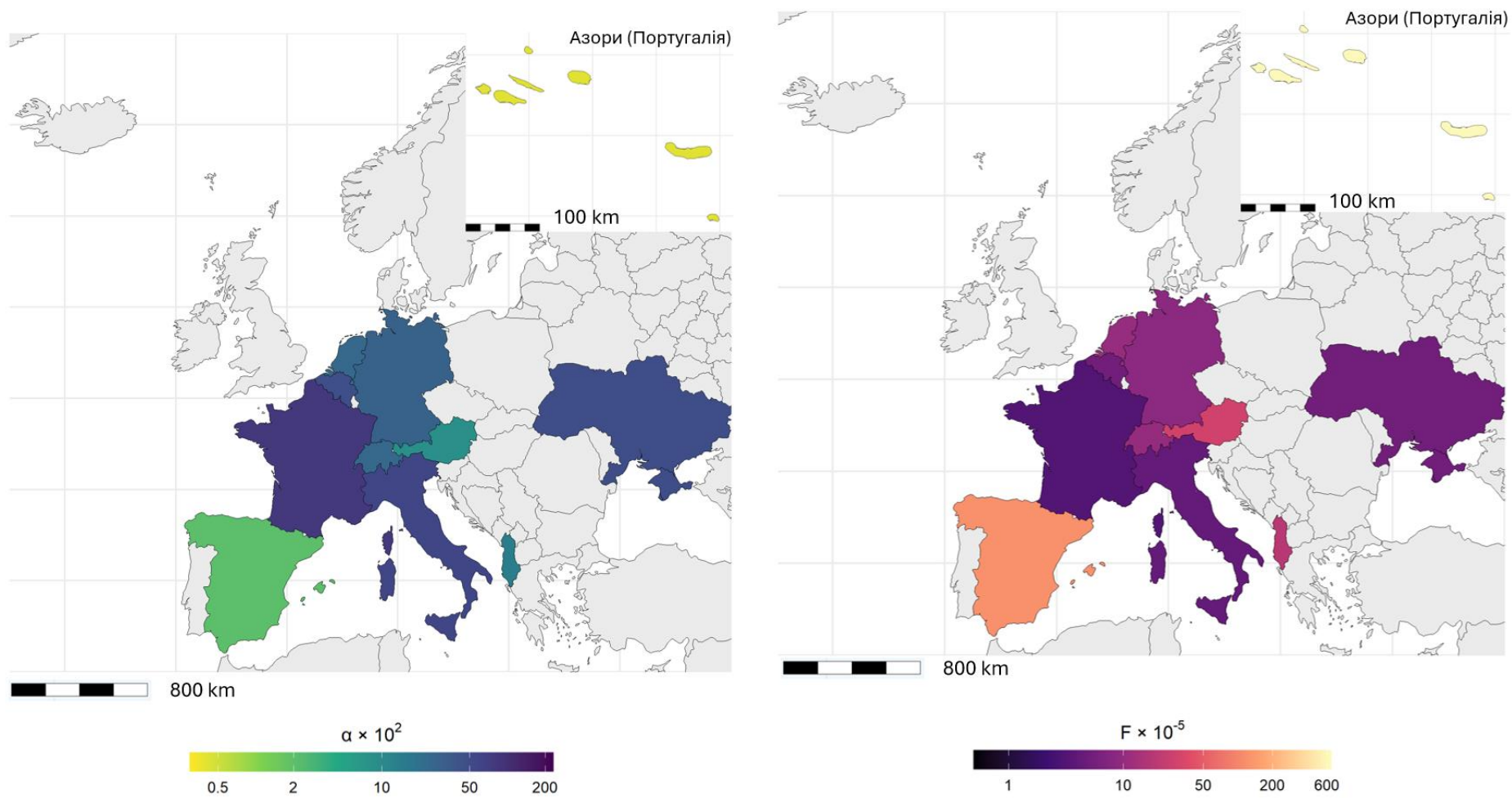


Рис. 1.2. Показники різноманітності прізвищ Фішера (α) та коефіцієнта інбридингу (F) у європейських популяціях. Дані наведено в табл. 1.1.

1.3.2. Острівні та ізольовані популяції

Острівні популяції становлять особливий інтерес для популяційної генетики, оскільки демонструють екстремальні значення показників унаслідок географічної ізоляції та обмеженої чисельності населення засновника.

Азорські острови, досліджені на основі даних 57 387 телефонних абонентів і 2 451 прізвища, демонструють надзвичайно високий рівень I^{jj} (243×10^{-4}) — найвищий серед усіх досліджених популяцій світу, — що відображає екстремальну концентрацію прізвищ, типову для ізольованих острівних популяцій із невеликою чисельністю населення-засновника. Низьке значення α ($0,41 \times 10^2$) та помірний індекс міграції ($\nu = 7 \times 10^{-4}$) вказують на те, що зовнішня міграція, переважно з континентальної Португалії, була достатньою для запобігання ще більшій концентрації прізвищ, яка спостерігалася б за умов повної ізоляції. Це демонструє складну взаємодію між ізоляцією та міграцією в острівних популяціях [15].

Республіка Саха (Якутія) є унікальним прикладом арктичної популяції з екстремальними кліматичними умовами. Дослідження 491 259 осіб (із 976 580 жителів) і 44 625 прізвищ виявило високі значення I^{jj} ($23,09 \times 10^{-4}$) та F ($57,74 \times 10^{-5}$) за низьких α ($4,33 \times 10^2$) і ν (8×10^{-4}) [83]. Високі значення показника ізонімії свідчать про те, що значна частина населення має однакові прізвища — це є наслідком довготривалого інбридингу та ефекту засновника. Історично населення Якутії не мало прізвищ і отримало їх після XVII століття від Російської імперії, що призвело до домінування прізвищ типу *Іванов* (10 824), *Васильєв* (6 151), *Слепцов* (5 153). Низькі рівні F спостерігаються у Якутському, Вілюйському та Мирнинському улусах (районах), тоді як найвищі — у поселеннях Арктичної зони (Верхоянський, Середньоколимський, Оленьоцький улуси) [83].

1.3.3. Азійські популяції

Японія демонструє надзвичайну стабільність у передачі прізвищ, що дає підстави припускати збереження структури успадкування протягом століть.

Аналіз проведено на основі кількох незалежних наборів даних, які ілюструють цю закономірність [88]. За набором даних А (715 815 осіб) отримано такі значення: $I^{jj} = 21,0 \times 10^{-4}$, $F = 52,5 \times 10^{-5}$, $\alpha = 4,76 \times 10^2$, $\nu = 6,64 \times 10^{-4}$ [88]. Набір даних В (2 647 011 осіб) показав $I^{jj} = 24,2 \times 10^{-4}$, $F = 60,50 \times 10^{-5}$, $\alpha = 4,12 \times 10^2$, $\nu = 1,56 \times 10^{-4}$ [88]. Дані Sakuma (350 000 осіб) дали $I^{jj} = 16,8 \times 10^{-4}$, $F = 42,0 \times 10^{-5}$, $\alpha = 5,95 \times 10^2$, $\nu = 16,98 \times 10^{-4}$ [88]. Аналіз файлів Tanaka включав чотири вибірки різного обсягу. За набором Tanaka 1 (53 682 особи): $I^{jj} = 17,7 \times 10^{-4}$, $F = 44,25 \times 10^{-5}$, $\alpha = 5,70 \times 10^2$, $\nu = 105,06 \times 10^{-4}$ [88]. Набір Tanaka 2 (62 492 особи) показав $I^{jj} = 23,2 \times 10^{-4}$, $F = 58,00 \times 10^{-5}$, $\alpha = 4,33 \times 10^2$, $\nu = 68,82 \times 10^{-4}$ [88]. Найменші вибірки — Tanaka 3 (3 424 особи) з показниками $I^{jj} = 22,9 \times 10^{-4}$, $F = 57,25 \times 10^{-5}$, $\alpha = 4,99 \times 10^2$, $\nu = 1272,81 \times 10^{-4}$ та Tanaka 4 (3 808 осіб) з $I^{jj} = 21,0 \times 10^{-4}$, $F = 52,50 \times 10^{-5}$, $\alpha = 5,43 \times 10^2$, $\nu = 1248,20 \times 10^{-4}$ [88].

Помірні рівні внутрішньої міграції та регіональної диференціації зумовили незначні коливання показників між вибірками. Це свідчить про те, що основна генетико-демографічна структура японської популяції залишається стабільною протягом тривалого періоду, тоді як локальна динаміка продовжує формувати регіональні особливості ономастичного ландшафту.

Китай є унікальним прикладом концентрації обмеженої кількості прізвищ серед надзвичайно чисельного населення. Дослідження, що охопило все населення країни — приблизно 1,277 млрд осіб і 7 327 унікальних прізвищ, — виявило такі показники: $I^{jj} = 260,0 \times 10^{-4}$, $F = 650,0 \times 10^{-5}$, $\alpha = 0,37 \times 10^2$ [45]. Індекс міграції становить $\nu = 0,2934 \times 10^{-4}$ [45], що відображає абсолютну кількість міграційних подій у масштабі всієї популяції; нормалізоване значення (приблизно 23 на 1 млн населення) вказує на помірно низьку інтенсивність припливу нових прізвищ. Високі значення I^{jj} та F відображають екстремальну концентрацію прізвищ.

Серед етнічних ханців використовується близько 3 000 прізвищ, однак лише 100 найпоширеніших охоплюють 85 % населення країни. До найчисельніших прізвищ належать Ван (王), Лі (李), Чжан (張), Лю (劉), Чжоу (周) та Чень (陳), які разом становлять значну частку населення. Така концентрація зумовлена історичними особливостями формування китайської системи прізвищ, коли обмежена кількість родових назв закріплювалася за великими кланами протягом тисячоліть, починаючи з епохи династії Хань (206 р. до н.е.–220 р. н.е.) [45].

Надзвичайно низький показник різноманітності ($\alpha = 0,37 \times 10^2$) [45] — вказує на обмежений ефективний фонд прізвищ відносно чисельності населення. Для порівняння: у Німеччині з населенням, приблизно у 40 разів меншим, α становить $28,55 \times 10^2$ [45], тобто майже у 77 разів вище. Коефіцієнт ізоляції за відстанню за Ласкером становить 0,64 [45], що свідчить про помірний вплив географічних бар'єрів на розподіл прізвищ.

Це пояснюється тим, що, незважаючи на значну територію, історична міграція відбувалася переважно в межах етнокультурних регіонів, зберігаючи регіональну специфіку розподілу навіть найпоширеніших прізвищ. Висока ізонімія та низька різноманітність відображають культурну традицію збереження родових прізвищ без суттєвих змін унаслідок міграцій або адміністративних реформ — на відміну від багатьох європейських популяцій, де русифікація, полонізація та інші соціально-політичні процеси призводили до масових змін прізвищ [45].

Локалізоване дослідження в місті Сяньтао на основі даних 223 327 осіб і 422 прізвищ виявило такі показники: $I^{jj} = 2,60 \times 10^{-4}$, $F = 6,50 \times 10^{-5}$, $\alpha = 38,49 \times 10^2$, $\nu = 172,30 \times 10^{-4}$ [91]. Чітко простежується так званий «ефект префектури»: міські центри мають тенденцію демонструвати більшу різноманітність прізвищ через вищі темпи міграції, тоді як сільські райони характеризуються меншою різноманітністю та вищою ізонімією [91].

Дослідження із використанням *CRSED* у Китаї охопило 7184 прізвища. Значення *CRSED* для всієї країни становило 0,040, що відповідає приблизно 288 поширеним прізвищам у розтягнутій експоненціальній категорії. У 362 префектурах спостерігалася значна варіативність: регіони з високим *CRSED* мають структуру прізвищ, подібну до інших територій через інтенсивну міграцію, тоді як регіони з низьким *CRSED* більшою мірою піддаються впливу генетичного дрейфу та умовної «мутації» прізвищ [20].

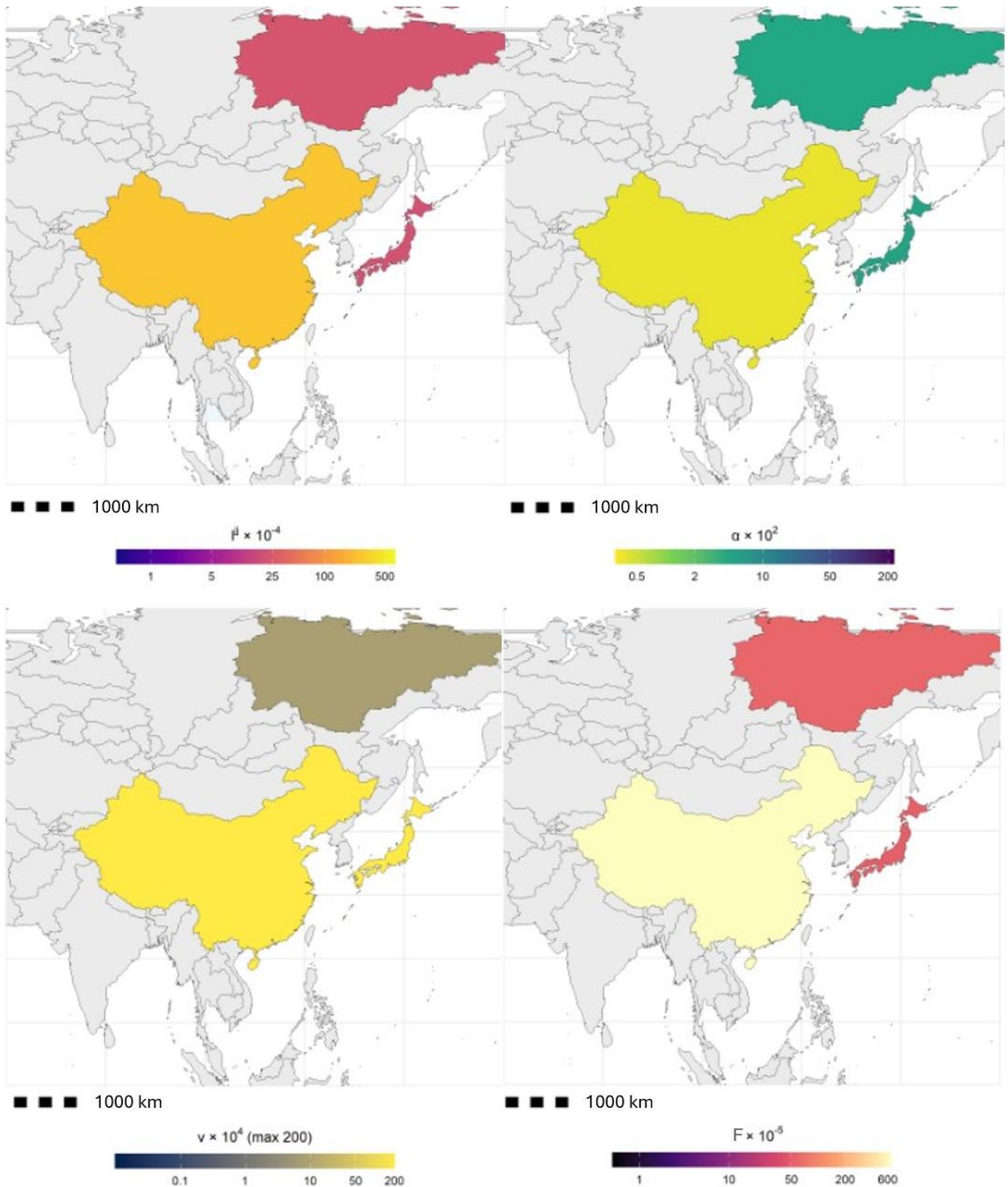


Рис. 1.3. Показники внутрішньопопуляційної ізонімії (I^{jj}), ефективної міграції Карліна — МакГрегора (v), різноманітності прізвищ Фішера (α) та коефіцієнта інбридингу (F) в азійських популяціях. Дані наведено в табл. 1.1.

1.3.4. Північноамериканські популяції

Сполучені Штати Америки, досліджені на основі понад 18 млн телефонних користувачів і 899 585 прізвищ із 247 міст у 48 суміжних штатах, демонструють такі показники: $I^{jj} = 7,32 \times 10^{-4}$, $F = 18,30 \times 10^{-5}$, $\alpha = 13,66 \times 10^2$,

$v = 0,76 \times 10^{-4}$ [12]. Відносно низький показник ізонімії вказує на високе генетичне різноманіття. Значний показник різноманітності прізвищ підкреслює вплив мультикультурного середовища. Найнижчий рівень різноманітності зафіксовано в місті Белфілд (318 прізвищ), найвищий — у Нью-Йорку (132 106 прізвищ) [12].

Техас, із набором даних 3 582 101 особа та 235 740 прізвищ, виявляє регіональні відмінності, особливо вздовж кордону ($I^{jj} = 8,13 \times 10^{-4}$, $F = 20,33 \times 10^{-5}$, $\alpha = 12,30 \times 10^2$, $v = 3,43 \times 10^{-4}$) [42], що відображає історичний іспанський колоніальний вплив і сучасні міграційні процеси. У регіонах типу Ларедо, де зберігаються традиційні сімейні структури, показник ізонімії є вищим. В урбанізованих центрах (Х'юстон, Даллас) інтенсивніші міграційні потоки зумовили більше розмаїття прізвищ [42].

1.3.5. Латиноамериканські популяції

Гондурас, з аналізом 4 348 021 особи та 14 662 різноманітних прізвищ демонструє відносно високі показники I^{jj} ($73,53 \times 10^{-4}$) та α ($1,36 \times 10^2$), але високі F ($183,82 \times 10^{-5}$) і v ($0,31 \times 10^{-4}$) [32]. Система подвійних прізвищ посилила генетичну однорідність на місцевому рівні завдяки збереженню розширених сімейних структур [32].

Аргентина на основі аналізу понад 22 млн. осіб і 414 441 прізвища демонструє такі показники: $I^{jj} = 23,70 \times 10^{-4}$, $F = 59,24 \times 10^{-5}$, $\alpha = 4,22 \times 10^2$, $v = 0,19 \times 10^{-4}$ [25]. Розподіл прізвищ відображає масовий приплив європейських іммігрантів наприкінці ХІХ — на початку ХХ століття, який змішався з генетичною спадщиною корінного населення країни. Помірні показники ізонімії та різноманітності свідчать про те, що, незважаючи на сильні регіональні традиції прізвищ, кумулятивний ефект широкомасштабної імміграції призвів до значної варіативності [25].

Болівія, досліджена на основі даних про 12 139 448 осіб і 174 922 унікальні прізвища (загалом 23 244 064 прізвища двобатьківського походження), демонструє дуже високі показники I^{jj} ($46,95 \times 10^{-4}$),

F ($117,37 \times 10^{-5}$), α ($2,11 \times 10^2$) та ν ($0,09 \times 10^{-4}$) [68]. Аналіз 50 найпоширеніших прізвищ виявив, що вони охоплюють 7 683 096 батьків, що дорівнює 30,05 % від загальної кількості батьківських прізвищ. Це вказує на відносну ізолюваність населення Болівії із сильними місцевими традиціями іменування, які зумовлюють високу концентрацію прізвищ і мінімальну міграцію [68].

Чилі з набором даних про 16 277 255 осіб і 72 667 прізвищ демонструє показники I^{jj} ($32,36 \times 10^{-4}$), F ($80,91 \times 10^{-5}$), α ($3,09 \times 10^2$) та ν ($0,19 \times 10^{-4}$) [14]. Проміжні значення свідчать про баланс між гомогенізуючим впливом іспанського колоніалізму та тривалою присутністю корінних груп (мапуче), чії особливі конвенції іменування сприяють локальному розмаїттю. Помірний показник міграції вказує на те, що, незважаючи на значні внутрішні переміщення із сільської місцевості до міських центрів, окремі регіони залишаються відносно ізолюваними [14].

Парагвай, досліджений на основі даних про 4 843 868 осіб і 39 021 прізвище, демонструє показники I^{jj} ($66,23 \times 10^{-4}$), F ($165,56 \times 10^{-5}$), α ($1,51 \times 10^2$) та ν ($0,31 \times 10^{-4}$) [24]. Високий показник ізонімії та низька різноманітність свідчать про сильну ендегамію й обмежений міграційний обмін між регіонами. Після колонізації багато громад у Парагваї залишалися відносно ізолюваними, зберігаючи обмежений пул прізвищ із незначними зовнішніми домішками [24].

Уругвай, аналіз якого базується на даних про 2 501 774 осіб і 70 395 прізвищ, демонструє показники I^{jj} ($33,11 \times 10^{-4}$), F ($82,78 \times 10^{-5}$), α ($3,02 \times 10^2$) та ν ($1,20 \times 10^{-4}$) [18]. Помірні показники свідчать про те, що європейська імміграція — переважно іспанська та італійська — сприяла формуванню різноманітного фонду прізвищ, проте регіональні відмінності зберігаються через локалізовану ендегамію та обмежений генний потік у певних регіонах [18].

Венесуела, досліджена на основі даних виборців 1991 року (3 905 487 осіб і 68 665 прізвищ), демонструє показники I^{jj} ($47,62 \times 10^{-4}$),

F ($119,05 \times 10^{-5}$), α ($2,10 \times 10^2$) та ν ($0,54 \times 10^{-4}$) [67]. Високий показник ізонімії та низький показник різноманітності свідчать про те, що, незважаючи на сучасну урбанізацію в окремих регіонах, зберігаються значні осередки ізоляції та локального генетичного дрейфу, особливо в географічно або політично ізольованих районах [67].

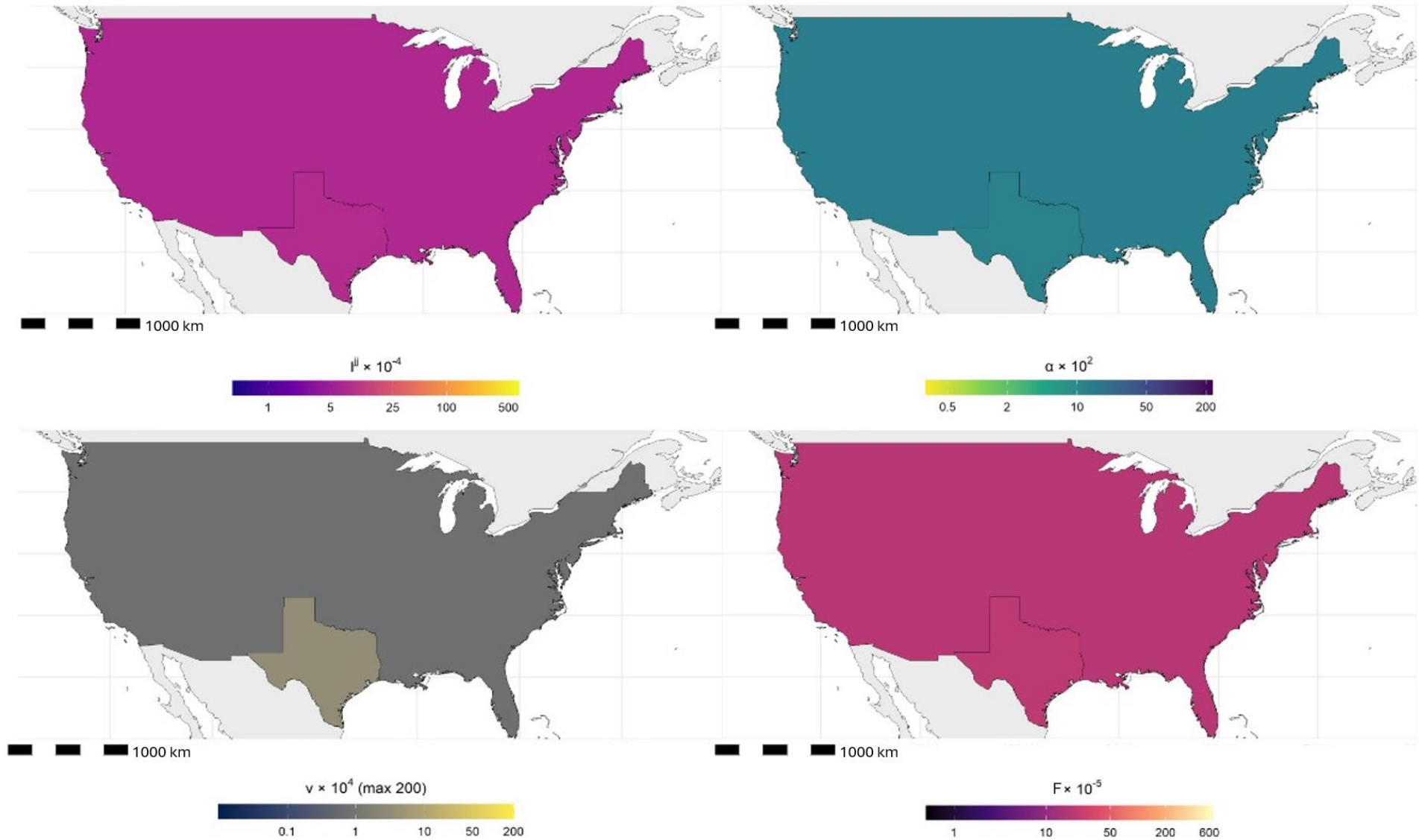


Рис. 1.4. Показники внутрішньопопуляційної ізонімії (I^j), ефективної міграції Карліна — МакГрегора (v), різноманітності прізвищ Фішера (α) та коефіцієнта інбридингу (F) у північноамериканських популяціях. Дані наведено в табл. 1.1.

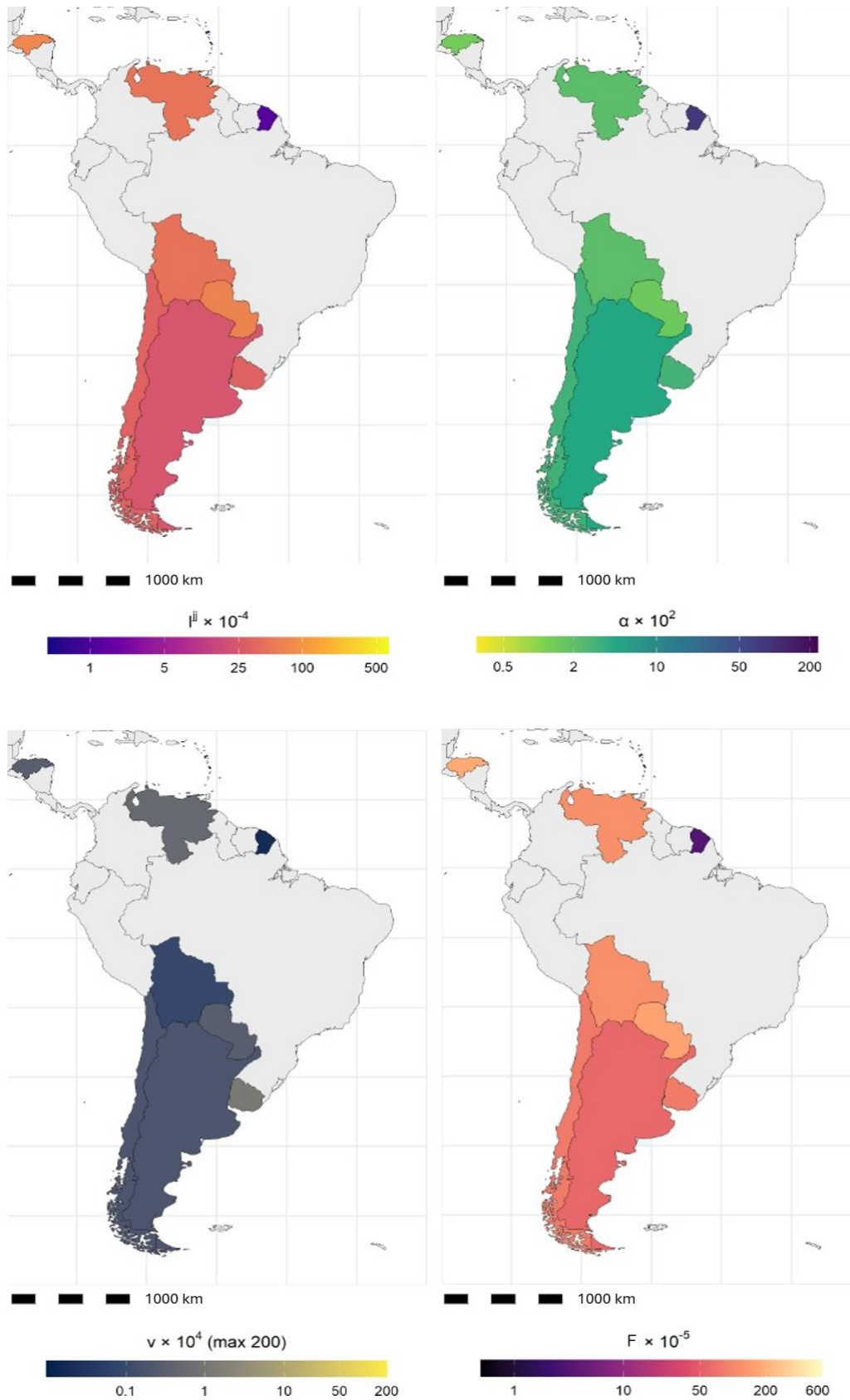


Рис. 1.5. Показники внутрішньопопуляційної ізонімії (I^j), ефективної міграції Карліна — МакГрегора (v), різноманітності прізвищ Фішера (α) та коефіцієнта інбридингу (F) у південноамериканських популяціях. Дані наведено в табл. 1.1.

Популяційні показники різних країн

Популяція	$I^{jj} \times 10^{-4}$	$F \times 10^{-5}$	$\alpha \times 10^2$	$\nu \times 10^{-4}$	r_L	r_E	r_N	N	k	Пос.
Європа										
Австрія	11,71	29,27	8,54	7,89	0,57	0,44	0,35	1 081 002	140 766	[7]
Албанія	7,54	18,84	13,27	4,32	0,62	0,71	0,56	3 068 447	37 184	[51]
Бельгія	2,15	5,38	46,46 (46,64)	4,20	0,72	-	-	1 118 004	137 380	[8]
Італія	1,71	4,27	58,55	11,60	0,63	0,57	0,51	5 043 580	215 623	[15]
Німеччина	3,50	8,76	28,55	5,54	0,51	0,48	0,51	5 150 310	462 526	[65]
Нідерланди	4,21	10,51	23,78	1,00	0,47	-	-	2 400 210	126 485	[10]
Франція	1,27	3,17	78,77 (78,86)	0,02	0,69	0,55	0,61	6 442 846	495 104	[73]
Іспанія	48,08	120,19	2,07	0,57	-	-	-	3 625 688	1 770 790	[69]
Подвійна	0,45	1,13	220,77	60,89	-	-	-	-	1 573 917	[69]
Батьківська	46,29	115,74	2,16	0,59	0,18	0,21	0,03	-	90 150	[69]
Материнська	50,10	125,63	1,99	0,55	0,13	0,26	0,08	-	106 723	[69]
Швейцарія	4,18	10,43	23,96	1,41	0,72	0,34	0,45	1 702 287	166 116	[11, 66]
Азори	243,0	390 (607,50)	0,41	7,00	-	-	0,73	57 387	2 451	[15]
Азія										
Японія (А)	21,0	52,50	4,76	6,64	-	-	-	715 815	-	[88]
Японія (В)	24,2	60,50	4,12	1,56	-	-	-	2 647 011	-	[88]
Японія (S)	16,8	42,00	5,95	16,98	-	-	-	350 000	4000	[88]
Японія (Т1)	17,7	44,25	5,70	105,06	-	-	-	53 682	-	[88]
Японія (Т2)	23,2	58,00	4,33	68,82	-	-	-	62 492	-	[88]
Японія (Т3)	22,9	57,25	4,99	1272,81	-	-	-	3 424	-	[88]
Японія (Т4)	21,0	52,50	5,43	1248,20	-	-	-	3 808	-	[88]
Китай*	260,0	650	0,37	0,2934	0,64	0,67	0,52	12768 x 10 ⁶	7 327	[45]
Сянтао	2,60	6,50	38,49	172,30	0,05	0,18	0,05	223 327	422	[91]
Якутія	23,09	57,74	4,33	8,00	0,51	0,63	0,69	491 259	44 625	[83]
Північна Америка										
США	7,32	18,30	13,66	0,76	0,21	-	0,24	18 022 000	899 585	[12]
Техас	8,13	20,33	12,30	3,43	0,09	0,49	0,34	3 582 101	235 740	[42]

Примітки: I^{jj} — показник ізонімії, F — компонента інбридингу, α — показник різноманітності прізвищ, ν — індекс міграції, r_L — ізоляція відстанню за Ласкером, r_E — ізоляція відстанню за Евклідовими відстанями, r_N — ізоляція відстанню за Неєм, N — чисельність вибірки, k — кількість прізвищ, Пос. — посилання, * — тотально досліджене населення, 1Т та 2Т — Файли Такака, S-файли Sakuma. Курсивом позначені статистики, розраховані В.В. Шубою.

Популяційні показники різних країн

Популяція	$I^{jj} \times 10^{-4}$	$F \times 10^{-5}$	$\alpha \times 10^2$	$\nu \times 10^{-4}$	r_L	r_E	r_N	N	k	Пос.
Латинська Америка										
Гондурас	73,53	183,82	1,36	0,31	0,65	0,63	0,55	4 348 021	14 662	[32]
Аргентина	23,70	59,24	4,22	0,19	-	-	-	22,6 x 10 ⁶	414 441	[25]
Болівія	46,95	117,37	2,11	0,09	0,55	0,55	0,50	23 244 064	174 922	[68]
Чілі	32,36	80,91	3,09	0,19	0,48	0,57	0,45	16 277 255	72 667	[14]
Парагвай	66,23	165,56	1,51	0,31	0,58	0,71	0,60	4 843 868	39 021	[24]
Уругвай	33,11	82,78	3,02	1,20	0,17	0,46	0,57	2 501 774	70 395	[18]
Венесуела	47,62	119,05	2,10	0,54	0,35	0,78	-	3 905 487	68 665	[67]

Примітки: I^{jj} — показник ізонімії, F — компонента інбридингу, α — показник різноманітності прізвищ, ν — індекс міграції, r_L — ізоляція відстанню за Ласкером, r_E — ізоляція відстанню за Евклідовими відстанями, r_N — ізоляція відстанню за Несм, N — чисельність вибірки, k — кількість прізвищ, Пос. — посилання, * — тотально досліджене населення, 1Т та 2Т — Файли Tanaka, S-файли Sakuma. Курсивом позначені статистики, розраховані В.В. Шубою.

1.4. Вплив історичних подій, географічних умов і культурних особливостей на структуру популяцій

1.4.1. Історичні чинники

Історичні події глибоко вплинули на динаміку популяцій і розподіл прізвищ. Північне Причорномор'я, що входить до складу сучасної України, було перехрестям міграційних шляхів, які з'єднували великий Євразійський степ із Центральною Європою. Географічне положення України зробило її місцем перетину східних та західних сусідів, взаємодія яких залишила слід у генетичному складі місцевих популяцій і відбилася в антропологічному обліку населення [1, 28, 31, 49, 58, 61, 71, 110].

Упродовж століть українські території перебували у складі різних державних утворень — Литовського князівства, Речі Посполитої, Польщі, Османської імперії, Австро-Угорщини, Московії, потім Російської імперії, — що зумовило їхню етнічну та культурну різноманітність [46, 82, 93]. У 1922 році Українська республіка стала частиною Радянського Союзу, однак її кордони змінювалися протягом усього ХХ століття. У серпні 1991 року Україна

набула статусу суверенної держави, успадкувавши території, що остаточно сформувалися після Другої світової війни [46, 82, 93].

Історичні події ХХ століття мали катастрофічний вплив на структуру українського населення й потенційно суттєво змінили розподіл прізвищ:

1. Голодомор 1932–1933 років [46, 82, 93].

Штучний голод, організований радянською владою, призвів до загибелі від 3,5 до 7 млн осіб, переважно в центральних і східних областях України (Київській, Полтавській, Харківській, Дніпропетровській, Одеській). Спустошені території заселяли мігрантами з Росії, Білорусі та інших регіонів СРСР. Це могло призвести до таких наслідків: 1) ефекту пляшкового горла — різке скорочення чисельності населення зумовило зменшення різноманітності прізвищ у найбільш постраждалих регіонах; 2) привнесення російських прізвищ — переселенці принесли з собою типові російські прізвища (*Іванов, Петров, Сидоров, Кузнєцов*), що збільшило їхню частку в загальному фонді; 3) порушення історичної наступності — традиційна популяційна структура, сформована протягом століть, була частково або повністю знищена в окремих районах; 4) Зміщення в бік молодших когорт — нові поселенці адаптували свої прізвища відповідно до радянських норм стандартизації.

2. Друга світова війна (1941–1945) [46, 82, 93].

Загибель 5-8 млн українців, масові депортації, примусові переміщення населення та знищення цілих населених пунктів суттєво вплинули на демографічну структуру країни. Особливо постраждали: 1) західні області (Волинь, Галичина) — партизанська війна, масові каральні операції, депортації; 2) Донбас і промислові центри — тривала окупація, примусова праця в Німеччині, руйнування міст; 3) південні регіони — окупація, Голокост (знищення єврейських громад з їхніми специфічними прізвищами).

3. Післявоєнні репресії та депортації (1944–1953) [46, 82, 93].

1) Депортація кримських татар (1944) — повне виселення понад 200 тис. осіб із Криму, що призвело до зникнення цілого пласту кримськотатарських прізвищ. 2) Боротьба з ОУН-УПА на заході України — депортації сімей

повстанців до Сибіру та Казахстану. 3) Переселення поляків (1944–1946) — обмін населенням між Польщею та УРСР, депортація близько 800 тис. поляків. 4) Виселення німців — депортація етнічних німців з України.

4. Індустріалізація та урбанізація (1950–1980) [46, 82, 93].

Масовий приплив робітників із Росії, Білорусі та інших республік СРСР до промислових центрів зумовив суттєві демографічні зміни, зокрема:

- 1) Донбас — будівництво шахт і заводів залучило мільйони мігрантів;
- 2) Придніпров'я — розвиток металургії та машинобудування;
- 3) Харків, Дніпро, Запоріжжя — формування багатонаціональних промислових міст;
- 4) Київ — розширення столиці, приплив кадрів з усього СРСР.

Розташування України як історичного перехрестя зумовило формування неоднорідного пулу прізвищ. В Австрії століття Габсбурзької монархії в поєднанні з фізичною ізоляцією, зумовленою Альпами, сприяли формуванню локалізованих генетичних кластерів (високі показники I^{jj} та F у сільській місцевості) поряд із міською різноманітністю, спричиненою міграцією індустріальної епохи [7].

Сталі кланові структури Албанії — фіси, зміцнені під час османського правління, — зумовили високий рівень повторюваності прізвищ та інбридингу [51]. Тривала історія Італії — від Стародавнього Риму через середньовічну фрагментацію до сучасного об'єднання — створила поліфілетичний ландшафт прізвищ із виразними регіональними відмінностями [15]. У Франції та Іспанії періоди революцій, колонізації та подальші хвилі імміграції залишили глибокі сліди в генетичній структурі [69, 73]. Система подвійних прізвищ в Іспанії зберігає лінії як по батьківській, так і по материнській лінії, забезпечуючи двобатьківське простеження родоводу [69].

1.4.2. Географічні чинники

Географічні об'єкти — гори, річки, болота та берегові лінії — історично визначали моделі розселення й міграційні маршрути. Великі рівнини України

та судноплавні річки сприяли міжрегіональному змішуванню, знижуючи показники ізонімії й інбридингу. Відкритий ландшафт полегшував міграційні рухи як у напрямку Схід–Захід (з Азії до Європи), так і в напрямку Північ–Південь (від лісової зони до Причорномор'я).

Однак Карпатський регіон представляє контраст: гірські хребти історично обмежували мобільність, створюючи ізольовані гірські долини з власними мікропопуляціями. Це має проявитися в підвищених значеннях I^{jj} та F для Закарпатської, Чернівецької та частково Івано-Франківської областей — аналогічно до ситуації в Альпах (Австрія, Швейцарія) [101].

В Австрії альпійські хребти обмежили мобільність у багатьох сільських районах, що призводить до вищих I^{jj} та F , тоді як міські центри демонструють підвищену різноманітність завдяки кращій транспортній доступності [7]. Пересічена місцевість Албанії також обмежує міграцію, зберігаючи концентровані кластери прізвищ [51]. Рівнинні й доступні ландшафти Бельгії та Нідерландів сприяли широкому переміщенню населення, що відображається в їхніх помірних і високих показниках міграції [8, 10]. В Італії та Німеччині різноманітні ландшафти зумовили регіональну диференціацію: гірські й сільські регіони демонструють вищий рівень генетичної ізоляції порівняно з міськими центрами [15, 65].

Острівні місця, такі як Азорські острови, природно обмежують потік генів, підсилюючи ефект засновника [15]. Різноманітний рельєф Японії в поєднанні з періодом історичної ізоляції зберігає стабільну, але регіонально диференційовану систему прізвищ [88]. Попри великі розміри країни та можливості для міграції, в Китаї історично склалася висока концентрація кількох домінантних прізвищ — завдяки централізованому управлінню та сильній культурній традиції [45].

Суворі арктичні умови Якутії наклали серйозні обмеження на міграцію, що призвело до високих значень I^{jj} та F [83]. У Північній Америці різноманітний ландшафт Сполучених Штатів — від відкритих рівнин до густих міських агломерацій — сприяв масштабній глобальній міграції,

що зумовило велике розмаїття прізвищ [12]. Техас відображає локалізовану історичну спадщину вздовж своїх прикордонних регіонів із чіткою диференціацією між прикордонними зонами та міськими центрами [42].

1.4.3. Культурні чинники

Культурні звичаї та соціальні норми відіграють вирішальну роль у формуванні структури прізвищ. Система подвійних прізвищ в Іспанії та Гондурасі зберігає материнську лінію і водночас сприяє більшому генетичному різноманіттю, знижуючи ризик ізоляції окремих родових ліній [32, 69]. Натомість суспільства із сильними традиціями ендогамії або вираженою клановою організацією, такі як Албанія та Якутія, демонструють вищу повторюваність прізвищ і підвищені показники інбридингу [51, 83].

Адміністративна політика також суттєво впливає на розподіл прізвищ. Раннє впровадження обов'язкової реєстрації прізвищ у Нідерландах під час французького правління створило стандартизовану систему іменування, що в окремих випадках призвело до високого рівня повторюваності прізвищ [10]. У Якутії населення отримало прізвища після XVII століття під впливом Російської імперії, що спричинило домінування типових російських прізвищ (*Іванов, Васильєв, Петров*) [83].

Релігійні та етнічні традиції також формують ономастичний ландшафт. У багатьох ісламських суспільствах традиції іменування відрізняються від європейських, що може впливати на розподіл прізвищ [74]. Єврейські громади зберігали специфічні традиції прізвищ, що відрізняли їх від навколишнього населення [64]. Корінні народи Америки часто мали власні системи іменування, які були трансформовані або замінені під час колонізації [12].

1.5. Популяційно-генетичні дослідження прізвищ в Україні

1.5.1. Молекулярно-генетичні дослідження

На початку XXI століття в Україні розпочалися систематичні дослідження генетичної структури населення з використанням молекулярних

маркерів ДНК. Такі наукові роботи передбачали польові виїзди до різних популяцій для збирання біологічного матеріалу. Найбільш масштабним став міжнародний проєкт «Генографія» (2005–2011), у якому брали участь науковці з Харкова. Фінансову підтримку цієї ініціативи надавали видавнича компанія журналу «National Geographic», корпорація IBM та благодійна організація родини Wait. Під час наукових експедицій по всій території держави було сформовано банк зразків генетичного матеріалу з детальною документацією етнічного походження учасників дослідження [101, 112–120].

Після завершення міжнародного проєкту наукова діяльність продовжувалася за рахунок бюджетних коштів, проте їхній обсяг значно поступався попередньому зовнішньому фінансуванню. Через брак фінансування дослідники звернулися до альтернативних методів. Аналіз прізвищ виявився доступним і економічно вигідним інструментом для продовження популяційно-генетичних досліджень [101].

Починаючи з 2014 року, події на сході України спричинили масштабні міграційні процеси, що призвели до депопуляції регіону. Із 24 лютого 2022 року міграційні процеси набули масового й катастрофічного характеру, значно посиливши негативні демографічні тенденції. Наслідки цих процесів доведеться оцінювати популяційним антропогенетикам, оскільки сучасні події неминуче вплинуть на структуру популяцій та їхню адаптивність. Після стабілізації ситуації масштаби змін можна буде оперативно оцінити за допомогою аналізу прізвищ, порівнявши нові показники з даними 2001 року, отриманими в межах цього дослідження. Локальності з найбільш вираженими відхиленнями стануть пріоритетними об'єктами для детального вивчення з використанням класичних та ДНК-маркерів.

Базовою одиницею часу в популяційній генетиці є тривалість покоління людини, що становить приблизно 20 років. З огляду на часовий інтервал від 2001 року нині настав період, коли можна кількісно оцінити генетичні трансформації, що відбулися в українських популяціях протягом одного покоління.

1.5.2. Перші дослідження прізвищ

Використання прізвищ як інструменту в популяційно-генетичних дослідженнях вперше запропонував Герман Джозеф Мюллер в особистому листуванні з Джеймсом Кроу [21]. У 1965 році Кроу, ґрунтуючись на принципах Мюллера, працях Сьюалла Райта та формулі, запропонованій Гюставом Малєко (1948), описав розрахунок показника інбридингу для популяції людей: $F = I^{jj}/4$ [16, 21, 22]. Пізніше, у 1967 році, Ясуда і Мортон запропонували формулу для розрахунку ізонімії [89].

1.5.3. Основні результати попередніх досліджень

Попередні дослідження показали, що Україна характеризується відносно низькими показниками внутрішньогрупової ізонімії ($I^{jj} = 2,1 \times 10^{-4}$) та коефіцієнта інбридингу ($F = 5,3 \times 10^{-5}$), що свідчить про значну різноманітність прізвищ і високий рівень історичного змішування [101]. Ці значення суттєво нижчі порівняно з географічно ізольованими популяціями та відображають позицію України як історичного перехрестя міграційних шляхів.

Загальна кількість унікальних прізвищ українського населення становить 697 147, що є другим показником у світі (після Сполучених Штатів) та найвищим серед європейських країн, для яких наразі проведено відповідні дослідження. Це узгоджується з максимальним рівнем генетичного різноманіття за маркерами Y-хромосоми, виявленим в українців порівняно з іншими народами Європи [101].

Було встановлено, що між часткою етнічних прізвищ серед 100 найпоширеніших і питомою вагою етносу існує прямий зв'язок, що відображено у коефіцієнтах кореляції: для українських прізвищ $r_s = 0,703$, для російських — $r_s = 0,739$, для «інших» — $r_s = 0,393$ [101]. Індекс ізонімії на рівні областей демонструє чітке зростання в напрямку з південного сходу на північний захід, що знаходить логічне пояснення в ландшафті країни та напрямках демографічних рухів, що відбувалися в історичний період [101].

1.5.4. Валідація методу ізонімії через молекулярно-генетичні дані

Для підтвердження достовірності методу ізонімії як непрямого генетичного маркера вкрай важливо зіставити результати, отримані на основі прізвищ, із даними прямого молекулярно-генетичного аналізу. Дослідження варіабельності Y-хромосоми в Україні [6, 40, 90, 112, 114, 116, 117, 121] слугують підґрунтям для таких порівнянь.

Коефіцієнт кореляції між матрицями відстаней, визначених за прізвищами українського населення, та матрицями, побудованими за Y-хромосомними маркерами становить $r = 0,58$, тоді як кореляція з матрицями відстаней, розрахованими за групами крові, сягає $r = 0,60$. Ці результати свідчать про те, що українські прізвища загалом здатні досить точно відобразити популяційну структуру та відмінності між різними групами населення [101].

1.6. Порівняльний аналіз популяційно-генетичних показників

Порівняльний аналіз популяційно-генетичних показників різних країн виявляє кілька фундаментальних закономірностей (Рис. 1.6). По-перше, взаємодія між географічною ізоляцією та міграцією є основним чинником, що визначає розмаїття прізвищ. Регіони, позначені природними бар'єрами або ізолюваністю, як правило, демонструють вищий рівень ізонімії (I^{jj}) та меншу різноманітність прізвищ (α), як це спостерігається в Албанії ($I^{jj} = 7,54 \times 10^{-4}$, $\alpha = 13,27 \times 10^2$) [51], Якутії ($I^{jj} = 23,09 \times 10^{-4}$, $\alpha = 4,33 \times 10^2$) [83], та Парагваї ($I^{jj} = 66,23 \times 10^{-4}$, $\alpha = 1,51 \times 10^2$) [24].

Натомість регіони, які зазнали значної міграції та культурного обміну — часто внаслідок урбанізації й економічного розвитку, — мають тенденцію до нижчого рівня ізонімії та більшого розмаїття прізвищ. Німеччина ($I^{jj} = 3,50 \times 10^{-4}$, $\alpha = 28,55 \times 10^2$) [65], Франція ($I^{jj} = 1,27 \times 10^{-4}$, $\alpha = 78,77 \times 10^2$) [73], та Сполучені Штати ($I^{jj} = 7,32 \times 10^{-4}$, $\alpha = 13,66 \times 10^2$) [12] демонструють цю закономірність.

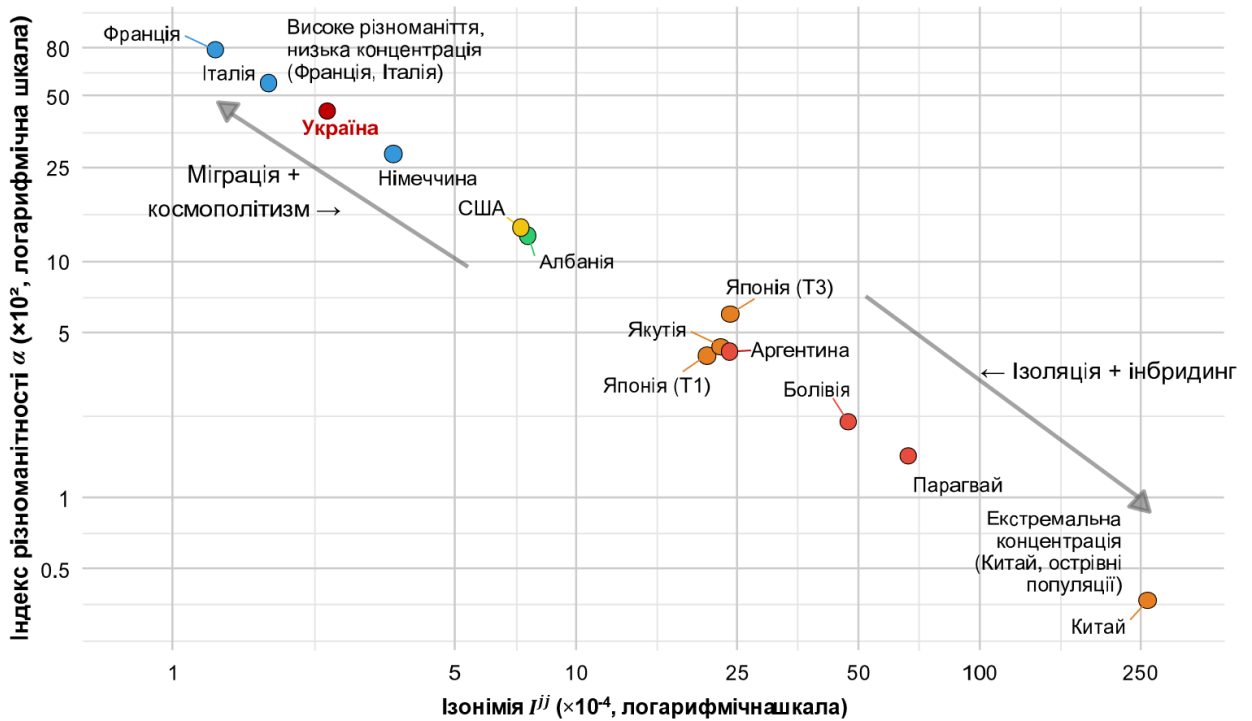


Рис. 1.6. Позичування України в системі координат «ізонімія — різноманітність» ($n = 13$ популяцій). Україна ($I^{jj} = 2,1 \times 10^{-4}$, $\alpha = 47,0 \times 10^2$) займає проміжне положення, що відображає баланс між історичною ізоляцією окремих регіонів (Карпати) та інтенсивним змішуванням у степових зонах. Логарифмічні шкали дають змогу відобразити весь діапазон — від ізольованих популяцій (Азори, Китай) до космополітичних (Франція, Італія).

По-друге, показники ізонімії (I^{jj}), коефіцієнт інбридингу (F), різноманітність прізвищ (α) та показник міграції (ν) є надійними вимірюваними індикаторами основних демографічних процесів. Високі значення I^{jj} та F зазвичай корелюють із популяціями, які пережили тривалу ізоляцію та ендогамію. Високі значення α та ν свідчать про динамічні міграційні процеси й постійний приплив нових прізвищ.

Існує чітка негативна кореляція між показником ізонімії та індексом різноманітності: популяції з високим I^{jj} демонструють низький α , і навпаки. Це підтверджує теоретичні очікування, оскільки висока концентрація окремих прізвищ (висока ізонімія) закономірно супроводжується низькою загальною різноманітністю. Аналогічна закономірність спостерігається між F та ν : ізольовані популяції з високим інбридингом (F) характеризуються низьким рівнем міграції (ν).

По-третє, коефіцієнти ізоляції за відстанню (r) варіюють залежно від географії та історії регіону. Високі значення r спостерігаються в Європі (Бельгія — $r = 0,72$ [8], Франція — $r = 0,55-0,69$ [73], Швейцарія — $r = 0,34-0,72$ [11, 66]), що відображає стабільні історичні поселенські структури та обмежену міграцію на великі відстані в доіндустріальну епоху. Нижчі значення r характерні для країн Нового Світу (США — $r = 0,21-0,24$ [12]), де масова міграція та мобільність населення послабили географічну залежність генетичної структури.

1.7. Методологічні обмеження та перспективи досліджень

Використання прізвищ як генетичних маркерів має певні методологічні обмеження, які необхідно враховувати під час інтерпретації результатів. Як зазначалося в розділі 1.1, метод ізонімії базується на припущенні про патрилійну передачу прізвищ та їхню стабільність у часі. Однак, крім фундаментальних обмежень (русифікація, конвергентне виникнення, непатрилійна передача), слід враховувати й специфічні джерела помилок.

По-перше, це соціально-демографічні чинники: незміна прізвища під час шлюбу, коли частина жінок зберігала дівоче прізвище або утворювала подвійні прізвища; добровільна зміна з соціальних мотивів — у першій половині ХХ століття відбувалася зміна прізвищ для уникнення дискримінації; помилки перепису та транслітерації (наприклад, *Сидоренко/Сидоров, Кьося/Кіося*).

По-друге, конвергентне виникнення прізвищ створює додаткову складність щодо високочастотних прізвищ. Поширені професійні прізвища (наприклад, *Коваль, Мельник, Кравець*) могли виникати незалежно в різних регіонах, оскільки відображали типові заняття. Це означає, що носії однакових високочастотних прізвищ не обов'язково мають спільного предка в межах досліджуваного періоду, що потенційно призводить до завищення коефіцієнтів спорідненості між парами районів із високою частотою таких прізвищ. Водночас рідкісні прізвища з більшою ймовірністю відображають

справжню генетичну спорідненість і є найбільш надійними маркерами для філогенетичних реконструкцій.

Незважаючи на ці обмеження, метод ізонімії залишається оптимальним інструментом для популяційно-генетичних досліджень у великих територіальних масштабах та в історичній перспективі. Основні переваги методу порівняно з молекулярно-генетичними підходами включають повнопопуляційне охоплення (аналіз всього населення проти незначної вибірки в типових ДНК-дослідженнях, що усуває вибіркоче зміщення), ретроспективність (можливість аналізу історичних даних без потреби в біологічних зразках), економічну ефективність та доступність даних з адміністративних джерел.

Перспективи подальших досліджень включають три основні напрямки. По-перше, валідацію через інтеграцію з молекулярною генетикою: збір Y-хромосомних даних для репрезентативної вибірки з усіх районів України з метою розрахунку поправочних коефіцієнтів, що враховують регіональну варіабельність точності методу ізонімії. Це дасть змогу оцінити кореляцію між показниками генетичної диференціації за прізвищами та молекулярними маркерами окремо для макрорегіонів із різною історією. По-друге, просторова деталізація: детальний аналіз на рівні об'єднаних територіальних громад для виявлення мікроструктури популяцій, особливо в гірських регіонах, де очікується підвищена генетична диференціація через географічну ізоляцію. По-третє, часова динаміка: порівняльний аналіз історичних переписів для реконструкції міграційних процесів із застосуванням байєсівського демографічного моделювання (Bayesian Skyline Plot). Особливу увагу буде приділено оцінюванню впливу катастрофічних подій ХХ століття на генетичну структуру регіональних популяцій.

Таким чином, метод ізонімії, незважаючи на методологічні обмеження, залишається незамінним інструментом для повнопопуляційного аналізу генетичної структури на великих часових масштабах, ефективно доповнюючи

молекулярно-генетичні методи у вивченні мікроеволюційних процесів та демографічної історії України.

УЗАГАЛЬНЕННЯ ТА ВИСНОВКИ ДО РОЗДІЛУ 1

Огляд світової літератури з популяційно-генетичних досліджень прізвищ виявляє кілька фундаментальних закономірностей, які мають важливе значення для розуміння структури популяцій і можуть бути застосовані для аналізу українського населення.

Взаємодія географічної ізоляції та міграції є основною детермінантою розмаїття прізвищ і генетичної структури популяцій. Природні бар'єри (гори, острови, великі відстані) історично обмежували потік генів і призводили до формування ізольованих популяцій із високою ізонімією, високим рівнем інбридингу та низькою різноманітністю прізвищ. Натомість доступні ландшафти, судноплавні річки й міські центри сприяли міграції та культурному обміну, що відображається в низькій ізонімії та високій різноманітності.

Система кількісних показників — ізонімія (I^{jj}), коефіцієнт інбридингу (F), індекс різноманітності Фішера (α), індекс міграції (ν), ентропія Шеннона (H), показник надмірності (R), генетичні відстані (Ласкера, Нея, Евкліда, хордова), індекс місця (IP_k), модель ізоляції відстанню та CRSED — становить комплексний інструментарій для оцінювання популяційної структури. Ці показники демонструють узгоджені закономірності під час порівняння різних популяцій і дають змогу реконструювати демографічну історію.

Історичні події глибоко впливають на генетичну структуру популяцій через міграційні процеси, завоювання, колонізацію та адміністративні реформи. Україна як історичне перехрестя між Сходом і Заходом зазнала впливу численних держав і культур, що сформувало її специфічний ономастичний ландшафт. Порівняльний аналіз показує, що Україна займає проміжне положення між західноєвропейськими популяціями та східноєвропейськими регіонами.

Культурні практики, такі як системи називництва, традиції ендогамії та адміністративна політика реєстрації прізвищ, суттєво впливають на їх розподіл. Система подвійних прізвищ у латиноамериканських країнах сприяє збереженню генетичного різноманіття, тоді як традиційна ендогамія в деяких суспільствах призводить до генетичної ізоляції.

Незважаючи на методологічні обмеження, пов'язані із соціокультурними процесами зміни прізвищ і можливою конвергентною еволюцією, метод ізонімії залишається потужним інструментом для популяційно-генетичних досліджень. Особливо цінною є можливість аналізу великих популяцій у ретроспективі без потреби в молекулярно-генетичному тестуванні.

Метою даного дослідження є детальний аналіз закономірностей поширення прізвищ в Україні на обласному та районному рівні для реконструкції міграційних маршрутів, оцінювання рівнів інбридингу та співвіднесення висновків із географічними особливостями регіону й його складною історичною спадщиною. Застосування комплексу популяційно-генетичних показників дасть змогу отримати багатовимірний портрет генетичної структури українського населення на початку XXI століття.

РОЗДІЛ 2 МАТЕРІАЛИ ТА МЕТОДИ

2.1. Джерело даних

Дисертацію виконано на кафедрі генетики і цитології Харківського національного університету імені В. Н. Каразіна в межах науково-дослідної роботи «Генетико-демографічні процеси в українських популяціях» (номер держреєстрації 0224U032384, 2021–2024 рр.).

Масив даних охоплює 46 431 365 записів про носіїв прізвищ, які представляють усе постійне населення України станом на початок ХХІ століття. Загальна кількість унікальних прізвищ становила 697 147. Розподіл населення за областями наведено в табл. 2.1. Чисельність варіювалася від 870 210 осіб у Чернівецькій області до 4 575 766 осіб у Донецькій області.

Набір даних складався виключно з прізвищ в агрегованому вигляді та не містив жодних прямих персональних ідентифікаторів, які дали б змогу встановити належність інформації до конкретної особи. Такий рівень знеособлення відповідає вимогам Закону України «Про захист персональних даних» (2010) та етичним стандартам проведення популяційно-генетичних досліджень. Дані подано у формі частотного розподілу прізвищ за адміністративно-територіальними одиницями без можливості деанонізації.

Масив включав прізвища всіх громадян України незалежно від етнічної належності. Відповідно до аналізу потрапили прізвища представників усіх національних груп: українські, російські, угорські (переважно в Закарпатській області), румунські (у Чернівецькій області), кримськотатарські (АР Крим), польські, єврейські та інші. Така стратегія дає змогу оцінити реальну популяційну структуру, сформовану історичними процесами міграції, асиміляції та адміністративно-територіальних змін упродовж останніх століть.

Жіночі й чоловічі форми прізвищ були об'єднані за базовою (чоловічою) формою. Наприклад, варіанти *Ковальський* і *Ковальська*, *Петренко* і *Шевчук* (чоловіча і жіноча форми) ураховували як єдине прізвище [101]. Таке рішення

обґрунтоване специфікою передачі прізвищ по батьківській лінії: розділення за родовими формами штучно завищило б загальну кількість унікальних варіантів (параметр S) без зміни генетичної інформації, оскільки носії обох форм походять від спільних предків. Цей підхід є стандартним у популяційно-генетичних дослідженнях на основі прізвищ [101] і дає змогу коректно застосовувати моделі, розроблені для аналізу Y-хромосомних маркерів.

Це дослідження є продовженням антропогенетичних праць, виконаних в Україні протягом останніх десятиліть із використанням прізвищ як генетично-демографічних маркерів [92, 97–105]. Застосований підхід базується на принципі ізонімії — методі, що дає змогу оцінювати генетичну структуру популяцій через аналіз розподілу успадкованих ідентифікаторів.

Таблиця 2.1

Характеристика вибірки даних за областями України

Область	N	S	D	Область	N	S	D
Черкаська	1 494 074	99 066	20	Миколаївська	1 163 805	122 422	19
Чернігівська	1 074 996	76 610	22	Одеська	2 143 688	179 591	26
Чернівецька	870 210	62 819	11	Полтавська	1 748 150	105 340	25
АР Крим	2 189 533	186 020	15	Рівненська	1 089 343	60 456	16
Дніпропетровська	3 134 360	188 651	22	Сумська	1 007 814	83 304	18
Донецька	4 575 766	282 661	18	Тернопільська	1 119 051	63 089	17
Харківська	2 818 552	182 765	27	Вінницька	1 905 140	112 371	27
Херсонська	1 144 565	117 012	18	Волинська	1 050 021	59 545	16
Хмельницька	1 524 052	98 551	20	Закарпатська	1 283 620	64 796	13
Кіровоградська	1 152 353	97 390	21	Запорізька	1 748 151	144 325	20
Київська	4 480 704	225 045	25	Житомирська	1 376 812	88 715	23
Луганська	2 229 101	173 457	18	Івано-	1 432 000	66 602	14
Львівська	2 675 504	132 189	20	Франківська			

Примітки: N – кількість носіїв прізвищ, S – кількість прізвищ, D – кількість районів.
Джерело: складено автором на основі отриманого масиву даних.

Детальні популяційно-генетичні характеристики всіх 25 областей і 490 адміністративних районів України, включно з показниками ізонімії, інбридингу, міграції та різноманітності прізвищ, подано в додатку Б (таблиці Б.1–Б.25).

2.2. Статистичний аналіз

У цьому дослідженні застосовано багаторівневий підхід до аналізу структури прізвищ (рис. 2.1). Аналіз охоплює п'ять основних етапів:

1. Розрахунок базових частотних характеристик (розділ 2.2.1): визначення відносних частот прізвищ (q_k^j) та показників ізонімії (I^{jj} , I^{ij}), які відображають імовірність збігу прізвищ усередині популяції та між популяціями;
2. Оцінювання різноманітності (розділ 2.2.2): обчислення показників, що характеризують внутрішньопопуляційну гетерогенність — ентропії Шеннона (H), ефективного числа прізвищ (N_2 , N_I), індексу Фішера (α) та коефіцієнта *CRSED*;
3. Аналіз популяційної структури (розділ 2.2.3): визначення рівня генетичної диференціації (F), показника міграції (ν) та регіональної специфічності (індекс місця $IP_k(R)$);
4. Міжпопуляційні порівняння (розділ 2.2.4): розрахунок матриць генетичних відстаней (Ласкера, Нея, Евкліда, хордової) для кількісного оцінювання подібності популяцій за розподілом прізвищ;
5. Просторовий аналіз (розділ 2.2.5): перевірка моделі ізоляції за відстанню з використанням географічних координат і тесту Мантеля для встановлення кореляції між генетичними та географічними відстанями.

Взаємозв'язки між показниками та послідовність їх обчислення подано на рис. 2.1 і 2.2.

Послідовність статистичного аналізу

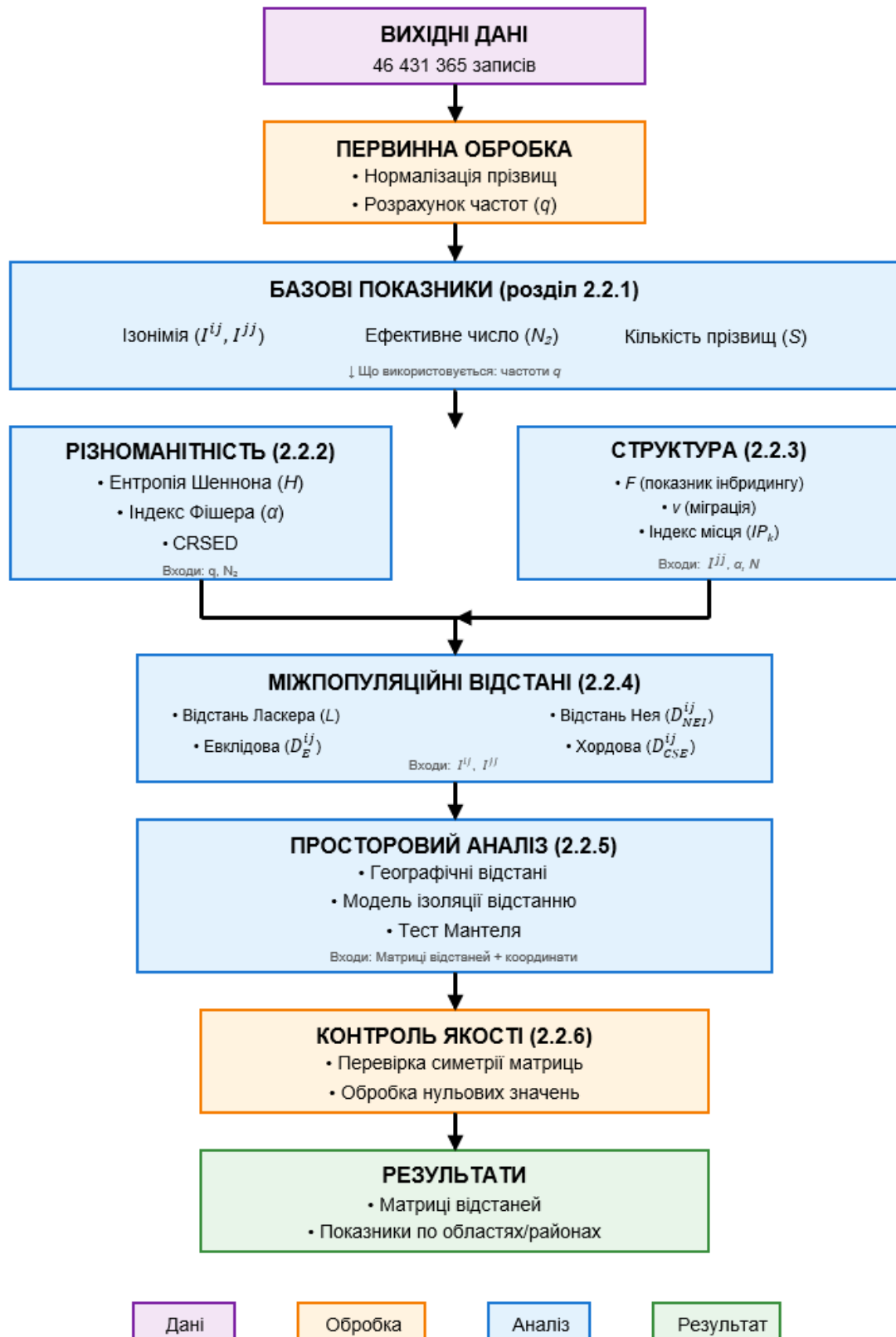


Рис. 2.1. Послідовність статистичного аналізу даних прізвищ та зв'язки між основними показниками.

2.2.1. Частоти прізвищ та показники ізонімії

Ці показники зосереджені на фундаментальних характеристиках розподілу прізвищ у популяції, зокрема на частотах та ймовірностях збігу. Вони слугують основою для подальших розрахунків.

Відносні частоти прізвищ отримано за формулою (2.1) [101]:

$$q_k^j = \frac{n_k^j}{N^j} \quad (2.1)$$

де n_k^j — кількість носіїв k -го прізвища в популяції j , та N^j — загальна чисельність популяції j .

Показник внутрішньопопуляційної ізонімії кількісно визначає ймовірність того, що дві випадково обрані особи мають однакове прізвище [29]. Його обчислюють шляхом підсумовування квадратів частот, оскільки квадрат частоти відображає ймовірність збігу для конкретного прізвища. Ізонімію в популяції j розраховували за формулою (2.2) [29]:

$$I^{jj} = \sum_k (q_k^j)^2 \quad (2.2)$$

де q_k^j — відносна частота (частка) k -го прізвища в популяції j , k — індекс конкретного прізвища.

Метод ізонімії базується на припущенні, що носії однакового прізвища мають спільного предка. Однак у деяких випадках ідентичні прізвища могли виникнути незалежно (наприклад, *Коваль*, *Ковальчук* — від професії коваля в різних регіонах). Як зазначалося раніше, це може призводити до завищення оцінок спорідненості, особливо для дуже поширених прізвищ. Тому інтерпретація результатів має враховувати як історичні (міграції, адміністративні межі), так і соціокультурні (традиції називництва) чинники.

Міжпопуляційна ізонімія (I^{ij}) — міра подібності розподілів прізвищ між популяціями i та j — визначається за формулою (2.3) [29]:

$$I^{ij} = \sum_k q_k^i \cdot q_k^j \quad (2.3)$$

де I^{ij} — скалярний добуток векторів частот для популяцій i та j , q_k^i — відносна частота (частка) k -го прізвища в популяції i , q_k^j — відносна частота (частка) k -го прізвища в популяції j , k — індекс конкретного прізвища.

Для оцінювання різноманітності прізвищ у межах кожної популяції розраховують ефективне число прізвищ (N_2 , зворотний індекс Сімпсона), що визначається як обернене значення ізонімії (формула 2.4) [33, 78]. Цей показник відображає, скільки рівних за частотою прізвищ потрібно, щоб відтворити ту саму картину збігів, що спостерігається у досліджуваній вибірці. На відміну від загальної кількості прізвищ (S), що значною мірою залежить від розміру вибірки та наявності рідкісних варіантів, N_2 ураховує повний спектр частот і є менш чутливим до одиничних або надзвичайно рідкісних прізвищ. Таким чином, N_2 є надійним показником внутрішньопопуляційної різноманітності й дає змогу коректно порівнювати популяції з різними розмірами вибірок¹.

$$N_2 = \frac{1}{\sum_k (q_k^j)^2} = \frac{1}{I^{jj}} \quad (2.4)$$

де N_2 — ефективне число прізвищ, що відповідає кількості рівночастих прізвищ, яка дала б таку ж ізонімію, як спостережений розподіл; q_k^j — відносна частота (частка) k -го прізвища в популяції j , I^{jj} — внутрішньопопуляційна ізонімія.

2.2.2. Показники різноманітності

Показник ентропії Шеннона H [76], визначає ентропію розподілу прізвищ, де q_i — це частота (або відносна частка) i -го прізвища в популяції. Цей показник характеризує ступінь невизначеності чи різноманітності в системі прізвищ. Якщо всі прізвища трапляються з однаковою частотою, H досягає свого максимального значення, що дорівнює $H_{max} = \log_2 N$, де N — загальна кількість особин. Якщо деякі прізвища домінують

¹ Наближення для N_2 — порівняльні, не абсолютні.

(тобто мають дуже високі частоти), H зменшується, оскільки невизначеність вибору прізвища зменшується. Показник H розраховують за формулою (2.5):

$$H^j = - \sum_k q_k^j \log_2 q_k^j \quad (2.5)$$

де H^j — ентропія в популяції j , q_k^j — відносна частота (частка) k -го прізвища в популяції j , k — індекс конкретного прізвища.

Також розраховують відповідне ефективне число N_1 [43, 65] — кількість однаково поширених прізвищ, що дала б таку саму ентропію, як і спостережений розподіл (формула 2.6):

$$N_1 = e^{H^j} \quad (2.6)$$

де N_1 — ентропійне ефективне число прізвищ, H^j — ентропія Шеннона.

Показник надмірності розподілу прізвищ R [101] (відомий також як «дефіцит ентропії») вимірює, наскільки реальна ентропія H^j поступається максимальній ентропії H_{max} . Чим вище значення R , тим менш рівномірним є розподіл прізвищ, тобто тим більше популяція зосереджена на певних прізвищах. Показник R розраховують за формулою (2.7) [101]:

$$R = 100 \left(1 - \frac{H^j}{H_{max}} \right) \quad (2.7)$$

де R — показник надмірності, H^j — ентропія в популяції j , H_{max} — максимальна можлива ентропія за умови рівномірного розподілу.

Таким чином, показник H є мірою загальної невизначеності або різноманітності в розподілі прізвищ, а показник R відображає ступінь надмірності (дефіциту різноманітності) порівняно з максимально можливим рівномірним розподілом.

Показник різноманітності прізвищ (α Фішера) [27] забезпечує оцінювання ефективної кількості прізвищ, наявних у популяції. Він слугує основою для оцінювання загальної генетичної мінливості та зазвичай виражається в одиницях (10^2). Високе значення α вказує на широку різноманітність прізвищ, що свідчить про значну історичну домішку

й інтенсивні міграційні процеси. Показник різноманітності прізвищ α обчислюють за формулою (2.8) [27]:

$$\alpha = \frac{Nv}{1 - v} \quad (2.8)$$

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

Для оцінювання різноманітності прізвищ застосовано коефіцієнт покриття розширеного експоненціального розподілу (*CRSED*), який визначає пропорцію прізвищ у частині розподілу, що підкоряється експоненціальному закону, порівняно зі степеневою частиною розподілу [20]. Для моделювання частот прізвищ, використовували розтягнутий експоненціальний розподіл, який продемонстрував ефективність у демографічних дослідженнях [20]. Ця модель охоплює прізвища, частоти яких підпорядковуються степеневому закону для найпоширеніших варіантів, а також експоненціальному зрізу — для менш поширених. Модель дозволяє доповнити стандартні методи, зокрема ізонімію, урахуванням експоненціальної частини розподілу частот прізвищ. Вищі значення *CRSED* вказують на більшу соціальну та географічну мобільність населення. Розподіл визначають за формулою (2.9) [20]:

$$P(n) = a \cdot n^{-b} \cdot \exp^{-\left(\frac{n}{c}\right)^d} \quad (2.9)$$

де $P(n)$ — кумулятивна ймовірність появи прізвищ із частотою, рівною або більшою за n ; a — параметр масштабу, що визначає нормування ймовірностей, щоб їх сума дорівнювала одиниці; b — початкова швидкість спадання в області степеневого закону; c — початок експоненційного зрізу; d — різкість експоненційного спадання.

Ця модель поєднує поведінку степеневого закону для поширених прізвищ та експоненційного зрізу для рідкісних прізвищ, ефективно охоплюючи весь розподіл [41]. Для підлаштування моделі до даних використано нелінійну регресію методом найменших квадратів для оцінювання параметрів a , b , c і d для кожної області. Підбір виконували за допомогою алгоритму Левенберга — Марквардта (LMA), який є ефективним для розв'язання нелінійних задач найменших квадратів. Алгоритм поєднує градієнтний спуск для обережних і стабільних кроків, коли

модель уже близька до оптимального розв'язку, та метод Гауса — Ньютона для більших і швидших кроків, коли відхилення моделі від емпіричних даних є суттєвим [53].

Показник *CRSED* був розрахований для оцінювання характеристик розподілу прізвищ у різних областях за підходом, застосованим у роботі J.Chen et al. (2019) [20], де *CRSED* допоміг виявити динаміку населення та регіональну подібність. Порогову частоту n_0 , яка розділяє степеневу та експоненціальну частину розподілу, визначали за формулою (2.10) [20]:

$$n_0 = c \cdot \left(\frac{b}{d}\right)^{\frac{1}{d}} \quad (2.10)$$

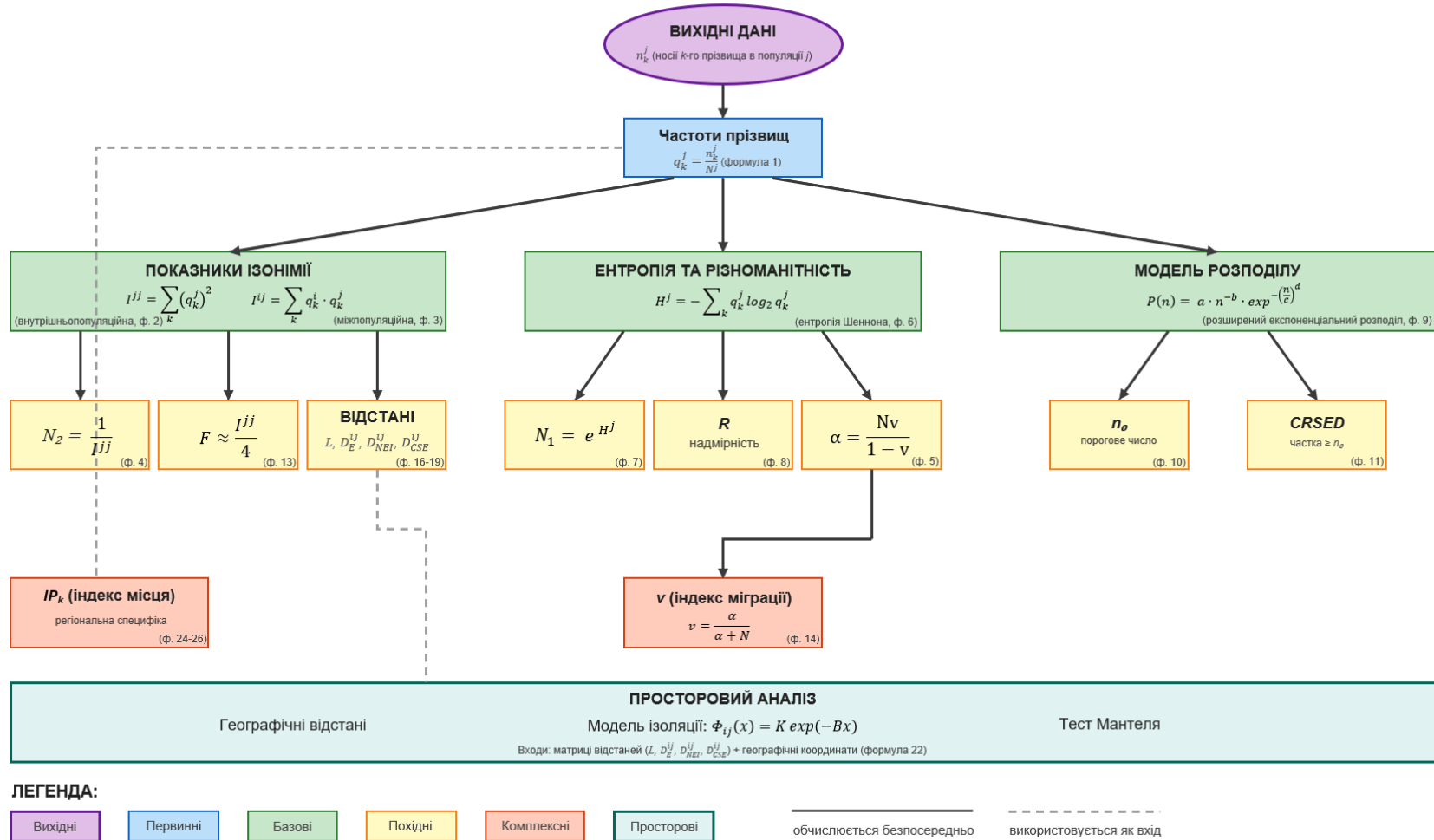
де b — початкова швидкість спадання в області степеневого закону; c — початок експоненційного зрізу; d — різкість експоненційного спадання.

Показник *CRSED* розраховують як частку прізвищ із частотами, що дорівнюють або перевищують n_0 (формула 2.11) [77]:

$$CRSED = \frac{\text{Кількість прізвищ з } n \geq n_0}{S} \quad (2.11)$$

де S — загальна кількість прізвищ.

Взаємозв'язки між показниками структури прізвищ



На **рис. 2.2** узагальнено взаємозв'язки між усіма розрахованими показниками. Стрілками позначено послідовність обчислень: кожен похідний показник розраховують на основі попередніх. Кольорове кодування відображає рівень складності — від первинних частот (блакитний колір) до комплексних метрик просторової структури (жовтогарячий колір).

2.2.3. Показники популяційної структури

Коефіцієнт інбридингу (F) — це показник, який кількісно характеризує ймовірність того, що дві випадково вибрані особи, які є носіями одного й того самого прізвища, генетично споріднені [21, 43, 84].

Застосування F до даних про прізвища базується на аналогії між передачею прізвищ по батьківській лінії та успадкуванням Y-хромосоми [34]. Однак слід враховувати обмеження цієї аналогії: прізвища можуть змінюватися внаслідок соціальних процесів (русифікація, полонізація), тоді як генетичні маркери залишаються стабільними. Тому наведені значення F слід інтерпретувати як евристичну оцінку популяційної структури, а не як абсолютні генетичні параметри [63, 66].

Евристичну оцінку диференціації за аналогією до F подають у вигляді формули (2.12) [16, 21, 22]:

$$F \approx \frac{I_{jj}}{4} \quad (2.12)$$

Ця апроксимація коректна за умови стабільної патрилійної передачі прізвищ протягом багатьох поколінь. Для України, де відбувалися масові зміни прізвищ (русифікація XVIII–XX ст., радянська уніфікація 1930-х), значення F може бути систематично заниженим порівняно з молекулярно-генетичними даними (розділ 1.5.4). Тому наведені оцінки слід інтерпретувати як нижню межу популяційної диференціації.

Показник міграції (v Карліна — МакГрегора) кількісно визначає ступінь потоку генів або обміну прізвищами між субпопуляціями [36, 60, 83]. Він виражається в одиницях 10^{-4} і відображає ступінь, до якого сучасна або історична міграція зумовила появу нових прізвищ серед населення. Вищі значення v свідчать про активну міграцію та підвищене генетичне змішування. Показник міграції v розраховується за формулою (2.13) [36, 60, 83]:

$$v = \frac{\alpha}{\alpha + N} \quad (2.13)$$

де v — інтенсивність міграції, α — показник Фішера, N — чисельність популяції.

2.2.4. Міжпопуляційні відстані

Ця група показників фокусується на відстанях і коефіцієнтах, що вимірюють подібність або відмінності між популяціями на основі розподілу прізвищ.

Для виявлення ізоляції за відстанню було розраховано лінійну кореляцію між відстанями за прізвищами (відстанями Ласкера, Евкліда та Нея) між населеними пунктами i та j і відповідними географічними відстанями.

Відстань Ласкера (L) вимірює ступінь відмінності між популяціями [44, 66]. Чим більше значення L , тим більше відмінність між популяціями (менше спільних прізвищ). Чим менше значення L , тим ближчі популяції (спільні прізвища трапляються частіше й із подібною частотою). Показник є логарифмічною мірою відмінності між популяціями та визначається через індекс Ласкера за формулою (2.14) [44, 66]:

$$L = -\ln(I^{ij}) \quad (2.14)$$

де I^{ij} — міжпопуляційна ізонімія.

Евклідова відстань [19] є мірою відмінності між двома популяціями на основі розподілу прізвищ. Якщо евклідова відстань дорівнює нулю, то популяції вважаються ідентичними за розподілом прізвищ. Якщо евклідова відстань наближається до одиниці, то популяції суттєво відрізняються. Показник D_E^{ij} , обчислений для векторів частот, розраховується за формулою (2.15) [19]:

$$D_E^{ij} = \sqrt{I^{ii} + I^{jj} - 2I^{ij}} \quad (2.15)$$

де I^{ij} — міжпопуляційна ізонімія, I^{jj} — внутрішньопопуляційна ізонімія.

Цей показник використовують для оцінювання ступеня спорідненості або відмінності між групами населення на основі частот спільних прізвищ.

Генетична відстань Нея (D_{NEI}^{ij}) використовується для оцінювання еволюційної та генетичної відстані між популяціями [55, 57, 66]. Чим більше значення показника D_{NEI} , тим менша генетична схожість між популяціями, тобто тим давніше вони розійшлися. Чим менше показник D_{NEI} , тим більш

спорідненими є популяції, тобто вони мають спільне походження або перебувають у стані активного генетичного обміну. Відстань Нея (адаптовану для аналізу прізвищ) визначають за формулою (2.16) [55, 57, 66]:

$$D_{NEI}^{ij} = -\ln \left(\frac{I^{ij}}{\sqrt{(I^{ii} \cdot I^{jj})}} \right) \quad (2.16)$$

де I^{ij} — міжпопуляційна ізонімія, I^{jj} — внутрішньопопуляційна ізонімія відповідних популяцій.

Хордова відстань Каваллі–Сфорца–Едвардса (D_{CSE}^{ij}) дає змогу виміряти, наскільки дві популяції подібні чи відрізняються за частотами генів або прізвищ [19]. Умовно кожен популяцію можна розглядати як точку в багатовимірному просторі, де кожен вимір відповідає певному гену або ознаці. Відстань між цими точками вимірюється по хорді (прямій лінії через сферу), а не по її поверхні; ця відстань зростає зі збільшенням розбіжностей у частотах ознак між популяціями. Отже, D_{CSE}^{ij} відображає генетичну «віддаленість» груп у наочний і порівняльний спосіб. Показник розраховують за формулою (2.17) [19]:

$$D_{CSE}^{ij} = \sqrt{2 \left(1 - \sum_k \sqrt{q_k^i \cdot q_k^j} \right)} \quad (2.17)$$

де q_k^i і q_k^j — відносні частоти прізвища k у популяціях i та j .

Індекс Ласкера (I_L^{ij}) — це показник спорідненості двох популяцій за розподілом прізвищ [44, 66]. Він визначає, наскільки подібними є популяції за частотами прізвищ (чим вищий його показник, тим більш подібними є популяції), і розраховується за формулою (2.18) [44, 66]:

$$I_L^{ij} = \sum_k \sqrt{q_k^i \cdot q_k^j} \quad (2.18)$$

де q_k^i і q_k^j — відносні частоти прізвища k у популяціях i та j .

Матриці відстаней обчислювали для районів із ненульовим значенням I^{jj} ; діагональні елементи встановлювали рівними нулю, забезпечуючи симетричність матриці.

2.2.5. Просторовий аналіз та регіональна специфіка

У цьому дослідженні спираємося на класичну модель «ізоляції за відстанню» [37, 47], згідно з якою випадкова спорідненість між двома локалізаціями i і j , розділеними лінійною відстанню x , описується експоненціальним законом (2.19) [37, 47]:

$$\Phi_{ij}(x) = K \exp(-Bx) \quad (2.19)$$

де $K = \Phi_{ij}(0)$ відображає середню спорідненість за нульової відстані (еквівалент середнього F), а параметр $B \propto \sqrt{\mu/\sigma^2}$ пов'язує інтенсивність «мутацій» прізвищ μ та дисперсію їх географічного переміщення σ^2 . Емпірично $\Phi_{ij}(x)$ оцінюється через частоту збігів прізвищ $I^{ij}(x)$ за формулою (2.20) [37, 47]:

$$\widehat{\Phi}_{ij}(x) = \frac{I^{ij}(x)}{4} \quad (2.20)$$

що переводить спостережувану частку ідентичностей у шкалу спорідненості, сумісну з F [14].

Географічні відстані між центрами районів обчислювали за формулою гаверсинуса (Haversine) з використанням радіуса Землі $r \approx 6371$ км (2.21) [79]:

$$D = 2r \arcsin \sqrt{\sin^2 \frac{\Delta\varphi}{2} + \cos\varphi_1 \cos\varphi_2 \sin^2 \frac{\Delta\lambda}{2}} \quad (2.21)$$

де d_{ij} — географічна дистанція між центроїдами популяцій i та j ; r — радіус Землі (≈ 6371 км); φ — широти; λ — довготи.

Повну матрицю географічних відстаней формували для пар із валідними координатами; для опису масштабу вказували мінімальне, максимальне, середнє значення та медіану.

Для перевірки статистичної значущості залежності $\widehat{\Phi}_{ij}(x)$ від x застосовувався тест Мантеля [48, 80]. Тест Мантеля використовують для

перевірки того, чи не виявлений випадковий зв'язок між сімейною спорідненістю та географічними розташуванням випадковим. Дані багаторазово випадково перемішували та щоразу повторно обчислювали кореляції. Якщо спостережувана кореляція виявляється значно сильнішою за будь-які «випадкові» варіанти, можна обґрунтовано говорити про ефект «ізоляції за відстанню» між спільнотами.

Було сформовано дві матриці — географічних відстаней D та оцінок спорідненості $\hat{\Phi}$, після чого обчислювали коефіцієнт кореляції Пірсона між їхніми позадіагональними елементами. Послідовно здійснювали перестановку рядків і стовпців однієї з матриць (999 рандомізацій) із двобічною перевіркою, що дозволяло отримати нульовий розподіл кореляцій та оцінити p -значення.

Індекс місця (IP_k) є кількісною мірою подібності розподілу прізвищ у конкретному регіоні до загальнонаціонального шаблону [4, 99, 101, 102]. Кожному прізвищу s присвоюється національний ранг $r(s)$ за його позицією у зведеному списку частот: найпоширеніше прізвище отримує ранг 1, десяте — 10, соте — 100 і т.д. У досліджуваному регіоні R формується множина $S_k(R)$ з k найчастіших прізвищ; національні ранги цих прізвищ підсумовуються та діляться на k (2.22):

$$IP_k(R) = \frac{1}{k} \sum_{s \in S_k(R)} r(s) \quad (2.22)$$

де $IP_k(R)$ — індекс місця для району (області, регіону) R , $S_k(R)$ — безліч k найчастіших прізвищ району (області, регіону), $r(s)$ — національний ранг прізвища s .

За $k = 5, 10$ або 20 отримують відповідно IP_5, IP_{10} та IP_{20} . Якщо значення $IP_k(R)$ близьке до загальнонаціонального стандарту IP_k^{nat} , це свідчить про те, що регіональні лідери за частотою збігаються з найпоширенішими в країні, що вказує на низьку регіональну специфіку. Зростання відхилення від еталону відображає наявність у регіоні прізвищ, які на загальнонаціональному рівні займають нижчі позиції, і, отже, характеризує регіон як більш своєрідний. Наприклад, якщо в регіоні X топ-5 прізвищ мають ранги 1, 4, 25, 7, 2, то (2.23):

$$IP_5(X) = \frac{1 + 4 + 25 + 7 + 2}{5} = 7.8 \quad (2.23)$$

що за загальнонаціонального значення 5,3 вказує на помірну відмінність регіону. Індекс місця цінується за наочність і можливість порівняння різних регіонів або часових зрізів, однак не враховує абсолютних частот прізвищ і є чутливим до вибору параметра k . У соціологічних і геногеографічних дослідженнях його часто доповнюють показником (2.24):

$$\Delta IP_k(R) = IP_k(R) - IP_k^{nat} \quad (2.24)$$

що дає змогу кількісно оцінити ступінь відхилення регіону від загальнонаціональної норми [4, 99, 101, 102].

2.2.6. Кластеризація та багатовимірні візуалізації

Агломеративну ієрархічну кластеризацію проводили методом Уорда («ward.D2») [54, 85] для евклідових метрик (зокрема для D_E^{ij} та евклідових координат після PCoA/MDS) [30, 38, 39].

Дендрограми візуалізували в лінійному та круговому форматах. Теплові карти будували з упорядкуванням рядків і стовпців відповідно до дендрограми. Для описових візуалізацій застосовували багатовимірне шкалювання (MDS) та аналіз головних координат (PCoA) [38, 39].

2.2.7. Контроль якості та числова стабільність

Після нормалізації даних перевіряли, що суми відносних частот у кожному стовпці матриць дорівнюють одиниці, щоб забезпечити коректність обчислень. Крім того, забезпечували симетричність матриць відстаней (шляхом дзеркального копіювання елементів) та встановлювали нульові значення на діагоналі відповідно до математичних вимог.

Особливу увагу приділяли обробці пар із нульовою міжпопуляційною ізонімією $I^{ij} = 0$, яка може виникати через відсутність спільних прізвищ між регіонами. В аналізах на рівні областей застосовували усічення $I^{ij} \leftarrow \max(I^{ij}, 10^{-15})$ перед логарифмуванням, щоб уникнути невизначених значень (наприклад, $\log(0) = -\infty$) і зберегти стабільність обчислень. В інших

аналізах (зокрема на рівні районів) пари з $I^{ij} = 0$ призводили до *NA* або *Inf*, які на етапі кластеризації замінювали на консервативне значення — поточний максимум матриці, збільшений на 10%. Стабільність отриманих висновків перевіряли шляхом варіювання порогів заміни (наприклад, від 5 % до 20 %), оцінюючи вплив на кінцеві метрики, такі як структура кластерів або значення відстаней. Це дозволило підтвердити робастність результатів до потенційних артефактів даних.

Зв'язок між показниками, розподіл яких не відповідає нормальному, оцінювали за методом Спірмена. Коефіцієнт кореляції розраховували за формулою [2, 81] (2.25):

$$r_s = 1 - \frac{6}{n(n^2 - 1)} \sum d^2 \quad (2.25)$$

де r_s — коефіцієнт кореляції Спірмена; n — число спостережень; d — різниця між рангами ознак, що спостерігалися.

2.3. Програмне забезпечення генетико-статистичного аналізу

Зберігання та аналіз інформації, а також підготовку списків прізвищ здійснювали з використанням програмного забезпечення Microsoft 365 (Microsoft Excel) і R 4.3.3. Парні та окремі коефіцієнти кореляції між матрицями генетичних відстаней розраховували за допомогою тесту Мантеля. Обчислення показників і багатовимірне шкалювання проводили з використанням мови програмування R 4.3.3 та Microsoft Excel [26, 62].

У R 4.3.3 [62] застосовували такі пакети (за призначенням): *geosphere* — обчислення географічних відстаней (формула гаверсинуса, функції *distHaversine*, *distm*); *ape* — аналіз головних координат (PCoA) та допоміжні функції для побудови дендрограм; *vegan* — тест Мантеля та суміжні багатовимірні процедури; *ggtree* — візуалізація дерев (у т. ч. радіальні дендрограми); *dplyr*, *tibble* — таблицна обробка даних і маніпуляції з фреймами; *stringr*, *stringdist* — обробка рядків і метрики рядкових відстаней; *clue* — утиліти для кластеризації/порівняння кластерних розбиттів.

Для великих обсягів даних застосовували потокове читання; розмір чанків (порції даних) добирали динамічно для оптимізації пам'яті. Вихідні дані експортували в табличному форматі (матриці, сумарні метрики) та у вигляді графіків для подальшої інтерпретації.

УЗАГАЛЬНЕННЯ ТА ВИСНОВКИ ДО РОЗДІЛУ 2

Методологічна основа дослідження базується на комплексному застосуванні методів популяційно-генетичного аналізу до унікального масиву даних про 46 431 365 носіїв 697 147 прізвищ з усіх областей України станом на 2001 рік.

Розроблений багаторівневий підхід охоплює п'ять основних категорій показників:

1. Показники внутрішньопопуляційної структури (I^{jj} , F , α , ν) дають змогу кількісно оцінити рівень генетичної ізоляції, ефективний розмір популяції та інтенсивність міграційних процесів;
2. Показники різноманітності (H , R , N_1 , N_2 , $CRSED$) забезпечують багатовимірну характеристику антропонімичного різноманіття. Інноваційне застосування методу $CRSED$ дає змогу виявляти демографічні ефекти міграційної історії через аналіз переходу від степеневого до експоненціального розподілу частот прізвищ;
3. Міжпопуляційні відстані (Ласкера, Нея, евклідова, хордова) використовують для кількісної оцінки генетичної подібності між популяціями. Порівняльний аналіз виявив специфічні переваги кожної метрики: евклідова дистанція (D_E^{ij}) оптимальна для виявлення внутрішньообласної гетерогенності, дистанція Ласкера (L) — для аналізу великомасштабних демографічних процесів, дистанція Нея (D_{NEI}^{ij}) — для реконструкції еволюційних взаємовідносин;
4. Просторовий аналіз через модель ізоляції за відстанню та тест Мантеля дає змогу перевірити кореляцію між генетичними та географічними відстанями. Індекс місця (IP_k) забезпечує оцінку регіональної специфічності;
5. Багатовимірні методи (кластеризація методом Уорда, PCoA, MDS) застосовують для виявлення прихованої структури популяційних взаємовідносин.

Програмну реалізацію виконано в R 4.3.3 із використанням спеціалізованих пакетів (*geosphere*, *ape*, *vegan*, *minpack.lm*, *ggtree*). Контроль якості передбачав перевірку симетрії матриць, валідацію нормалізації та обробку нульових значень через усічення ($\max(I^{ij}, 10^{-15})$). Стабільність результатів верифіковано варіюванням порогів заміни від 5 % до 20 %.

Топологічна стабільність ключових кластерів у трьох метриках відстаней підтверджує надійність виявленої структури та її відповідність реальним філогенетичним взаємовідносинам субпопуляцій України. Розроблена методологія забезпечує високу роздільну здатність аналізу на всіх ієрархічних рівнях і є адекватним інструментом для досягнення мети дисертаційного дослідження.

РОЗДІЛ 3 РЕЗУЛЬТАТИ ТА ОБГОВОРЕННЯ

3.1. Загальноукраїнські значення показників

Популяційно-генетична структура України характеризується виразною регіональною гетерогенністю, яка відображає складну взаємодію історичних, географічних та демографічних факторів. Згідно з даними М. Ю. Горпинченка, загальноукраїнський індекс внутрішньогрупової ізонімії (I^{jj}) оцінюється у $2,1 \times 10^{-4}$, проте цей показник демонструє суттєву просторову варіабельність.

Регіональний аналіз виявляє закономірний градієнт: південні, східні та західні регіони характеризуються порівняно нижчими значеннями ($2,2 \times 10^{-4}$, $2,7 \times 10^{-4}$ та $2,4 \times 10^{-4}$ відповідно), тоді як центральні та північні області демонструють підвищені показники ($3,2 \times 10^{-4}$ та $3,4 \times 10^{-4}$ відповідно) [101]. Порівняльний аналіз популяційних показників усіх областей України (додаток Б) виявив значну регіональну варіабельність за всіма досліджуваними параметрами.

Внутрішньообласна варіація ізонімії (I^{jj}) є ще більш вираженою. Наприклад, на сході значення коливаються від $2,0 \times 10^{-4}$ до $2,9 \times 10^{-4}$, на півдні — від $1,6 \times 10^{-4}$ до $2,6 \times 10^{-4}$, у центральній частині України — від $2,8 \times 10^{-4}$ до $5,3 \times 10^{-4}$, на півночі — від $3,5 \times 10^{-4}$ до $5,3 \times 10^{-4}$, на заході — від $2,3 \times 10^{-4}$ до $6,8 \times 10^{-4}$. Цей південно-східний — північно-західний градієнт відображає фундаментальні відмінності в демографічній історії регіонів.

Території з високою історичною мобільністю населення, зокрема розташовані вздовж основних торговельних шляхів (наприклад, Крим), характеризуються нижчою ізонімією внаслідок інтенсивного генетичного потоку. Натомість області з природними географічними бар'єрами, такими як гірські масиви Карпат чи лісисті простори Полісся (Волинська та Закарпатська області), демонструють підвищені значення ізонімії, що узгоджується з гіпотезою про відносну ізоляцію популяцій.

Іншим ключовим показником є різноманітність прізвищ, що вимірюється індексом різноманітності α і становить для України

в середньому $47,0 \times 10^2$. Регіональний розподіл цього показника виявляє чітку зворотну кореляцію з ізонімією: найвища різноманітність спостерігається у південних ($44,8 \times 10^2$) та західних ($41,4 \times 10^2$) областях, проміжні значення характерні для східних ($37,1 \times 10^2$) і центральних регіонів ($31,7 \times 10^2$), тоді як найнижчу різноманітність демонструють північні області ($29,4 \times 10^2$).

Ця закономірність узгоджується з теоретичними очікуваннями: вища різноманітність прізвищ асоціюється зі зниженою ймовірністю кровноспоріднених шлюбів і меншим ризиком інбридингу. Ентропія розподілу прізвищ (H), яка в середньому для України становить 15,1, додатково підтверджує цей зв'язок: більш рівномірний розподіл прізвищ (вища ентропія) корелює з меншим ризиком споріднених союзів [101].

Міграційні процеси відіграють ключову роль у формуванні антропонімічного різноманіття. Міграційний індекс (v), який відображає швидкість уведення нових прізвищ у популяцію, оцінюється для України у $1,01 \times 10^{-4}$. Просторовий розподіл цього показника демонструє виразну неоднорідність: максимальні значення характерні для південних областей ($5,6 \times 10^{-4}$), мінімальні — для центральних ($3,4 \times 10^{-4}$), із помітною внутрішньорегіональною варіабельністю.

Чіткий південно-східний — північно-західний градієнт міграційного індексу узгоджується з історичними даними про колонізацію південних степових територій у XVI–XVIII століттях та індустріалізацію східних регіонів у XX столітті. Порівняно низька міграційна ефективність України відносно багатьох європейських країн свідчить про відносну стабільність генофонду протягом останніх століть [101].

Коефіцієнт інбридингу (F) безпосередньо оцінює генетичні ризики, пов'язані з кровною спорідненістю. Загальноукраїнське значення F становить $5,3 \times 10^{-5}$ і маскує значну міжобласну варіацію: від $4,0 \times 10^{-5}$ у Криму до $17,9 \times 10^{-5}$ на Волині. Просторовий розподіл коефіцієнта інбридингу демонструє конгруентність із розподілом ізонімії: нижчі значення характерні

для регіонів з високою мобільністю населення, вищі — для географічно ізольованих областей.

Оскільки споріднені союзи підвищують імовірність гомозиготності за рецесивними алелями, ці результати мають важливе значення для системи охорони здоров'я та медико-генетичного консультування, особливо в регіонах з підвищеними значеннями показника [101].

Усі популяційно-генетичні показники, наведені в підрозділі 3.1, розраховані М. Ю. Горпинченком [101], на основі бази даних, яка охоплює 46 449 268 носіїв прізвищ з усіх регіонів України.

3.2. Аналіз розподілу прізвищ за методом CRSED

Для поглибленого аналізу структури антропонімичного різноманіття застосовано метод CRSED (*Coverage Ratio of Stretched Exponential Distribution*), який дає змогу кількісно охарактеризувати пропорцію прізвищ у частині розподілу, що підпорядковується експоненціальному закону, порівняно зі степеневим [20, 52, 77].

3.2.1. Теоретичні основи методу

Розподіл частот прізвищ у великих популяціях має складну, бімодальну структуру, що відображає еволюційну історію антропонімів. Поширені прізвища, які виникли в ранні періоди формування антропонімичних систем, встигли досягти високої частоти й широкого географічного поширення. Їхній розподіл підпорядковується степеневому закону (power law) [3, 52], що є типовим для систем із кумулятивною перевагою («*the rich get richer*») — механізмом, за якого високочастотні варіанти передаються дедалі більшій кількості нащадків, посилюючи свою представленість у популяції (рис. 3.1).

Рідкісні прізвища, що з'явилися в пізніші періоди або виникли локально, демонструють іншу модель розподілу. Спад їхніх частот підпорядковується експоненціальному закону, характерному для стохастичних, неагрегованих подій без механізму накопичення переваг (рис. 3.1). Точка переходу між цими

двома режимами розподілу, визначена параметром n_0 моделі CRSED, маркує критичний поріг між «старими» та «новими» прізвищами в популяції.

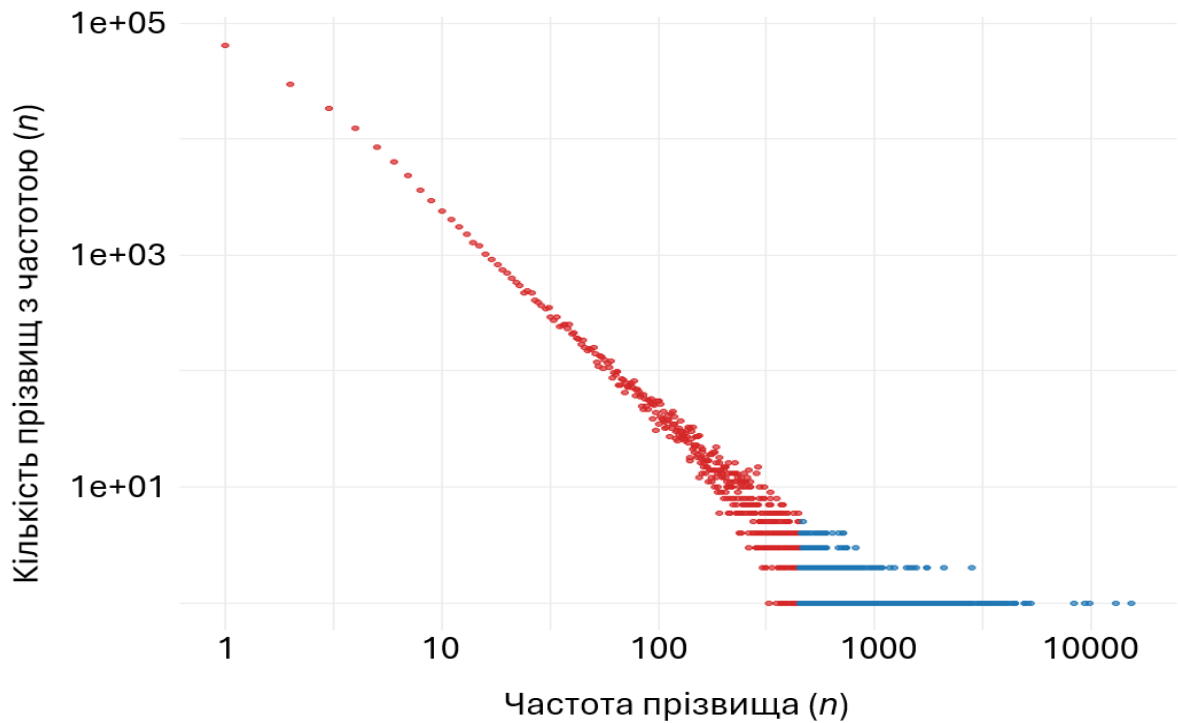


Рис. 3.1. Спад за степеневим законом та експоненціальний хвіст, визначені за розподілом частот прізвищ щодо CRSED-порогу, на прикладі Харківської області (синій — область степеневого закону, червоний — експоненціальна область). Поріг, що розділяє ці області, відповідає n_0 , як визначено моделлю CRSED [77].

3.2.2. Регіональна варіація параметрів CRSED

Аналіз 25 областей України виявив значну варіацію параметрів моделі CRSED (табл. 3.1, рис. 3.1 – 3.4). Параметр b , який характеризує швидкість спадання частоти поширених прізвищ, демонструє діапазон від від’ємних значень (Івано-Франківська область: $b = -0,025$) до $0,119$ (АР Крим). Області з низькими значеннями b ($b \leq 0,05$) — Львівська, Чернівецька, Рівненська, Івано-Франківська, Чернігівська, Тернопільська, Волинська, Хмельницька, Вінницька, Сумська — характеризуються повільним зменшенням частоти домінуючих прізвищ, що свідчить про обмежений набір поширених антропонімів із поступовим переходом до менш частих варіантів [77].

Від’ємне значення параметра b в Івано-Франківській області відображає відхилення форми розподілу від стандартної степеневої апроксимації —

явище, ймовірно, пов'язане зі специфічними культурно-історичними особливостями гуцульської популяції [77].

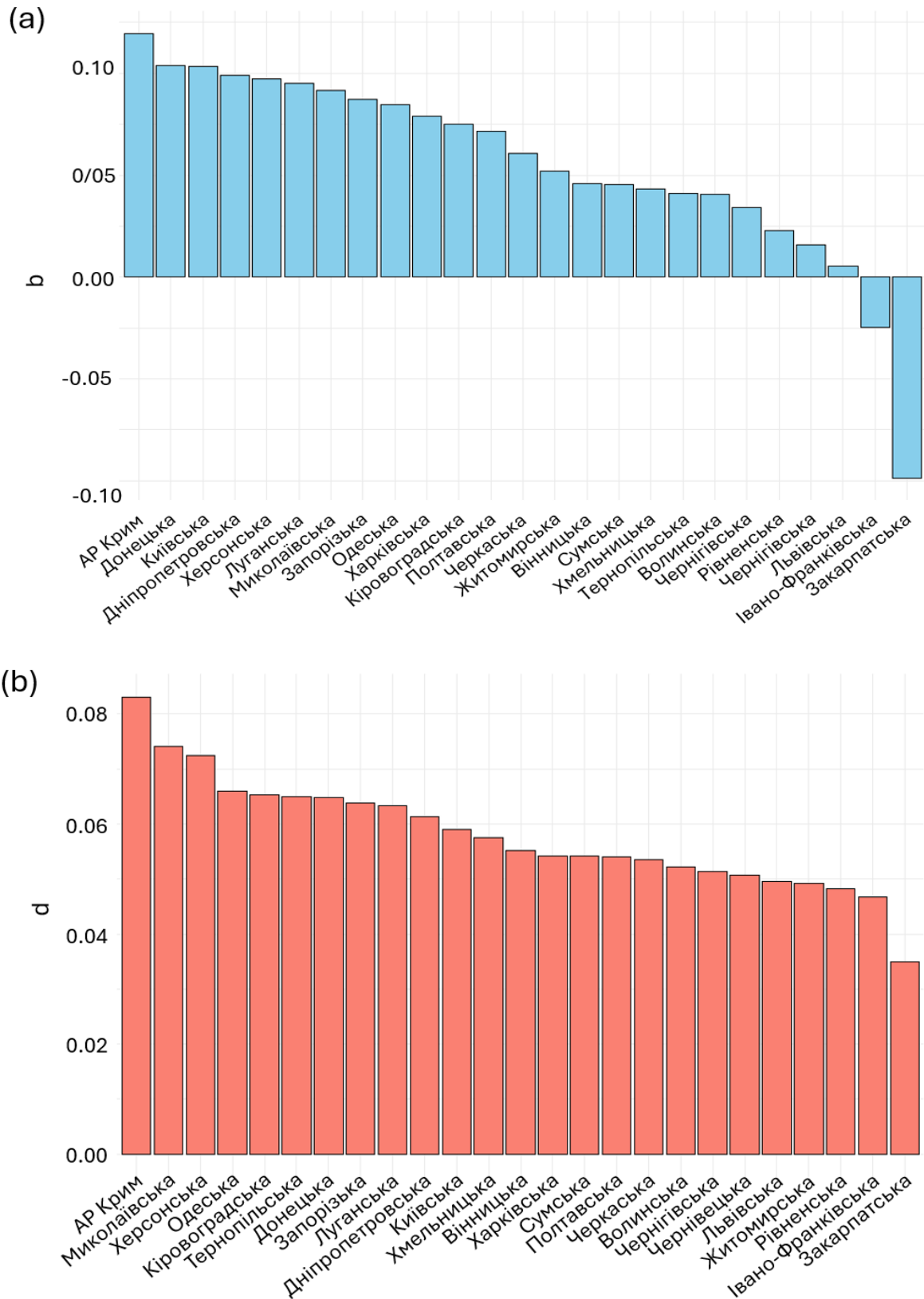


Рис. 3.2. Обласні варіації параметрів CRSED: (a) — параметр b (коефіцієнт степеневі шкали), (b) — параметр d (сила експоненційного згасання). Області відсортовані за спаданням значення параметра [77].

Області з помірними значеннями b ($0,05 < b \leq 0,08$) — Житомирська, Черкаська, Полтавська, Кіровоградська, Харківська — демонструють більш виражене зменшення частоти поширених прізвищ, що вказує на швидший перехід від домінуючих до рідкісних варіантів і підвищену загальну різноманітність антропонімів [77].

Області з високими значеннями b ($b > 0,08$) — Одеська, Запорізька, Миколаївська, Луганська, Херсонська, Дніпропетровська, Закарпатська, Київська, Донецька області та АР Крим — характеризуються різким спаданням частоти після невеликої кількості найпоширеніших прізвищ. Така структура типова для урбанізованих регіонів з інтенсивними міграційними процесами та етнічною гетерогенністю [77].

Параметр c визначає положення точки переходу між степеневою та експоненціальною ділянками розподілу.

Параметр d , що відображає крутість експоненціального спадання для рідкісних прізвищ, варіює у відносно вузькому діапазоні — від 0,482 (Рівненська область) до 0,830 (АР Крим) [77]. Області з нижчими значеннями d ($d \leq 0,50$) — Закарпатська, Івано-Франківська, Рівненська, Житомирська, Львівська — демонструють повільніше зменшення частоти рідкісних прізвищ, що свідчить про значний внесок низькочастотних варіантів у загальне антропонімичне різноманіття [77]. Більшість областей характеризується помірними значеннями d ($0,50 < d \leq 0,70$), тоді як області з високими значеннями ($d > 0,70$) — Крим, Миколаївська та Херсонська — демонструють швидке зниження частоти рідкісних прізвищ до мінімальних значень [77].

3.2.3. Класифікація областей за CRSED

На основі значень CRSED області України класифіковано на три категорії: Клас I ($0 \leq CRSED \leq 0,01$), Клас II ($0,01 < CRSED \leq 0,05$), Клас III ($CRSED > 0,05$) (Рис. 3.3, 3.4). Використання логарифмічної шкали для порогових значень зумовлене значною варіацією показника CRSED, що дає змогу змістовно диференціювати популяції [87, 86]. **Клас I** представляє області, розподіл прізвищ у яких найбільш наближений до степеневої функції,

тоді як **Клас III** охоплює області, для яких розподіл прізвищ наближається до розтягнуто-експоненціальної функції.

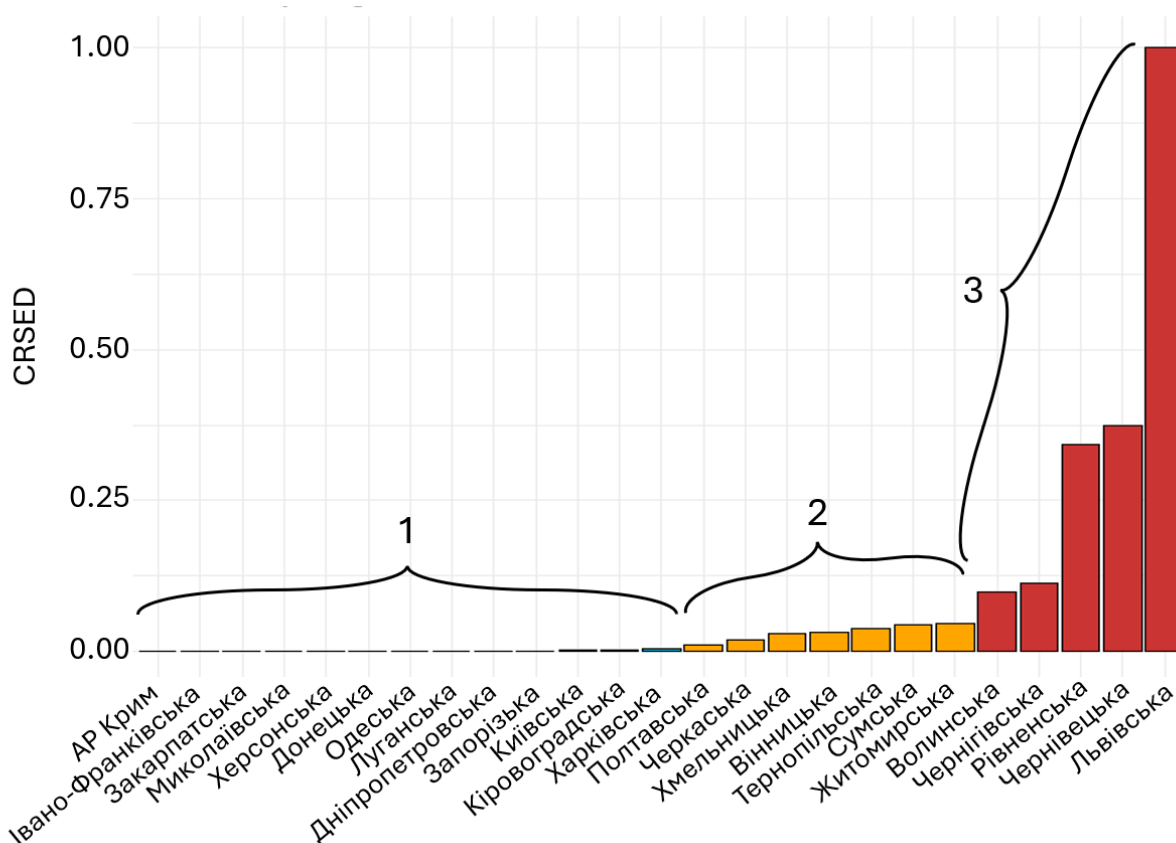


Рис. 3.3. Значення CRSED для областей України, розподілені за трьома класами CRSED [77].

Області класу I (АР Крим, Закарпатська, Івано-Франківська, Миколаївська, Херсонська, Донецька, Одеська, Луганська, Дніпропетровська, Запорізька, Київська, Кіровоградська, Харківська області) з низькими значеннями CRSED розміщені переважно на сході та півдні України, у степовій зоні (рис. 3.4а). Ці землі були спустошені монгольською навалою XIII століття й поступово заселялися українськими колоністами починаючи з XV–XVI століть. Області характеризуються високою часткою етнічних меншин [93, 107]: кримські татари; грецькі громади (переселені з Криму до Приазов'я наприкінці XVIII століття); російська діаспора (сформована після депортації кримських татар у 1944 році); єврейські, болгарські, румунські та молдавські спільноти в Одеській області. Зазначені історичні процеси можуть бути одним із чинників, що формують сучасні значення CRSED у цих областях.

У середині ХХ століття східні області зазнали інтенсивної індустріалізації, що спричинило масовий приплив робітників із різних республік СРСР. У західній частині України також наявні області з низьким *CRSED*: Закарпаття з великою угорською діаспорою та Івано-Франківська область, де понад 99 % населення ідентифікують себе як гуцули [110]. Параметр *d* (табл. 3.1, рис. 3.2b) фіксує повільне зниження частоти рідкісних прізвищ, тобто їхній кумулятивний внесок у загальний антропонімічний масив залишається значним. Відповідно до класифікації Chen et al. [20], популяції класу I належать до категорії «регіонів імміграції».

Області класу II (Полтавська, Черкаська, Хмельницька, Вінницька, Тернопільська, Сумська, Житомирська області) зосереджені переважно в центральній Україні. Їхнє географічне положення зумовлює проміжні значення популяційно-демографічних характеристик (середнє $I^{jj} = 4,5 \times 10^{-4}$). Репресії 1930-х років (розкуркулення 1929–1932 рр., Голодомор 1932–1933 рр.) призвели до загибелі та депортації від 3 до 7 млн осіб, переважно в центральних областях [93, 111].

Події Другої світової війни спричинили подальші примусові переміщення: численні втрати на фронтах, депортації до Німеччини, післявоєнні репатріації та переселення робітників із відбудованих східних регіонів [107]. Таким чином, населення областей класу II формувалося під впливом багаторазових депортацій, репатріацій і економічних міграцій, що спричинили постійні зміни етнодемографічного складу. За класифікацією Chen et al. [20], ці регіони належать до категорії «рекультивованих».

Області класу III (Львівська, Рівненська, Чернівецька, Волинська, Чернігівська області) характеризуються найвищими значеннями *CRSED* та підвищеними індексами ізонімії (середнє $I^{jj} = 5,0 \times 10^{-4}$). Високі значення ізонімії свідчать про значну концентрацію носіїв поширених прізвищ, тоді як високі величини *CRSED* фіксують суттєве зниження частоти рідкісних прізвищ. Географічно ці регіони локалізовані переважно в західній і північній частинах України — у лісовій зоні Полісся та передгір'ях Карпат.

Історична демографія цих територій характеризується масштабними еміграційними хвилями XVI–XVII століть, спрямованими на колонізацію східних степових земель [93]. Протягом наступних століть ці області тривалий час перебували під владою сусідніх імперій, що спричиняло прикордонне змішування різноетнічного населення. Особливо значущим демографічним потрясінням стали примусові переселення 1939 року, коли внаслідок радянської анексії східної Польщі велика частина місцевого населення підлягала депортації до Сибіру [111].

Згідно з класифікацією Chen et al. [20], ці області належать до категорії «регіонів еміграції», оскільки їхня сучасна демографічна структура сформувалася під визначальним впливом відтоку населення. Ці демографічні події відображаються у низьких значеннях CRSED та високих значеннях параметра b , зафіксованих у відповідних регіонах (табл. 3.1)

3.2.4. Порівняльний аналіз і методологічні висновки

Відносно низьке значення коефіцієнта кореляції Спірмена ($r = 0,44$) вказує на те, що CRSED враховує форму розподілу, зокрема рідкісні прізвища, на відміну від ізонімії, яка фокусується на поширених [77]. Згідно з Chen [20], на відміну від показника випадкової ізонімії, що майже повністю визначається частотою найпоширеніших прізвищ, CRSED відображає загальну форму розподілу, зокрема баланс між домінуючими та рідкісними прізвищами. Таким чином, CRSED доповнює аналіз ізонімії, включаючи інформацію про рідкісні прізвища, яка за використання класичних методів залишається поза увагою.

Порівняння українських результатів із китайським дослідженням [20] виявляє як структурні паралелі, так і суттєві відмінності. В обох країнах регіони з низькими значеннями CRSED (гірські та автономні зони в Китаї; Карпати та Полісся в Україні) характеризуються високою антропонімічною консервативністю та мінімальною інтеграцією зовнішніх прізвищ. «Рекультивовані регіони» китайського Північного Сходу та Заходу, а також степово-приазовські простори України демонструють підвищений CRSED унаслідок інтенсивних зовнішніх демографічних втручань.

Оцінені параметри та значення CRSED у 25 регіонах

Область	N	S	$I^{jj} \times 10^{-4}$	Параметри					CRSED	K
				a	b	c	d			
АР Крим	2 189 533	186 020	1,61	1,624	0,119	45955	0,830	0,000000	1	
Івано-Франківська	1 432 000	66 602	4,62	0,980	-0,025	1955	0,467	0,000000	1	
Закарпатська	1 283 620	64 796	6,09	0,990	-0,099	427	0,349	0,000000	1	
Миколаївська	1 163 805	122 422	2,41	1,383	0,092	22911	0,740	0,000196	1	
Херсонська	1 144 565	117 012	2,57	1,409	0,097	21384	0,724	0,000248	1	
Донецька	4 575 766	282 661	2,04	1,506	0,104	38369	0,649	0,000460	1	
Одеська	2 143 688	179 591	2,10	1,374	0,085	25213	0,660	0,000663	1	
Луганська	2 229 101	173 457	2,4	1,432	0,095	23533	0,634	0,000755	1	
Дніпропетровська	3 134 360	188 651	2,81	1,434	0,099	26681	0,614	0,000983	1	
Запорізька	1 748 151	144 325	2,53	1,366	0,087	19453	0,638	0,001136	1	
Київська	4 480 704	225 045	3,51	1,431	0,103	24071	0,590	0,001693	1	
Кіровоградська	1 152 353	97 390	3,59	1,262	0,075	12929	0,654	0,002485	1	
Харківська	2 818 552	182 765	2,92	1,349	0,079	15765	0,542	0,004481	1	
Полтавська	1 748 150	105 340	4,09	1,263	0,071	8610	0,540	0,010480	2	
Черкаська	1 494 074	99 066	5,08	1,199	0,060	7489	0,536	0,018654	2	
Хмельницька	1 524 052	98 551	5,05	1,105	0,043	7845	0,575	0,029812	2	
Вінницька	1 905 140	112 371	5,32	1,127	0,046	7930	0,552	0,031521	2	
Тернопільська	1 119 051	63 089	3,20	1,126	0,041	6165	0,650	0,038073	2	
Сумська	1 007 814	83 304	4,47	1,145	0,046	5886	0,542	0,044356	2	
Житомирська	1 376 812	88 715	5,34	1,203	0,052	5093	0,492	0,046971	2	
Волинська	1 050 021	59 545	7,14	1,116	0,041	3546	0,521	0,098211	3	
Чернігівська	1 074 996	76 610	4,57	1,123	0,034	4246	0,514	0,114099	3	
Рівненська	1 089 343	60 456	6,77	1,099	0,023	2658	0,482	0,343175	3	
Чернівецька	870 210	62 819	4,29	1,078	0,016	3426	0,507	0,373072	3	
Львівська	2 675 504	132 189	2,3	1,057	0,005	5471	0,495	1.000000	3	

Примітки: N — кількість носіїв прізвищ, S — кількість прізвищ, I^{jj} — показник ізонімії для тотального списку, a — параметр масштабу, що нормує ймовірності, b — початкова швидкість спадання в області степеневого закону; c — початок експоненційного зрізу; d — різкість експоненційного спадання, K — клас.

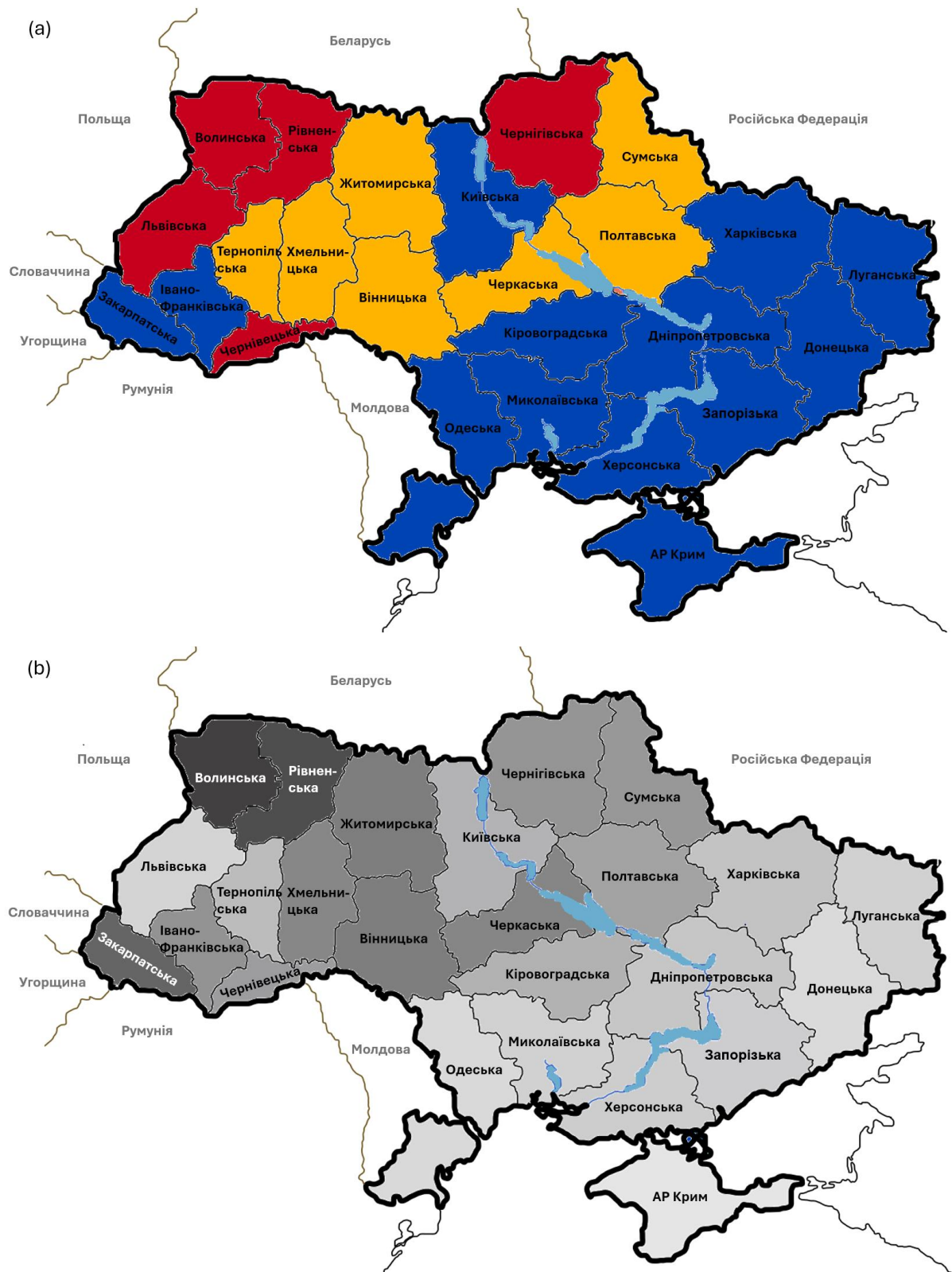


Рис. 3.4. Карти класифікації CRSED (a) та ізонімії (b) в областях України. (Карта CRSED поділяє області за пороговими значеннями розподілу частот прізвищ: група 1 — синій колір, група 2 — жовтий, група 3 — червоний. Карта ізонімії відображає подібність прізвищ у межах областей, де темніші відтінки вказують на вищі значення ізонімії) [77].

Проте часові рамки й природа антропонімічних традицій суттєво різняться. Китайські прізвища мають історію понад чотири тисячоліття та відображають давні демографічні події, складні для історичної реконструкції. Українські прізвища сформувалися протягом останніх трьох століть і мають чітку історіографічну прив'язку до конкретних міграційних, військових та політичних подій [107, 108, 111]. Це означає, що в українському контексті CRSED переважно відображає межі антропонімічної консервативності та документовані історичні хвилі міграції, тоді як у китайському випадку він радше сигналізує про масштаб імперських чи династійних втручань у систему формування прізвищ.

Отже, застосування методу CRSED є методологічно обґрунтованим у межах поставлених завдань дослідження: він доповнює індекси, розраховані на підставі показника ізонімії, дозволяючи виявити тонші демографічні й антропонімічні ефекти міграційної історії, а також дає змогу порівнювати регіональні ономастичні закономірності з іншими культурами в межах уніфікованої методології [20].

3.3. Районні значення показників: детальний просторовий аналіз

Перехід від обласного до районного рівня аналізу розкриває значно більшу просторову гетерогенність популяційно-генетичних показників (рис. 3.5). Підвищена варіабельність відображає локальні демографічні процеси, географічні особливості та мікрорегіональну міграційну динаміку.

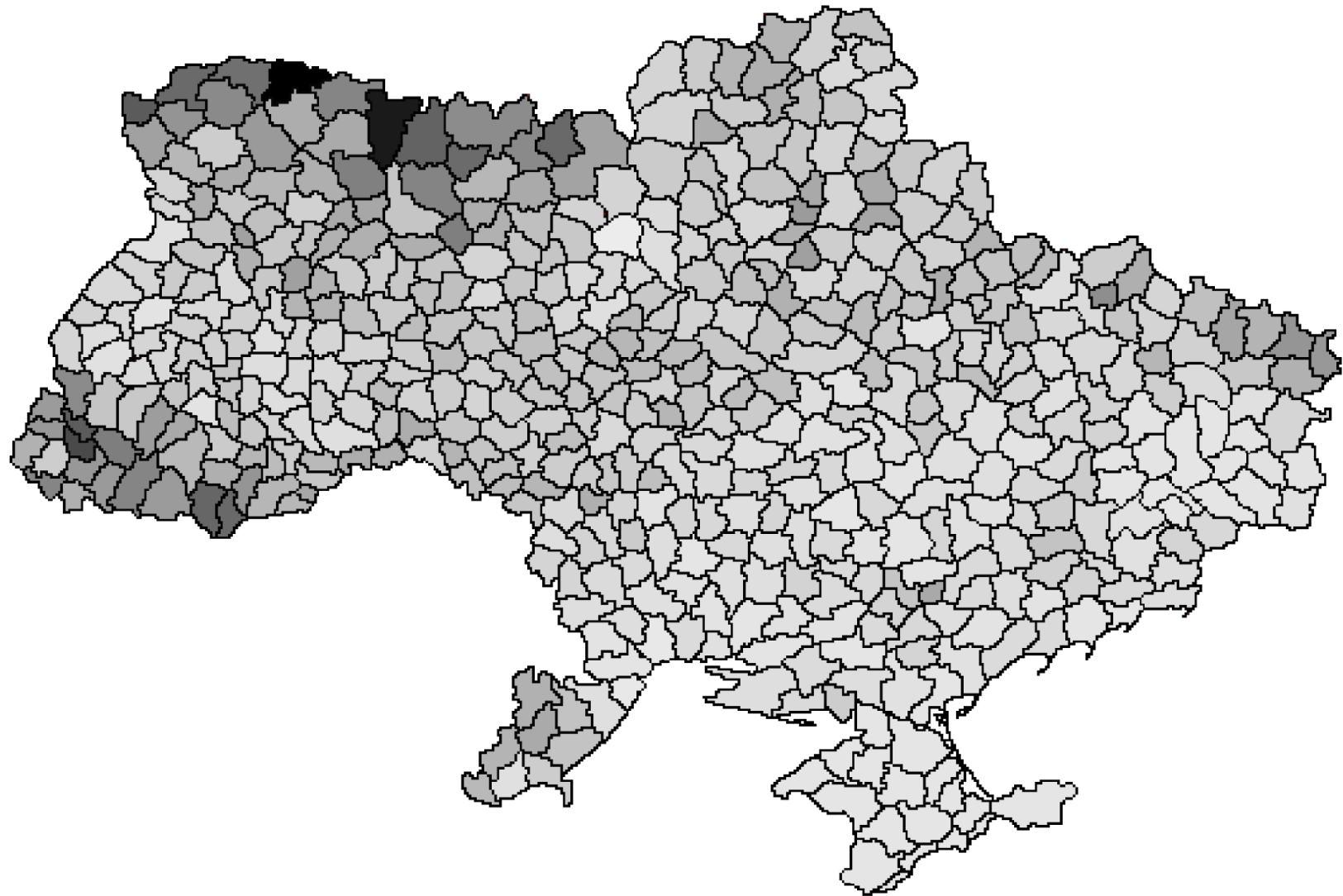


Рис. 3.5. Розподіл показників інбридингу в районних популяціях України (світлі зони відповідають низьким значенням показника, темні — високим)

3.3.1. Схід України

Схід України демонструє помітну варіабельність популяційно-генетичних показників, пов'язаних із прізвищем. У Харківській, Донецькій та Луганській областях індекс ізонімії (I^{jj}) становить відповідно $2,5 \times 10^{-4}$, $2,0 \times 10^{-4}$ та $2,4 \times 10^{-4}$. У межах Харківської області I^{jj} змінюється від $3,3 \times 10^{-4}$ у м. Харкові до $25,6 \times 10^{-4}$ у Печенізькому районі. Так само в Донецькій області значення коливаються від $2,8 \times 10^{-4}$ у Донецько-Макіївській міській агломерації до $10,3 \times 10^{-4}$ у Лиманському районі, тоді як у Луганській області індекс варіює від $3,8 \times 10^{-4}$ у м. Луганську до $23,7 \times 10^{-4}$ у Марківському районі [122] (рис. 3.6).

На сході України коефіцієнт інбридингу становить $6,3 \times 10^{-5}$ у Харківській області, $5,1 \times 10^{-5}$ у Донецькій та 6×10^{-5} у Луганській [122]. На районному рівні мінімальний показник інбридингу в Харківській області зафіксовано в м. Харкові ($8,3 \times 10^{-5}$), максимальний — у Печенізькому районі ($63,9 \times 10^{-5}$). У Донецькій області найнижчий рівень інбридингу спостерігається в обласному центрі та місті-супутнику Макіївці (7×10^{-5}), тоді як у Лиманському районі — $25,7 \times 10^{-5}$. У Луганській області районні показники коливаються в межах від $8,7 \times 10^{-5}$ у Попаснянському районі до $59,3 \times 10^{-5}$ у Марківському районі, що підкреслює суттєві відмінності в рівнях спорідненості серед місцевого населення [122].

Ці закономірності відображають демографічні процеси, пов'язані з індустріалізацією та урбанізацією регіону. У Харківській області, зокрема, м. Харків і прилеглі до нього райони характеризуються найнижчими показниками інбридингу, що зумовлено інтенсивною внутрішньою міграцією з різних регіонів України та з-за кордону. Міграційні потоки, спричинені розвитком промисловості та функціонуванням численних закладів вищої освіти, сприяли зростанню різноманітності прізвищ і, відповідно, зниженню ймовірності кровноспоріднених союзів. Порівняльний аналіз карт просторового розподілу індексів ізонімії з розташуванням великих промислових підприємств засвідчує, що райони, які приваблюють значну

кількість мігрантів, зазвичай характеризуються нижчими значеннями ізонімії та коефіцієнта випадкового інбридингу. Це, своєю чергою, підсилює загальну генетичну й демографічну неоднорідність населення.

Показник ентропії (H), який відображає невизначеність імовірної зустрічі з тим чи іншим прізвищем, у цілому по Харківській області становить 14,3. На районному рівні його значення варіюють від 9,7 у Печенізькому районі до 32,3 у м. Харкові. Показник R у межах області змінюється від 23,2 у Кегичівському районі до 30,4 в обласному центрі. У Донецькій області загальне значення H становить приблизно 15; районні показники коливаються від 11,2 в Олександрівському районі до 14,1 в Донецько-Макіївській міській агломерації. Середнє значення R дорівнює 32,2 і варіює в межах від 22,9 до 30,7. У Луганській області загальний H становить 31, а районні значення коливаються від 9,9 у Марківському районі до 13,4 у Попаснянському районі. Загальний показник R дорівнює 31 і змінюється від 24 у Сорокинському районі до 30 у Марківському районі [122]. Спостережувані розбіжності між районними та обласними значеннями можна пояснити кумулятивними ефектом: агрегація ідентичних прізвищ у межах різних районів підвищує різноманітність на регіональному рівні, що зумовлює вищі інтегральні показники для області порівняно з її окремими адміністративними одиницями.

Індекс різноманітності прізвищ (α) демонструє обернену залежність від масштабу аналізу. У Харківській області регіональне значення показника становить $39,5 \times 10^2$, тоді як на районному рівні воно варіює від 4×10^2 у Печенізькому районі до $19,4 \times 10^2$ у Лозівському районі. У Донецькій області регіональний показник досягає 49×10^2 , а районні значення коливаються від $9,9 \times 10^2$ у Лиманському районі до $35,7 \times 10^2$ у Донецько-Макіївській міській агломерації. У Луганській області індекс регіонального різноманіття становить 31×10^2 , тоді як районні значення охоплюють діапазон від $4,3 \times 10^2$ у Марківському районі до 29×10^2 у Попаснянському районі, що свідчить про вищу однорідність у межах окремих районів порівняно з регіоном у цілому [122].

Міграція залишається домінуючим чинником формування структури населення. Індекс міграції (v) кількісно відображає зміни у складі популяції внаслідок припливу мігрантів. У Харківській області регіональний показник міграції становить $22,6 \times 10^{-4}$, тоді як на районному рівні він істотно варіює — від $30,4 \times 10^{-4}$ у м. Харкові до $465,4 \times 10^{-4}$ у Близнюківському районі. У Донецькій області середньообласне значення показника дорівнює $10,9 \times 10^{-4}$, тоді як на районному рівні воно зростає до $27,8 \times 10^{-4}$ у Донецько-Макіївській міській агломерації та досягає максимального значення $464,4 \times 10^{-4}$ у Нікольському районі. У Луганській області регіональний показник міграції становить 31×10^{-4} , а районні значення коливаються від $69,3 \times 10^{-4}$ у м. Луганську до $322,9 \times 10^{-4}$ у Міловському районі [122]. Такі значні розбіжності свідчать про високу інтенсивність міграційних процесів в районному рівні, що зумовлює меншу стабільності демографічної структури локальних спільнот порівняно з регіоном у цілому.

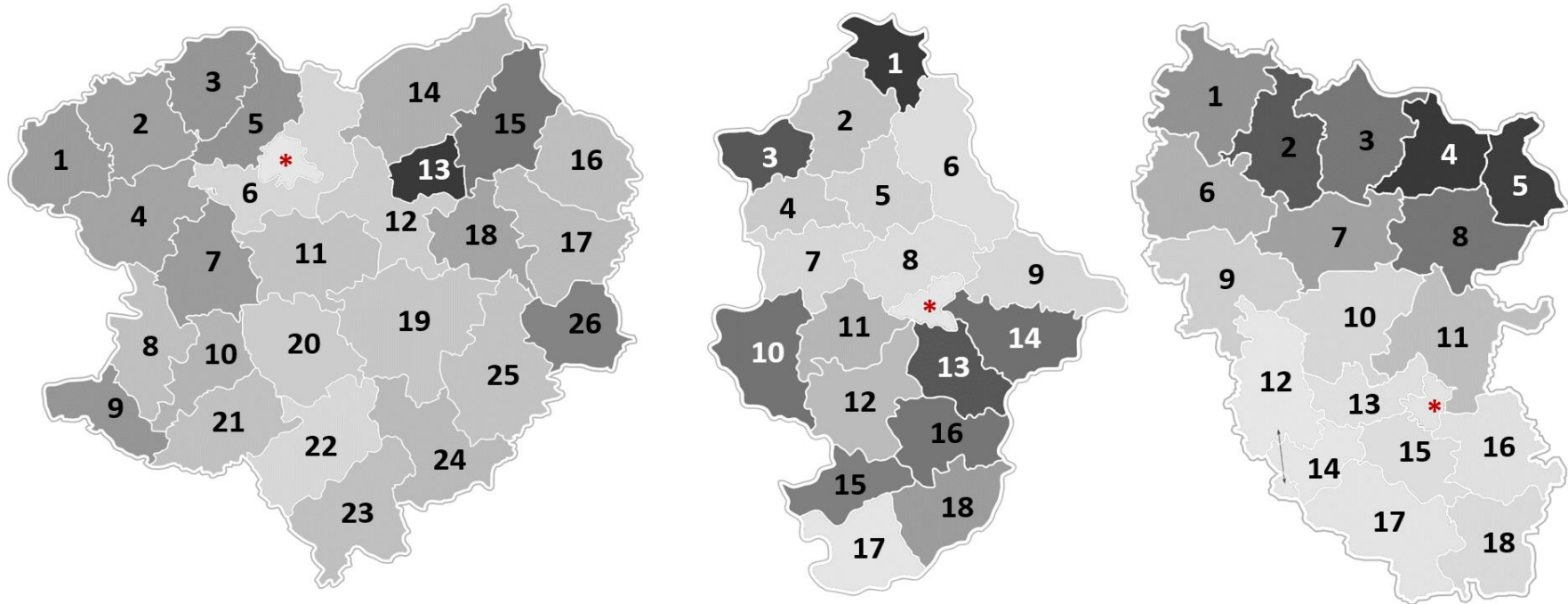


Рис. 3.6. Просторовий розподіл показників ізонімії в районних популяціях Харківської, Донецької та Луганської областей (світлі зони відповідають низьким значенням показника, темні — високим)

Райони Харківської області: * — м. Харків; 1. Краснокутський; 2. Богодухівський; 3. Золочівський; 4. Валківський та Коломацький; 5. Дергачівський; 6. Харківський; 7. Нововодолазький; 8. Красноградський; 9. Зачепилівський; 10, Кегичівський; 11. Зміївський; 12. Чугуївський; 13. Печенізький; 14. Вовчанський; 15. Великобурлуцький; 16. Дворічанський; 17. Куп'янський; 18. Шевченківський; 19. Балаклійський; 20, Первомайський; 21. Сахновщинський; 22. Лозівський; 23. Близнюківський.

Райони Донецької області: * — м. Донецьк; 1. Краснолиманський; 2. Слов'янський; 3. Олександрівський; 4. Добропільський; 5. Костянтинівський; 6. Артемівський; 7. Красноармійський; 8. Ясинуватський; 9. Шахтарський; 10, Великоновосілківський; 11. Мар'їнський; 12. Волноваський; 13. Старобешівський; 14. Амвросіївський; 15. Володарський; 16. Тельманівський; 17. Першотравневий; 18. Новоазовський.

Райони Луганської області: * — м. Луганськ; 1. Троїцький; 2. Білокуракинський; 3. Новопсковський; 4. Марківський; 5. Міловський; 6. Сватівський; 7. Старобільський; 8. Біловодський; 9. Кременський; 10, Новоайдарський; 11. Станично-Луганський; 12. Попаснянський; 13. Слов'яносербський; 14. Перевальський; 15. Лутугинський; 16. Краснодонський; 17. Антрацитівський; 18. Свердловський.

3.3.2. Північ України

Північ України демонструє значну територіальну варіабельність популяційно-генетичних показників (рис. 3.7, 3.8). Аналіз індексу внутрішньопопуляційної ізонімії (I^{jj}), коефіцієнта інбридингу (F), індексу міграції (v), а також показників різноманітності прізвищ, ентропії (H) і надмірності розподілу прізвищ (R) виявляє чіткі відмінності як між областями, так і в межах окремих районів.

У Київській області загальний індекс ізонімії становить $3,5 \times 10^{-4}$; на районному рівні його значення варіюють від $3,2 \times 10^{-4}$ у м. Києві до $36,4 \times 10^{-4}$ у Поліському районі [126]. У Чернігівській області регіональний показник I^{jj} дорівнює $4,3 \times 10^{-4}$; районні значення коливаються від $8,9 \times 10^{-4}$ у м. Чернігові до $22,4 \times 10^{-4}$ у Срібнянському районі. У Сумській області загальний індекс становить $4,5 \times 10^{-4}$, а районні значення змінюються від $5,5 \times 10^{-4}$ у м. Суми до $19,7 \times 10^{-4}$ у Липоводолинському районі. У Житомирській області зафіксовано найвище значення I^{jj} — $5,3 \times 10^{-4}$; на районному рівні показник варіює від $4,5 \times 10^{-4}$ у м. Житомирі до $38,2 \times 10^{-4}$ в Олевському районі.

Коефіцієнт інбридингу (F) також демонструє помітні регіональні та локальні варіації. У Київській області загальне значення F становить $8,8 \times 10^{-5}$ і варіює у межах від $8,2 \times 10^{-5}$ у м. Києві до 91×10^{-5} у Поліському районі [126]. У Чернігівській області регіональний показник дорівнює $10,7 \times 10^{-5}$; на районному рівні він змінюється від $22,2 \times 10^{-5}$ у м. Чернігові до 56×10^{-5} у Срібнянському районі. У Сумській області F становить $11,2 \times 10^{-5}$, а районні значення коливаються від $13,6 \times 10^{-5}$ у м. Суми до $49,2 \times 10^{-5}$ у Липоводолинському районі. У Житомирській області регіональний показник дорівнює $13,3 \times 10^{-5}$ і варіює у межах від $11,2 \times 10^{-5}$ у м. Житомирі до $95,5 \times 10^{-5}$ в Олевському районі.

Показники розподілу прізвищ додатково ілюструють регіональну неоднорідність. На обласному рівні індекс ентропії (H) є відносно стабільним: 14,3 у Київській області, 13,3 у Чернігівській та Житомирській, 13,4 у Сумській. Проте на районному рівні значення H демонструють значно

ширшу варіацію: у Київській області — від 9,4 у Поліському районі до 13,2 у Києво-Святошинському районі; у Чернігівській — від 9,8 у Срібнянському районі до 11,9 у Прилуцькому районі [126]; у Сумській — від 10,1 у Липоводолинському районі до 12,8 у м. Суми; у Житомирській — від 9,5 у Лугинському районі до 13,1 у м. Житомирі. Індекс надмірності розподілу прізвищ (R) доповнює ці спостереження: середні обласні значення становлять 35,4 у Київській області, 32,8 у Чернігівській, 33,8 у Сумській та 35 у Житомирській, тоді як районні показники характеризуються подібною варіативністю.

Різноманітність прізвищ, виражена індексом різноманітності (α), є нижчою на районному рівні порівняно з обласним. У Київській області значення показника варіюють від $2,8 \times 10^2$ у Поліському районі до $30,9 \times 10^2$ у м. Києві [126]. У Чернігівській області індекс різноманітності змінюється від $4,6 \times 10^2$ у Срібнянському районі до $12,7 \times 10^2$ у Прилуцькому районі. У Сумській області значення коливаються від $5,2 \times 10^2$ у Липоводолинському районі до $18,5 \times 10^2$ у м. Суми, а в Житомирській області – від $2,6 \times 10^2$ в Олевському районі до $22,6 \times 10^2$ у м. Житомирі.

Міграційні процеси суттєво вплинули на популяційно-генетичну структуру північної частини України. Регіональний індекс міграції (ν) варіює від $6,3 \times 10^{-4}$ у Київській області до $26,8 \times 10^{-4}$ у Чернігівській. Однак на районному рівні значення показника є майже на порядок вищими: у Київській області вони змінюються від $11,9 \times 10^{-4}$ у м. Києві до $333,3 \times 10^{-4}$ у Володарському районі [126]; у Чернігівській — від 181×10^{-4} у м. Чернігові до $316,8 \times 10^{-4}$ у Срібнянському районі; у Сумській — від 107×10^{-4} у м. Суми до $403,8 \times 10^{-4}$ у Путивльському районі; у Житомирській — від $55,8 \times 10^{-4}$ в Олевському районі до 421×10^{-4} у Брусилівському районі.

Найвищий рівень генетичного різноманіття та міграції спостерігається у місті Києві, що відображає його статус столиці та потужного економічного центру, який приваблює мігрантів з усієї країни [126]. Натомість сільські території, особливо в межах Полісся, характеризуються нижчими показниками

різноманітності та вищими значеннями інбридингу, що свідчить про відносну ізолюваність і демографічну замкнутість локальних популяцій.

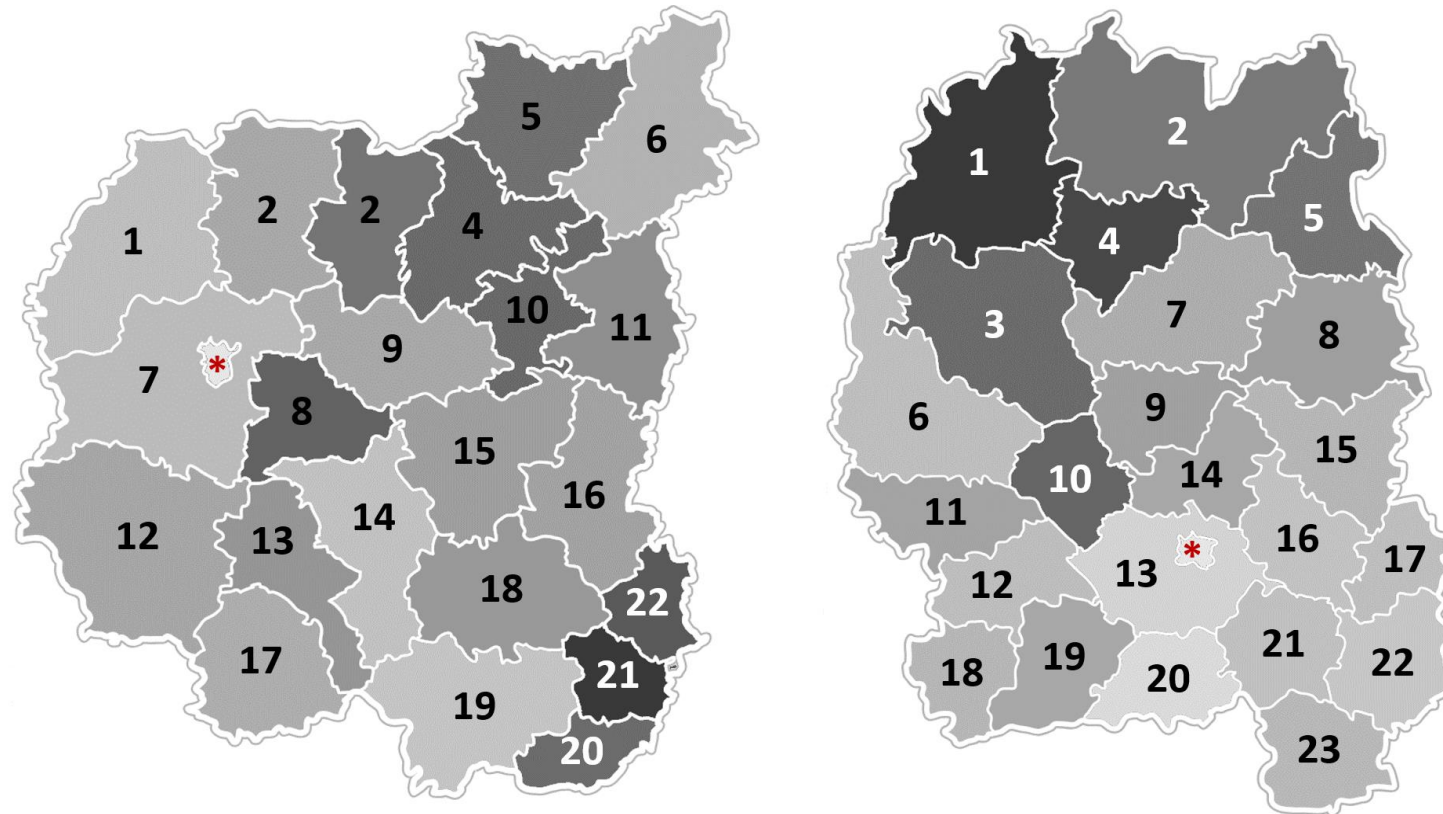


Рис. 3.7. Просторовий розподіл показників ізонімії в районних популяціях Чернігівської та Житомирської областей (світлі зони відповідають низьким значенням показника, темні — високим)

Райони Чернігівської області: * — м. Чернігів; 1. Ріпкинський; 2. Городнянський; 3. Щорський; 4. Корюківський; 5. Семенівський; 6. Новгород-Сіверський; 7. Чернігівський; 8. Куликівський; 9. Менський; 10, Сосницький; 11. Коропський; 12. Козелецький; Носівський; 14. Ніжинський; 15. Борзнянський; 16. Бахмацький; 17. Бобровицький; 18. Ічнянський; 19. Прилуцький; 20, Варвинський; 21. Срібнянський; 22. Талалаївський.

Райони Житомирської області: * — м. Житомир; 1. Олевський; 2. Овруцький; 3. Ємільчинський; 4. Лугинський; 5. Народицький; 6. Новоград-Волинський; 7. Коростенський; 8. Малинський; 9. Володарсько-Волинський; 10, Червоноармійський; 11. Баранівський; 12. Романівський; 13. Житомирський; 14. Черняхівський; 15. Радомишльський; 16. Коростишівський; 17. Брусилівський; 18. Любарський; 19. Чуднівський; 20, Бердичівський; 21. Андрушівський; 22. Попільнянський; 23. Ружинський.

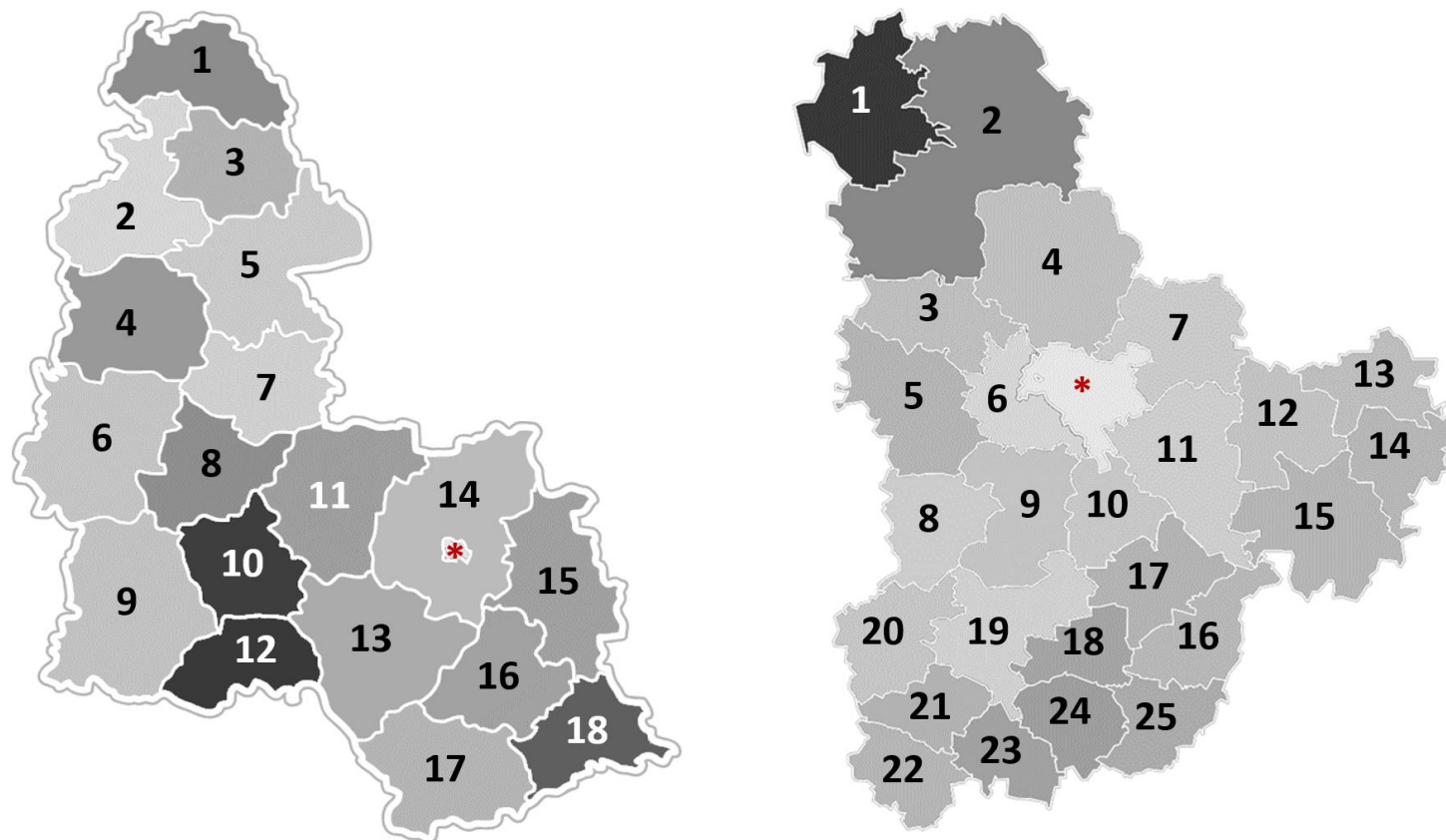


Рис. 3.8. Просторовий розподіл показників ізонімії в районних популяціях Сумської та Київської областей (світлі зони відповідають низьким значенням показника, темні — високим)

Райони Сумської області: * — м. Суми; 1. Середино-Будський; 2. Шостківський; 3. Ямпільський; 4. Кролевецький; 5. Глухівський; 6. Конотопський; 7. Путивльський; 8. Буринський; 9. Роменський; 10, Недригайлівський; 11. Білопільський; 12. Липоводолинський; 13. Лебединський; 14. Сумський; 15. Краснопільський; 16. Тростянецький; 17. Охтирський; 18. Великописарівський.

Райони Київської області: * — м. Київ; 1. Поліський; 2. Іванківський; 3. Бородянський; 4. Вишгородський; 5. Макарівський; 6. Києво-Святошинський; 7. Броварський; 8. Фастівський; 9. Васильківський; 10, Обухівський; 11. Бориспільський; 12. Баришівський; 13. Згурівський; 14. Яготинський; 15. Переяслав-Хмельницький; 16. Миронівський; 17. Кагарлицький; 18. Рокитнянський; 19. Білоцерківський; 20, Сквирський; 21. Володарський; 22. Тетіївський; 23. Ставищенський; 24. Таращанський; 25. Богуславський

3.3.3. Центр України

Центральна частина України демонструє популяційно-генетичні характеристики, які загалом є співставними з показниками Східної та Північної України, однак на районному рівні спостерігається виражена внутрішньорегіональна варіабельність. Загальні значення випадкової ізонімії індекс ізонімії (I^{jj}) для центральних областей становлять: у Вінницькій — $5,3 \times 10^{-4}$, у Черкаській — $5,1 \times 10^{-4}$, у Полтавській — $4,1 \times 10^{-4}$, у Дніпропетровській — $2,8 \times 10^{-4}$, у Кіровоградській — $3,6 \times 10^{-4}$. У межах цих регіонів значення I_{jj} суттєво варіюють. У Вінницькій області районні показники коливаються від $4,7 \times 10^{-4}$ у м. Вінниці до $18,4 \times 10^{-4}$ у Мурованокуриловецькому районі. У Черкаській області I^{jj} змінюється від $4,2 \times 10^{-4}$ у м. Черкасах до $16,1 \times 10^{-4}$ у Катеринопільському районі. У Полтавській області значення варіюють від $3,9 \times 10^{-4}$ у м. Полтаві до $21,2 \times 10^{-4}$ у Чорнухинському районі. Дніпропетровська область характеризується діапазоном від $2,9 \times 10^{-4}$ у м. Дніпрі до $12,1 \times 10^{-4}$ у Петриківському районі, тоді як у Кіровоградській області показник змінюється від $3,6 \times 10^{-4}$ у Кропивницькому до $13,5 \times 10^{-4}$ у Вільшанському районі (рис. 3.9 – 3.11).

Коефіцієнт інбридингу (F) додатково відображає варіабельність на районному рівні. У Вінницькій області значення F змінюється від $11,9 \times 10^{-5}$ у м. Вінниці до $46,1 \times 10^{-5}$ у Мурованокуриловецькому районі. У Черкаській області спостерігається подібний діапазон: від $10,4 \times 10^{-5}$ у м. Черкасах до $40,3 \times 10^{-5}$ у Катеринопільському районі. У Полтавській області коефіцієнт варіює від $9,8 \times 10^{-5}$ у м. Полтаві до 53×10^{-5} у Чорнухинському районі. У Дніпропетровській області F коливається від $7,3 \times 10^{-5}$ у м. Дніпрі до $30,1 \times 10^{-5}$ у Петриківському районі, тоді як у Кіровоградській області показник змінюється від $9,6 \times 10^{-5}$ у м. Кропивницькому до $33,7 \times 10^{-5}$ у Вільшанському районі.

Показник ентропії (H), що характеризує невизначеність у прогнозуванні прізвища особи, є відносно стабільним на обласному рівні: 13,7 у Вінницькій

області, 13,5 у Черкаській, 13,6 у Полтавській, 14,5 у Дніпропетровській та 14 у Кіровоградській.

Водночас на районному рівні значення H демонструють ширшу варіацію: у Вінницькій області — від 10,3 у Чернівецькому районі до 13,2 у м. Вінниці; у Черкаській — від 10,7 у Лисянському районі до 13,4 у м. Черкасах; у Полтавській — від 9,9 у Чорнухинському районі до 13,3 у Кременчуцькому районі; у Дніпропетровській — від 10,8 у Петриківському районі до 14,1 у м. Дніпрі; у Кіровоградській — від 10,6 у Вільшанському районі до 13,4 у м. Кропивницькому. Індекс надмірності розподілу прізвищ (R), що відображає нерівномірність їх розподілу, доповнює ці спостереження. Середні обласні значення становлять 34,4 у Вінницькій області, 34,3 у Черкаській, 34 у Полтавській, 32,8 у Дніпропетровській та 30,7 у Кіровоградській, тоді як районні показники характеризуються відповідною внутрішньорегіональною варіативністю.

Індекс різноманітності прізвищ (α) у районних популяціях є нижчим порівняно з обласним рівнем. У Вінницькій області значення показника варіюють від $5,5 \times 10^2$ у Мурованокуриловецькому районі до $21,2 \times 10^2$ у м. Вінниці. У Черкаській області індекс різноманітності змінюється від $6,3 \times 10^2$ у Катеринопільському районі до $24,2 \times 10^2$ у м. Черкасах, тоді як у Полтавській області він коливається від $4,9 \times 10^2$ у Чорнухинському районі до $25,9 \times 10^2$ у м. Полтаві. У Дніпропетровській області районні значення становлять від $8,6 \times 10^2$ у Петриківському районі до $34,4 \times 10^2$ у м. Дніпрі, а в Кіровоградській області – від $7,7 \times 10^2$ у Вільшанському районі до $26,2 \times 10^2$ у м. Кропивницькому.

Регіональний індекс міграції (ν) у центральних областях варіює у таких межах: у Вінницькій області — $9,9 \times 10^{-4}$, у Черкаській — $13,2 \times 10^{-4}$, у Полтавській — $15,8 \times 10^{-4}$, у Дніпропетровській — $10,9 \times 10^{-4}$, у Кіровоградській — $24,2 \times 10^{-4}$. Водночас на районному рівні значення показника є майже на порядок вищими. У Вінницькій області ν змінюється від $52,4 \times 10^{-4}$ у м. Вінниці до $266,8 \times 10^{-4}$ в Оратівському районі; у Черкаській

— від $77,8 \times 10^{-4}$ у м. Черкаси до $233,8 \times 10^{-4}$ у Христинівському районі; у Полтавській — від $81,8 \times 10^{-4}$ у Кременчуцькому районі до $442,5 \times 10^{-4}$ у Машівському районі; у Дніпропетровській — від $37,3 \times 10^{-4}$ у м. Дніпрі до $575,3 \times 10^{-4}$ у Юр'ївському районі. У Кіровоградській області районні значення індексу міграції коливаються від 98×10^{-4} у м. Кропивницькому до $748,3 \times 10^{-4}$ в Устинівському районі.

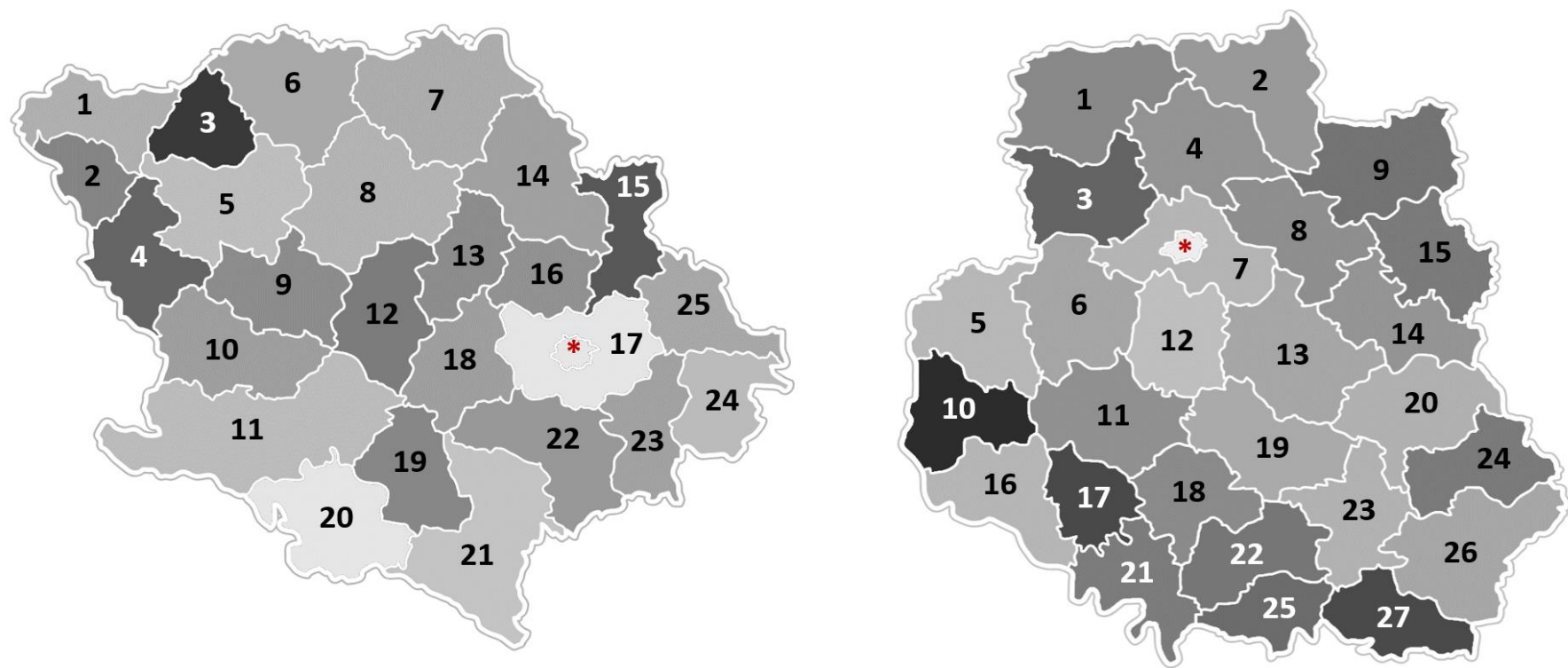


Рис. 3.9. Просторовий розподіл показників ізонімії в районних популяціях Вінницької та Полтавської областей (світлі зони відповідають низьким значенням показника, темні — високим)

Райони Вінницької області: * — м. Вінниця; 1. Хмільницький; 2. Козятинський; 3. Літинський; 4. Калинівський; 5. Барський; 6. Жмеринський; 7. Вінницький; 8. Липовецький; 9. Погребищенський; 10. Мурованокуріловецький; 11. Шаргородський; 12. Тиврівський; 13. Немирівський; 14. Іллінецький; 15. Оратівський; 16. Могилів-Подільський; 17. Чернівецький; 18. Томашпільський; 19. Тульчинський; 20. Гайсинський; 21. Ямпільський; 22. Крижопільський; 23. Тростянецький; 24. Теплицький; 25. Піщанський; 26. Бершадський; 27. Чечельницький.

Райони Полтавської області: * — м. Полтава; 1. Пирятинський; 2. Гребінківський; 3. Чорнухинський; 4. Оржицький; 5. Лубенський; 6. Лохвицький; 7. Гадяцький; 8. Миргородський; 9. Хорольський; 10. Семенівський; 11. Глобинський; 12. Великобагачанський; 13. Шишацький; 14. Зіньківський; 15. Котелевський; 16. Диканський; 17. Полтавський; 18. Решетилівський; 19. Козельщинський; 20. Кременчуцький; 21. Кобеляцький; 22. Новосанжарський; 23. Машівський; 24. Карлівський; 25. Чутівський.

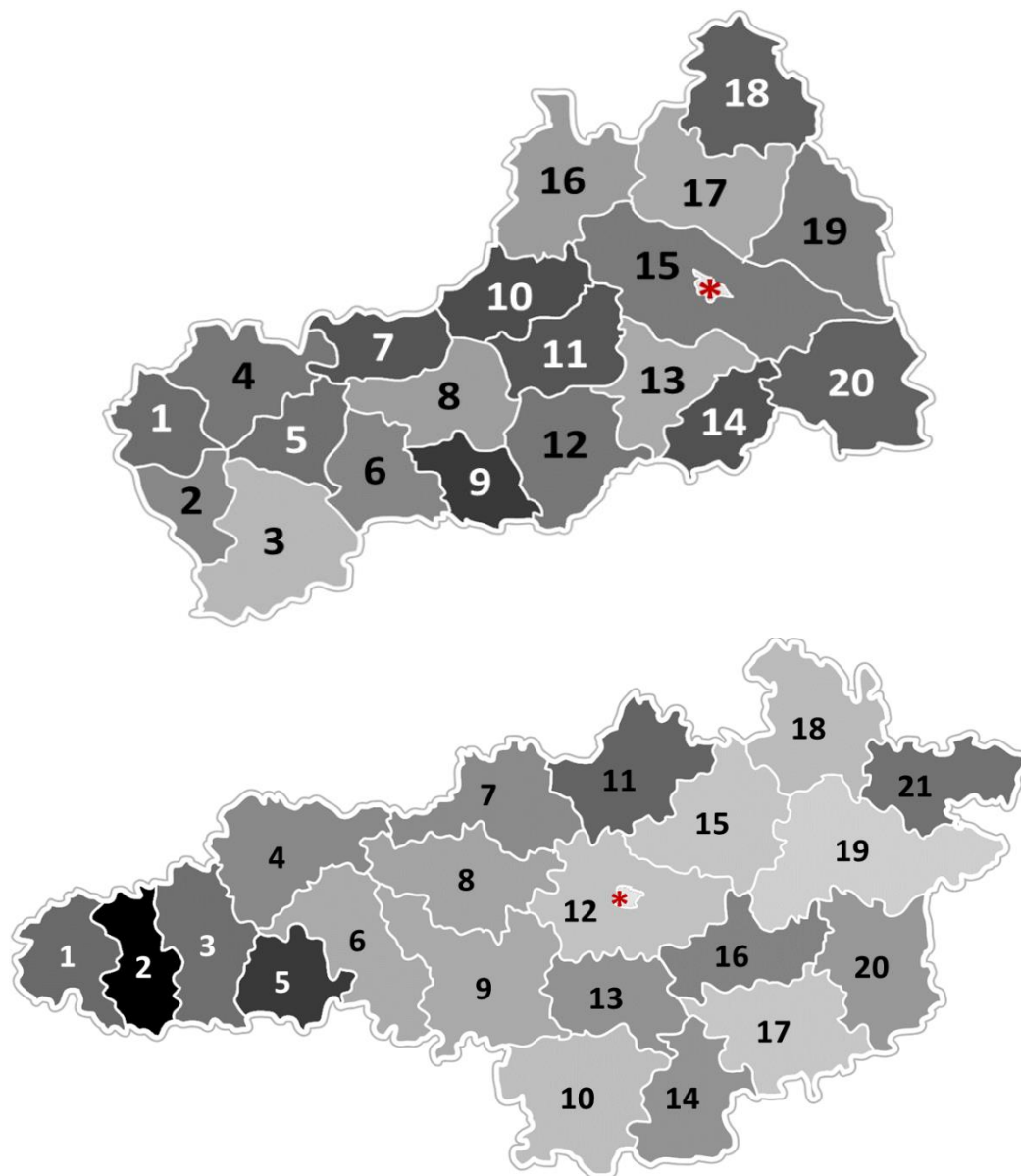


Рис. 3.10. Просторовий розподіл показників ізонімії в районних популяціях Черкаської та Кіровоградської областей (світлі зони відповідають низьким значенням показника, темні — високим)

Райони Черкаської області: * — м. Черкаси; 1. Монастирищенський; 2. Христинівський; 3. Уманський; 4. Жашківський; 5. Маньківський; 6. Тальнівський; 7. Лисянський; 8. Звенигородський; 9. Катеринопільський; 10, Корсунь-Шевченківський; 11. Городищенський; 12. Шполянський; 13. Смілянський; 14. Кам'янський; 15. Черкаський; 16. Канівський; 17. Золотоніський; 18. Драбівський; 19. Чорнобаївський; 20, Чигиринський.

Райони Кіровоградської області: * — м. Кропивницький; 1. Гайворонський; 2. Ульянівський; 3. Голованівський; 4. Новоархангельський; 5. Вільшанський; 6. Добровеличківський; 7. Новомиргородський; 8. Маловисківський; 9. Новоукраїнський; 10, Бобринецький; 11. Олександрівський; 12. Кіровоградський; 13. Компаніївський; 14. Устинівський; 15. Знам'янський; 16. Новгородківський; 17. Долинський; 18. Світловодський; 19. Олександрійський; 20, Петрівський; 21. Онуфріївський.

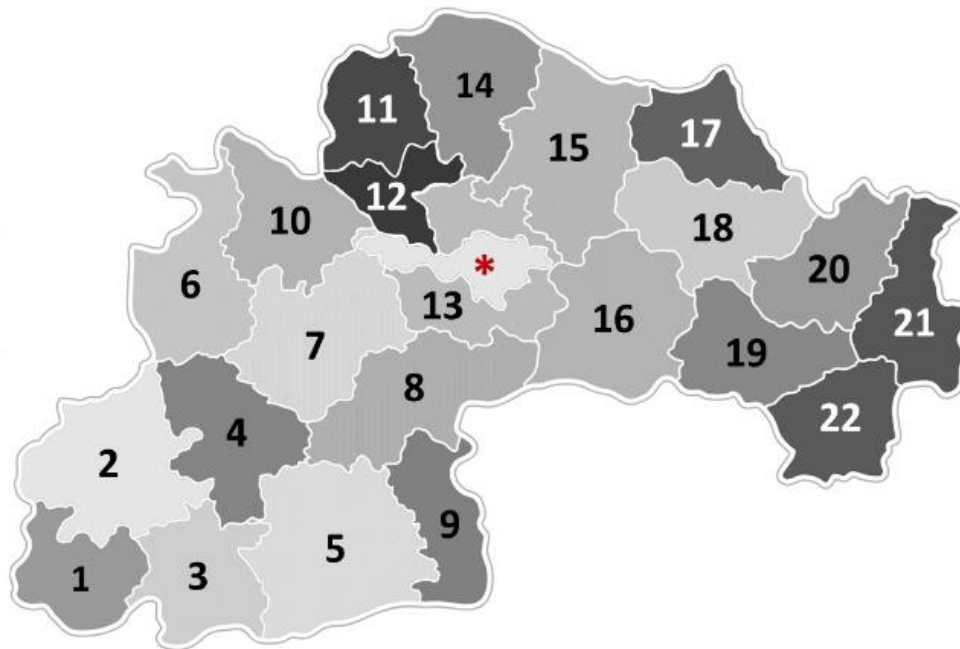


Рис. 3.11. Просторовий розподіл показників ізонімії в районних популяціях Дніпропетровської області (світлі зони відповідають низьким значенням показника, темні — високим)

Райони Дніпропетровської області: * — м. Дніпро; 1. Широківський; 2. Криворізький; 3. Апостолівський; 4. Софіївський; 5. Нікопольський; 6. П'ятихатський; 7. Криничанський; 8. Солонянський; 9. Томаківський; 10. Верхньодніпровський; 11. Царичанський; 12. Петриківський; 13. Дніпропетровський; 14. Магдалинівський; 15. Новомосковський; 16. Синельниківський; 17. Юр'ївський; 18. Павлоградський; 19. Васильківський; 20. Петропавлівський; 21. Межівський; 22. Покровський.

3.3.4. Захід України

Захід України характеризується складною популяційно-генетичною структурою, сформованою під впливом історичних, культурних, соціальних і географічних чинників (рис. 3.12 – 3.14). Аналіз показника внутрішньогрупової ізонімії (I^{jj}) виявляє значну варіабельність як на обласному рівні (Закарпатська — $6,1 \times 10^{-4}$; Чернівецька та Івано-Франківська — $4,6 \times 10^{-4}$; Львівська — $2,3 \times 10^{-4}$; Тернопільська — $3,2 \times 10^{-4}$; Хмельницька — $5,1 \times 10^{-4}$; Волинська — $7,1 \times 10^{-4}$ [123]; Рівненська — $6,8 \times 10^{-4}$), так і на районному рівнях.

У Закарпатській області значення I^{jj} змінюються від $3,6 \times 10^{-4}$ у м. Ужгороді до $47,9 \times 10^{-4}$ у Свалявському районі; у Чернівецькій області показник I^{jj} варіює від $3,4 \times 10^{-4}$ у м. Чернівцях до $36,2 \times 10^{-4}$ у Путивльському

районі.

В Івано-Франківській області I^{jj} коливається від $3,7 \times 10^{-4}$ у м. Івано-Франківську до $36,3 \times 10^{-4}$ у Верховинському районі. У Львівській області значення змінюються від $2,3 \times 10^{-4}$ у м. Львові до $26,5 \times 10^{-4}$ у Турківському районі [125]. У Тернопільській області I^{jj} варіює від 3×10^{-4} у м. Тернополі до $21,4 \times 10^{-4}$ у Шумському районі. У Хмельницькій області показник змінюється від $4,5 \times 10^{-4}$ у м. Хмельницькому до $17,5 \times 10^{-4}$ у Славутському районі. У Волинській області значення I^{jj} становлять від $6,3 \times 10^{-4}$ у м. Луцьку до $40,5 \times 10^{-4}$ у Шацькому районі [123], тоді як у Рівненській області — від $6,3 \times 10^{-4}$ у м. Рівному до $98,5 \times 10^{-4}$ у Заріченському районі.

Коефіцієнт інбридингу (F) у Західній Україні також демонструє значну варіабельність. У Закарпатській області значення F змінюються від $8,9 \times 10^{-5}$ в м. Ужгороді до 120×10^{-5} у Свалявському районі, що свідчить про суттєво вищий рівень спорідненості в більш ізольованих територіях. Подібні закономірності простежуються і в інших західних регіонах. У Чернівецькій області F варіює від $8,4 \times 10^{-5}$ у м. Чернівцях до $90,6 \times 10^{-5}$ у Путильському районі; в Івано-Франківській області — від $9,2 \times 10^{-5}$ у м. Івано-Франківську до $90,8 \times 10^{-5}$ у Верховинському районі; у Львівській області — від $5,7 \times 10^{-5}$ у м. Львові до $66,1 \times 10^{-5}$ у Турківському районі [125]. У Тернопільській області показник змінюється від $7,6 \times 10^{-5}$ у м. Тернополі до $53,6 \times 10^{-5}$ у Шумському районі; у Хмельницькій — від $11,2 \times 10^{-5}$ у м. Хмельницькому до $43,8 \times 10^{-5}$ у Славутському районі; у Волинській — від $15,8 \times 10^{-5}$ у м. Луцьку до $101,3 \times 10^{-5}$ у Шацькому районі [123]; у Рівненській області — від $15,8 \times 10^{-5}$ у м. Рівному до $246,2 \times 10^{-5}$ у Заріченському районі. Гірські райони стабільно демонструють вищі коефіцієнти інбридингу, що зумовлено природними бар'єрами, які обмежують генетичний потік. Натомість міські центри характеризуються нижчими значеннями показника внаслідок інтенсивної міграції та ширших соціокультурних контактів.

Показники розподілу прізвищ додатково підкреслюють регіональну неоднорідність Західної України. Індекс ентропії (H), який кількісно відображає невизначеність прогнозування прізвища окремої особи, набуває найвищих значень у Львівській області (приблизно 14,2) [125], тоді як найнижчі значення зафіксовано в Закарпатській області (близько 12,8). На районному рівні показник варіює в широких межах.

Міські центри та прикордонні райони, як правило, характеризуються підвищеними значеннями H , що відображає більшу антропонімічну та генетичну різноманітність, зумовлену історичними міграційними процесами та міжкультурною взаємодією. Натомість ізольовані гірські території демонструють нижчі значення ентропії, що свідчить на більш однорідний розподіл прізвищ і відносну демографічну замкнутість популяцій.

Значення індексу надлишковості прізвищ (R) у західних областях України також демонструють суттєву внутрішньорегіональну варіабельність. У Закарпатській області показник коливається від 22,6 у м. Ужгород до 42,8 у Тячівському районі; у Чернівецькій області — від 25,7 у м. Чернівці до 36,5 у Путивльському районі; в Івано-Франківській області — від 27,7 у м. Івано-Франківськ до 40,1 у Верховинському районі; у Львівській області — від 25,8 у Пустомитівському районі до 36,2 у Турківському районі [125]. У Тернопільській області R варіює від 24,4 у Тернопільському районі до 32,6 у Шумському районі; у Хмельницькій області — від 25,3 у Деражнянському районі до 31,7 у Славутському районі; у Волинській області — від 27,2 у Володимир-Волинському районі до 41 у Любешівському районі [123]; у Рівненській області — від 27,8 у Дубенському районі до 43,9 у Рокитнівському районі. Індекс різноманітності прізвищ (α) у Західній Україні демонструє чітке зниження на районному рівні порівняно з обласними центрами. Зокрема, у Закарпатській області районні значення варіюють від $2,1 \times 10^2$ у Свалявському районі до $28,6 \times 10^2$ у м. Ужгороді. Аналогічні закономірності простежуються в Чернівецькій, Івано-Франківській, Львівській, Волинській, Тернопільській, Хмельницькій та Рівненській

областях, де міські центри стабільно характеризуються вищим рівнем антропомічного розмаїття порівняно з периферійними районами.

Індекс міграції (v) суттєво перевищує регіональні значення на районному рівні, що підкреслює динамічний вплив сучасних міграційних процесів на структуру локальних популяцій. У Закарпатській області v коливається від $20,5 \times 10^{-4}$ у Тячівському районі до $242,6 \times 10^{-4}$ у м. Ужгороді. Подібні диспропорції зафіксовано й в інших регіонах: у Чернівецькій області — від $81,1 \times 10^{-4}$ у Сторожинецькому районі до $168,7 \times 10^{-4}$ у Герцаївському районі; в Івано-Франківській області — від $49,8 \times 10^{-4}$ у Надвірнянському районі до $183,5 \times 10^{-4}$ у Рогатинському районі; у Львівській області — від $62,3 \times 10^{-4}$ у м. Львові до $212,9 \times 10^{-4}$ у Перемишлянському районі [125]. У Тернопільській області показник варіює від $133,3 \times 10^{-4}$ у Шумському районі до $406,1 \times 10^{-4}$ у Підгаєцькому районі; у Хмельницькій області — від $74,5 \times 10^{-4}$ у Славутському районі до $293,7 \times 10^{-4}$ у Старосинявському районі; у Волинській області — від 55×10^{-4} у Ратнівському районі до $232,9 \times 10^{-4}$ у Турійському районі [123]; у Рівненській області — від $32,4 \times 10^{-4}$ у Зарічненському районі до $264,4 \times 10^{-4}$ у Демидівському районі.

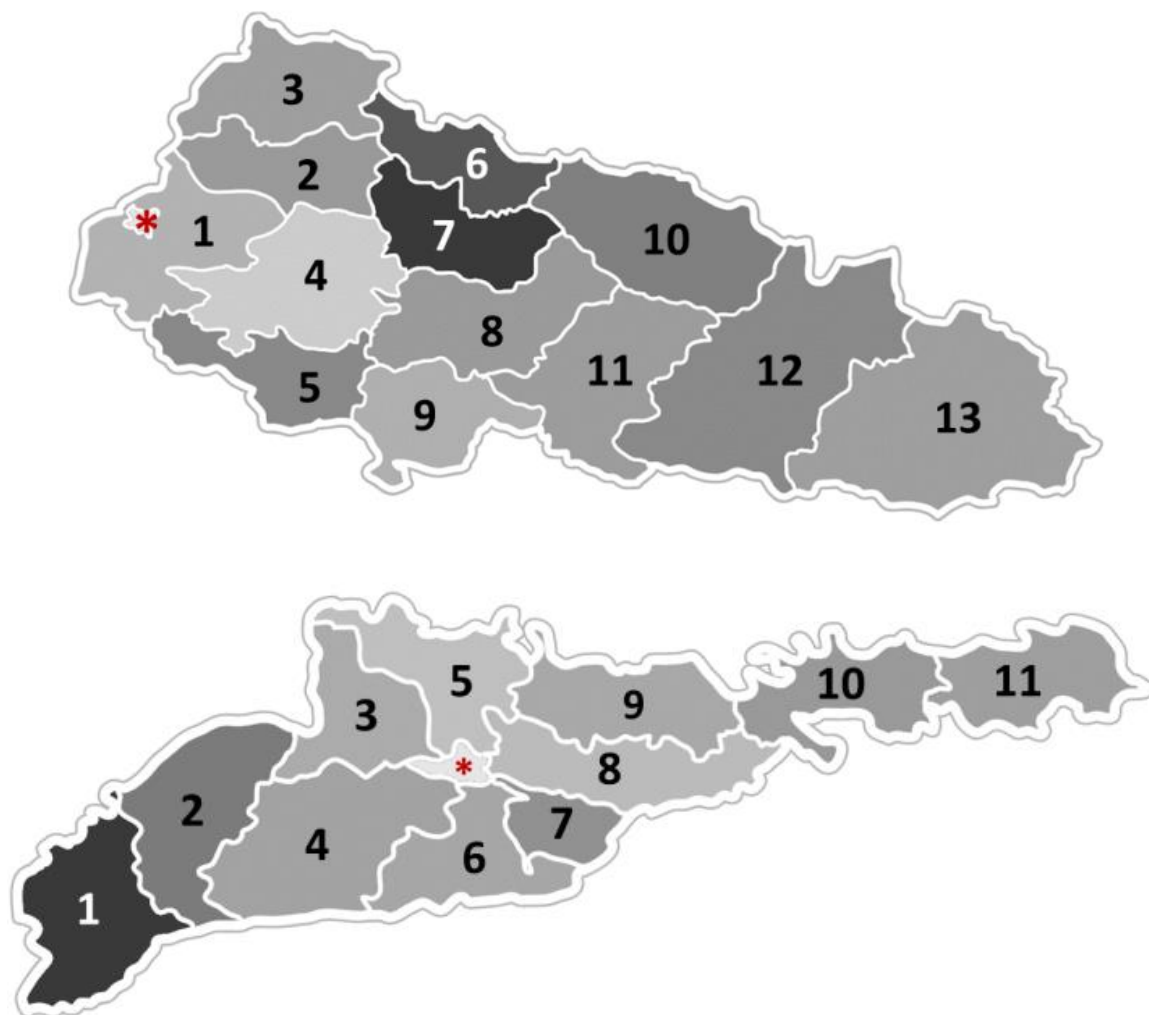


Рис. 3.12. Просторовий розподіл показників ізонімії в районних популяціях Закарпатської та Чернівецької областей (світлі зони відповідають низьким значенням показника, темні — високим)

Райони Черкаської області: * — м. Ужгород; 1. Ужгородський; 2. Перечинський; 3. Великоберезнянський; 4. Мукачівський; 5. Берегівський; 6. Воловецький; 7. Свалявський; 8. Іршавський; 9. Виноградівський; 10, Міжгірський; 11. Хустський; 12. Тячівський; 13. Рахівський.

Райони Чернівецької області: * — м. Чернівці; 1. Путильський; 2. Вижницький; 3. Кіцманський; 4. Сторожинецький; 5. Заставнівський; 6. Глибоцький; 7. Герцаївський; 8. Новоселицький; 9. Хотинський; 10, Кельменецький; 11. Сокирянський.

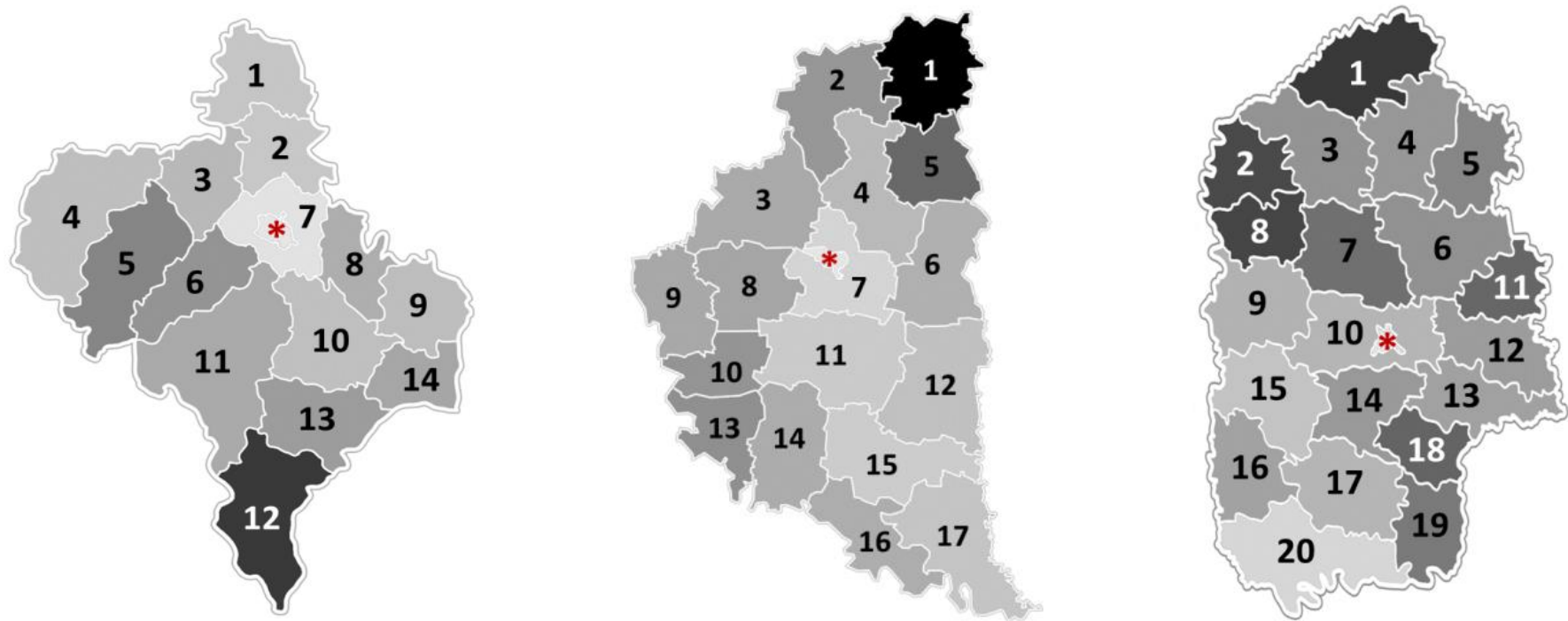


Рис. 3.13. Просторовий розподіл показників ізонімії в районних популяціях Івано-Франківської, Тернопільської та Хмельницької областей (світлі зони відповідають низьким значенням показника, темні — високим)

Райони Івано-Франківської області: * — м. Івано-Франківськ; 1. Рогатинський; 2. Галицький; 3. Калуський; 4. Долинський; 5. Рожнятівський; 6. Богородчанський; 7. Тисменицький; 8. Тлумацький; 9. Городенківський; 10. Коломийський; 11. Надвірнянський; 12. Верховинський; 13. Косівський; 14. Снятинський.

Райони Тернопільської області: * — м. Тернопіль; 1. Шумський; 2. Кременецький; 3. Зборівський; 4. Збарзький; 5. Лановецький; 6. Підволочиський; 7. Тернопільський; 8. Козівський; 9. Бережанський; 10. Підгаєцький; 11. Терехівський; 12. Гусятинський; 13. Монастирський; 14. Бучацький; 15. Чортківський; 16. Заліщицький; 17. Борщівський;

Райони Хмельницької області: * — м. Хмельницький; 1. Славутський; 2. Білогірський; 3. Ізяславський; 4. Шепетівський; 5. Полонський; 6. Старокостянтинівський; 7. Красилівський; 8. Теофіпольський; 9. Волочеський; 10. Хмельницький; 11. Старосинявський; 12. Летичівський; 13. Деражнянський; 14. Ярмолинецький; 15. Городецький; 16. Чемеровецький; 17. Дунаєвецький; 18. Вінковоцький; 19. Новоушицький; 20. Кам'янець-Подільський.

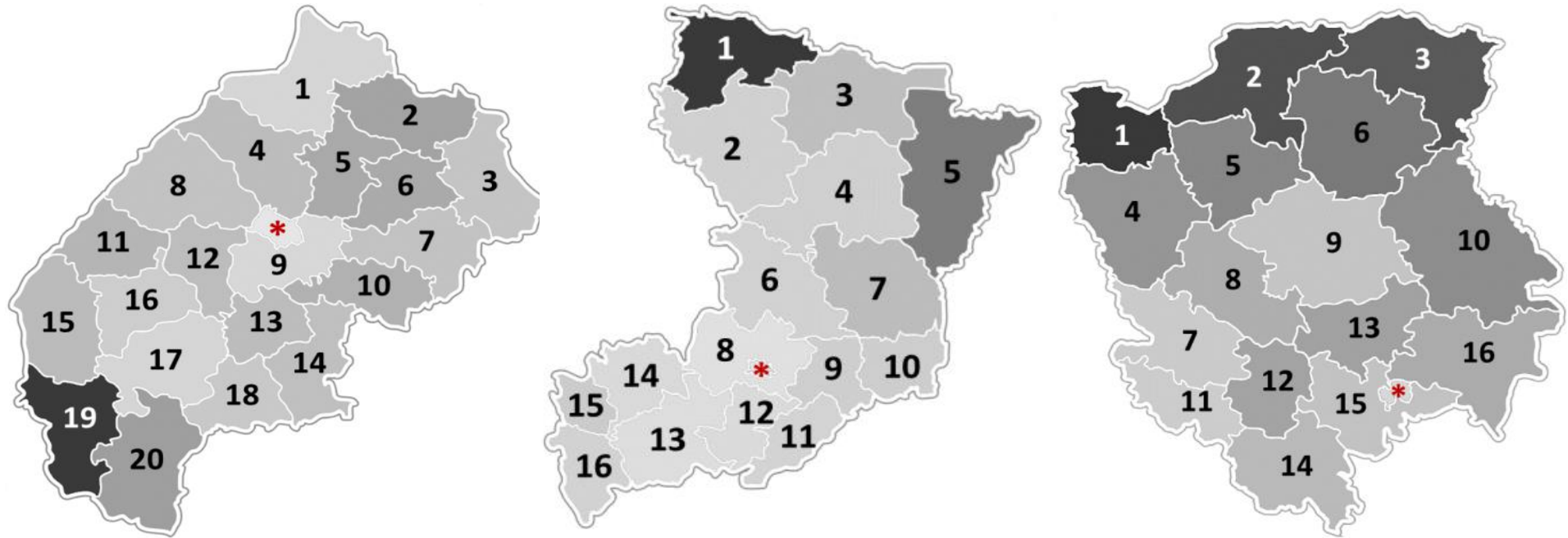


Рис. 3.14. Просторовий розподіл показників ізонімії в районних популяціях Львівської, Рівненської та Волинської областей (світлі зони відповідають низьким значенням показника, темні — високим)

Райони Львівської області: * — м. Львів; 1. Сокальський; 2. Радехівський; 3. Бродівський; 4. Жовківський; 5. Кам'янка-Бузький; 6. Буський; 7. Золочівський; 8. Яворський; 9. Пустомитівський; 10. Перемишлянський; 11. Мостиський; 12. Городоцький; 13. Миколаївський; 14. Жидачівський; 15. Старосамбірський; 16. Самбірський; 17. Дрогобицький; 18. Стрийський; 19. Турківський; 20. Сколівський.

Райони Рівненської області: * — м. Рівне; 1. Зарічненський; 2. Володимирецький; 3. Дубровицький; 4. Сарненський; 5. Рокитнівський; 6. Костопільський; 7. Березнівський; 8. Рівненський; 9. Гошанський; 10. Корецький; 11. Острозький; 12. Здолбунівський; 13. Дубенський; 14. Млинівський; 15. Демидівський; 16. Радивилівський.

Райони Волинської області: * — м. Луцьк; 1. Шацький; 2. Ратнівський; 3. Любешівський; 4. Любомльський; 5. Старовижівський; 6. Камінь-Каширський; 7. Володимир-Волинський; 8. Турійський; 9. Ковельський; 10. Маневицький; 11. Іваничівський; 12. Локачинський; 13. Рожищенський; 14. Горохівський; 15. Луцький; 16. Ківерцівський.

3.3.5. Південь України

Південь України характеризується специфічною структурою популяційно-генетичних показників, які, попри відносну загальну однорідність антропонімичного ландшафту, демонструють суттєві локальні варіації (рис. 3.15, 3.16). Показник внутрішньогрупової ізонімії (I^{jj}) у регіоні загалом є низьким (Одеська область — $2,1 \times 10^{-4}$, Миколаївська — $2,4 \times 10^{-4}$, Херсонська — $2,6 \times 10^{-4}$, Запорізька — $2,5 \times 10^{-4}$ [124], АР Крим — $1,6 \times 10^{-4}$); однак на районному рівні спостерігаються виразні відхилення.

Зокрема, в Одеській області I^{jj} варіює від $2,1 \times 10^{-4}$ у м. Одесі до $18,6 \times 10^{-4}$ в Арцизькому районі. Аналогічні диспропорції фіксуються в інших південних областях: у Миколаївській області показник становить $2,6 \times 10^{-4}$ у м. Миколаєві проти 11×10^{-4} у Врадіївському районі; у Херсонській області — від $2,8 \times 10^{-4}$ у м. Херсоні до $12,4 \times 10^{-4}$ у Верхньорогачицькому районі; у Запорізькій області — від $2,6 \times 10^{-4}$ у м. Запоріжжі до 19×10^{-4} у Великобілозерському районі [124]; в Автономній республіці Крим — від $1,6 \times 10^{-4}$ у м. Сімферополі до $3,4 \times 10^{-4}$ у Білогірському районі.

Коефіцієнт інбридингу (F) також демонструє виразну внутрішньорегіональну мінливість. В Одеській області F варіює від $5,2 \times 10^{-5}$ в м. Одесі до $46,5 \times 10^{-5}$ в Арцизькому районі. Подібні закономірності спостерігаються і в Миколаївській області — від $6,5 \times 10^{-5}$ у м. Миколаєві до $27,5 \times 10^{-5}$ у Врадіївському районі; у Херсонській області — від $6,9 \times 10^{-5}$ у м. Херсоні до $30,9 \times 10^{-5}$ у Верхньорогачицькому районі; у Запорізькій області — від $6,6 \times 10^{-5}$ у м. Запоріжжі до $47,6 \times 10^{-5}$ у Великобілозерському районі [124]; в Автономній Республіці Крим — від $3,4 \times 10^{-5}$ у Джанкойському районі до $8,5 \times 10^{-5}$ у Білогірському районі.

Метрики поширення прізвищ додатково конкретизують регіональні особливості Півдня України. Показник ентропії (H) демонструє значну варіабельність: в Одеській області його значення становить 14,6 у місті Одесі та 13,2 в Комінтернівському районі. У Миколаївській області H коливається

від 10,9 у Врадіївському районі до 13,2 у Вознесенському районі; у Херсонській області — від 10,7 у Верхньорогачицькому районі до 13,9 у м. Херсоні; у Запорізькій області — від 10,1 у Великобілозерському районі до 14,1 у м. Запоріжжі [124]. В Автономній Республіці Крим значення H змінюється від 1,7 у м. Сімферополі до 9,3 у Роздольненському районі. Індекс надлишковості прізвищ (R) також характеризується суттєвими коливаннями. В Одеській області він варіює від 17,3 в м. Іллічівську до 34,4 в Болградському районі. Подібні диспропорції зафіксовано у Миколаєвській області — від 17,9 у Веселинівському районі до 26,6 у Березнегуватському; у Херсонській області — від 20,1 у Генічеському районі до 24,3 у Цюрупинському районі; у Запорізькій області — від 17,2 у Розівському районі проти 28,5 у м. Запоріжжі [124]. В Автономній Республіці Крим максимальне значення R зафіксовано в м. Сімферополі (90,9), мінімальне — у Роздольненському районі (38).

Різноманітність прізвищ, вимірювана індексом різноманітності (α), має тенденцію до зниження на районному рівні порівняно з обласними центрами. В Одеській області значення α варіює від $5,4 \times 10^2$ в Арцизькому районі до $48,6 \times 10^2$ в м. Одесі. Подібні закономірності простежуються і в інших південних областях: у Миколаївській області — від $9,5 \times 10^2$ у Врадіївському районі до $38,8 \times 10^2$ у м. Миколаєві; у Херсонській області — від $8,5 \times 10^2$ у Верхньорогачицькому районі до $36,7 \times 10^2$ у м. Херсоні; у Запорізькій області — $5,5 \times 10^2$ у Великобілозерському районі до $37,8 \times 10^2$ у м. Запоріжжі [124]; в Автономній Республіці Крим — від $30,8 \times 10^2$ у Білогірському районі до $63,2 \times 10^2$ у м. Сімферополі.

Індекс міграції (ν) на Півдні України демонструє ще більш виражені розбіжності між районними та регіональними значеннями. В Одеській області показник коливається від $51,8 \times 10^{-4}$ у м. Одесі до $633,8 \times 10^{-4}$ у Великомихайлівському районі. Подібні закономірності зафіксовано в Миколаєвській області — від $87,7 \times 10^{-4}$ у м. Миколаєві проти $853,5 \times 10^{-4}$ у Веселенівському районі; у Херсонській області — від $119,1 \times 10^{-4}$

у м. Херсоні до 663×10^{-4} у Високопільському районі; у Запорізькій області — від $42,8 \times 10^{-4}$ у м. Запоріжжі до $1168,6 \times 10^{-4}$ у Розівському районі [124]; в Автономній Республіці Крим — від $166,3 \times 10^{-4}$ у м. Сімферополі до $1134,5 \times 10^{-4}$ у Роздольненському районі.

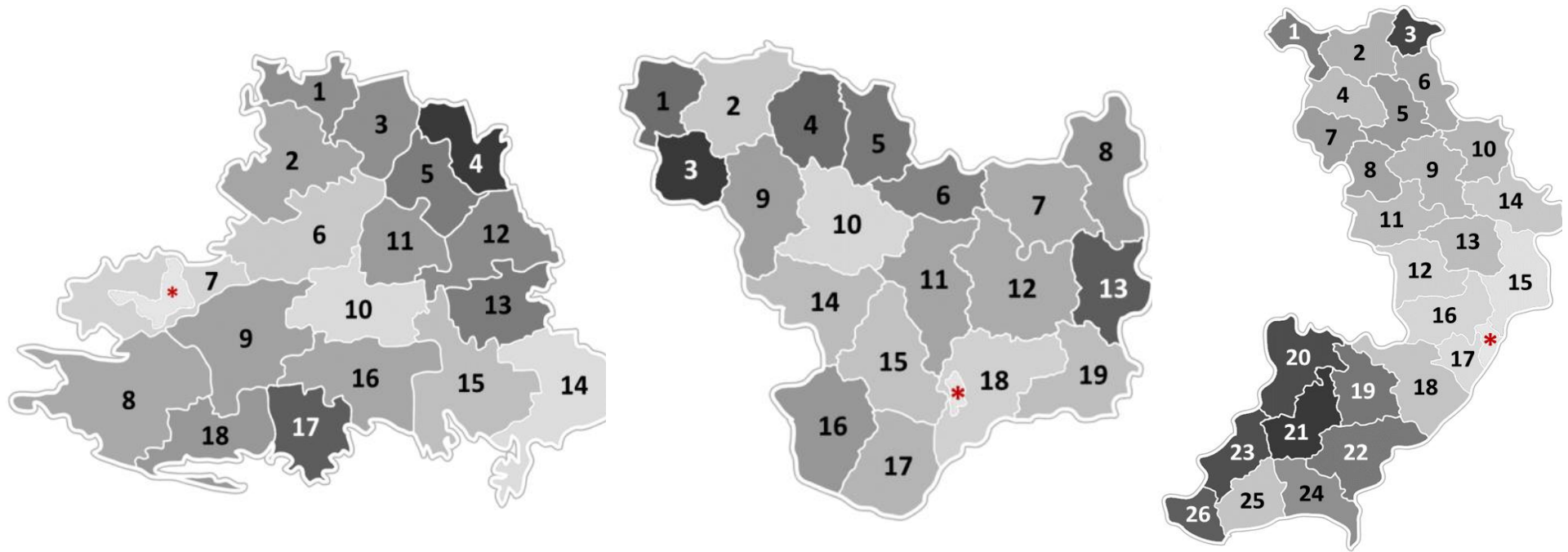


Рис. 3.15. Просторовий розподіл показників ізонімії в районних популяціях Миколаївської, Херсонської та Одеської областей (світлі зони відповідають низьким значенням показника, темні — високим)

Райони Миколаївської області: * — м. Миколаїв; 1. Кривоозерський; 2. Первомайський; 3. Врадіївський; 4. Арбузинський; 5. Братський; 6. Сланецький; 7. Новобузький; 8. Казанківський; 9. Доманівський; 10. Вознесенський; 11. Новоодеський; 12. Баштанський; 13. Березнегуватський; 14. Веселинівський; 15. Миколаївський; 16. Березанський; 17. Очаківський; 18. Жовтневий; 19. Снігурівський.

Райони Херсонської області: * — м. Херсон; 1. Високопільський; 2. Великоолександрівський; 3. Нововоронцовський; 4. Верхньорогачицький; 5. Великопетиський; 6. Бериславський; 7. Білозерський; 8. Голопристанський; 9. Цюрупинський; 10. Каховський; 11. Горностаївський; 12. Нижньосірогозький; 13. Іванівський; 14. Генічеський; 15. Новотроїцький; 16. Чаплинський; 17. Каланчацький; 18. Скадовський.

Райони Одеської області: * — м. Одеса; 1. Кодимський; 2. Балтський; 3. Савранський; 4. Подільський; 5. Ананьївський; 6. Любашівський; 7. Окнянський; 8. Захарівський; 9. Ширяївський; 10. Миколаївський; 11. Великомихайлівський; 12. Роздільнянський; 13. Іванівський; 14. Березівський; 15. Лиманський; 16. Біляївський; 17. Овідіопольський; 18. Білгород-Дністровський; 19. Саратський; 20. Тарутинський; 21. Арцизький; 22. Татарбунарський; 23. Болградський; 24. Кілійський; 25. Ізмаїльський; 26. Ренійський.

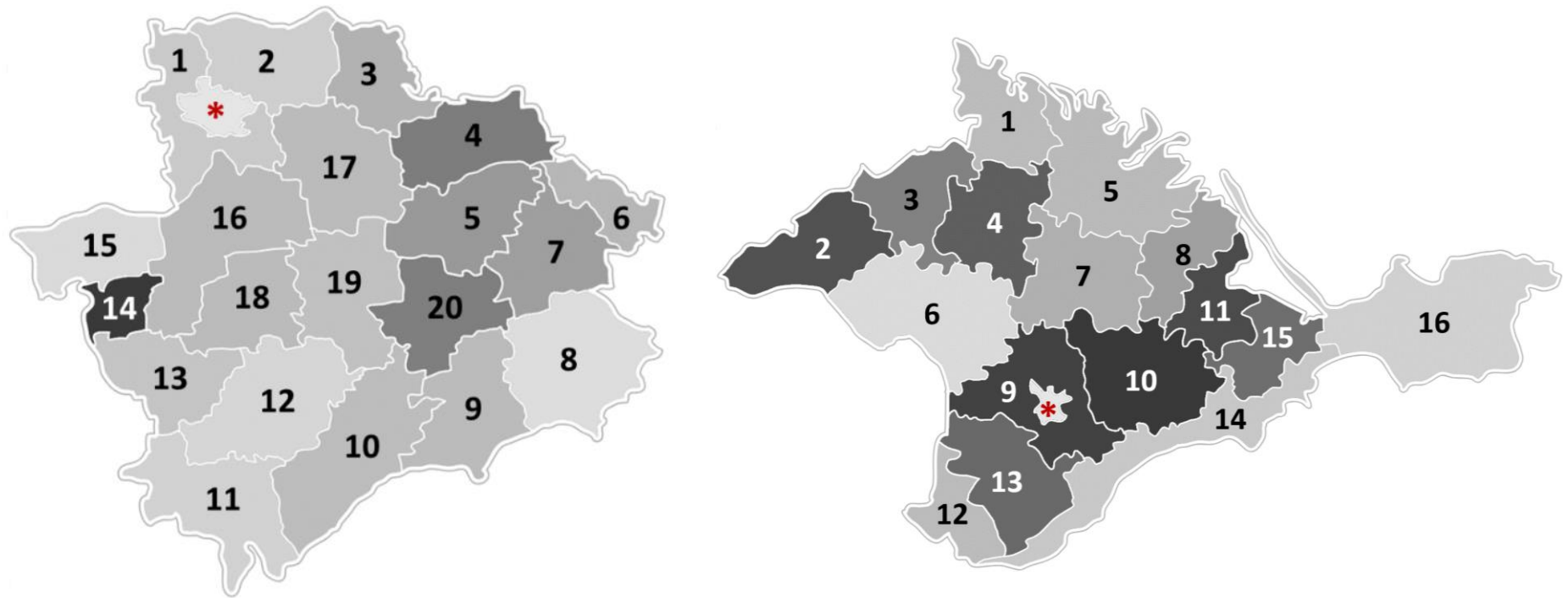


Рис. 3.16. Просторовий розподіл показників ізонімії в районних популяціях Запорізької області та Кримського півострова (світлі зони відповідають низьким значенням показника, темні — високим)

Райони Вінницької області: * — м. Запоріжжя; 1. Запорізький; 2. Вільнянський; 3. Новомиколаївський; 4. Гуляйпільський; 5. Пологівський; 6. Розівський; 7. Куйбишевський; 8. Бердянський; 9. Приморський; 10. Приазовський; 11. Якимівський; 12. Мелітопольський; 13. Веселівський; 14. Великобілозерський; 15. Кам'янсько-Дніпровський; 16. Василівський; 17. Оріхівський; 18. Михайлівський; 19. Токмацький; 20. Чернігівський;

Райони Кримського півострову: * — м. Сімферополь; 1. Красноперекопський; 2. Чорноморський; 3. Роздольненський; 4. Первомайський; 5. Джанкойський; 6. Сакський; 7. Красногвардійський; 8. Нижньогірський; 9. Сімферопольський; 10. Білогірський; 11. Советський; 12. м. Севастополь; 13. Бахчисарайський; 14. Яли Бою; 15. Кіровський; 16. Ленінський.

3.4. Міжпопуляційні генетичні дистанції та кластерний аналіз

Застосування методу Уорда до матриць генетичних дистанцій між 490 районами України виявило складну ієрархічну структуру, конфігурація якої варіює залежно від обраної метрики. Оптимальна кількість кластерів становить 11 для евклідової дистанції та дистанції Нея і 10 — для дистанції Ласкера, що свідчить про високу, хоча й не абсолютну, топологічну стабільність кластерної архітектури.

На попередньому етапі аналізу хордову відстань (D_{CSE}^{ij}) було включено до розрахунку матриць попарних дистанцій, однак отримані результати кластеризації були близькими до результатів, одержаних за стандартною евклідовою відстанню, і не змінювали основної інтерпретації структури даних. З огляду на це у фінальному аналізі як основну метрику було використано евклідову відстань (D_E^{ij}).

3.4.1. Кластерна структура за евклідовою дистанцією

Кластеризація на основі евклідової дистанції виявила 11 чітко окреслених кластерів з розміром від 5 до 107 районів, що відображає значну ієрархічну складність структури (Рис. 3.17, 3.18).

Кластер 1 (5 районів, монообласний) включає виключно райони Одеської області: Саратський, Ренійський, Болградський, Арцизький, Тарутинський. Це найменший кластер у всій системі. Його виразна ізоляція від решти області (інші 21 район Одещини розподілені між кластерами 5 та 10) відображає високий рівень внутрішньообласної гетерогенності Одеської області — найбільшої серед усіх областей України.

Зазначені п'ять районів формують компактний південно-західний блок Одещини з унікальною демографічною історією. Ренійський район належить до Придунав'я і зазнав інтенсивного молдавсько-румунського впливу внаслідок близькості до дельти Дунаю. Болградський, Арцизький і Тарутинський райони входять до зони болгарської колонізації початку ХІХ століття: після Російсько-турецької війни 1806–1812 років відбулося масове

переселення болгар до Бессарабії, де сформувалися компактні етнічні поселення на півдні Одещини [46, 82, 93]. Саратовський район також характеризується значною болгарською присутністю та розташований у південно-західній периферії області. Генетична ізоляція цієї групи районів відображає поєднання трьох чинників: етнічної специфіки (болгарсько-молдавський субстрат), географічної периферійності (прикордонне розташування з Молдовою та Румунією) та історичні відокремленості (Придунав'я та Буджак мали окрему адміністративну історію у складі Бессарабської губернії).

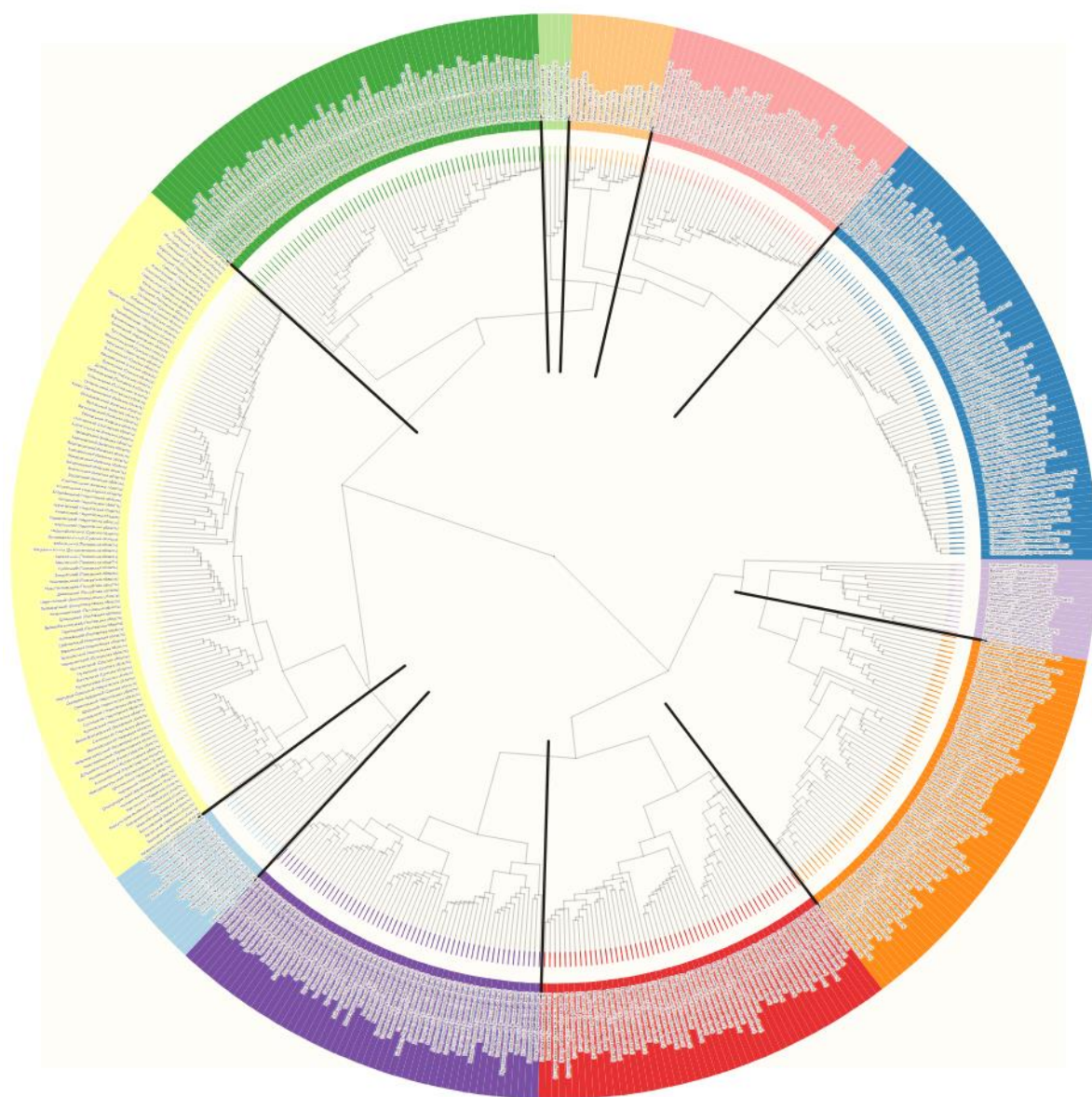


Рис. 3.17. Дендрограма ієрархічної кластеризації 490 районів України, побудована методом Уорда за евклідовою дистанцією. Інтерактивна версія доступна за [ПОСИЛАННЯМ](#).

Кластер 2 (63 райони, 4 області) об'єднує переважно донбаські та слобожанські райони: Луганську (18 районів), Харківську (26 районів), Донецьку області (18 районів) та 1 район Сумської області. Це найбільш компактний східний індустріальний кластер, що відображає спільну історію формування під впливом індустріалізації XIX–XX століть [46, 82, 93]. Включення всіх 18 районів Луганської та всіх 18 районів Донецької областей підтверджує винятково високий рівень генетичної спорідненості Донбасу як історично монопромислового регіону з інтенсивними міграційними потоками робітничого населення. Повна представленість Харківської області (26 районів), засвідчує її глибоку інтеграцію до східного індустріального блоку, попри подвійність історичного розвитку — як частини Слобожанщини та як потужного освітньо-наукового центру. Натомість включення лише одного району Сумської області (з 18 загалом) свідчить про відносну відокремленість більшості території Сумщини від східного індустріального ядра. Це узгоджується з її історичною роллю транзитної зони між Лівобережною Україною та Донбасом і пояснює більшу спорідненість Сумської області з північно-центральними областями [46, 82, 93].

Кластер 3 (107 районів, 8 областей) є найбільшим у структурі та об'єднує переважно центральні й частково північні області України: Черкаську (14 з 20 районів), Київську (19 з 25), Кіровоградську (7 з 21), Запорізьку (1 з 20), Чернігівську (22 з 22), Сумську (17 з 18), Полтавську (24 з 25) та Дніпропетровську (3 з 22) області. Цей мегакластер відображає північно-центральну та частково лівобережну філогенетичну спільність. Домінування Сумської (17 з 18), Полтавської (24 з 25) та повна представленість Чернігівської області (22 з 22) підтверджує їхню належність до лівобережно-поліського антропонімічного ареалу. Київська область представлена 19 з 25 районів, що свідчить про значну внутрішньообласну гетерогенність столичного регіону: решта шість районів розподілені між кластерами 4, 5 та 6. Часткова інтеграція Запорізької (1 район) і Дніпропетровської (3 райони) областей указує на їхню переважну належність до південно-східного

індустріального блоку, тоді як окремі периферійні райони демонструють спорідненість із центральним кластером.

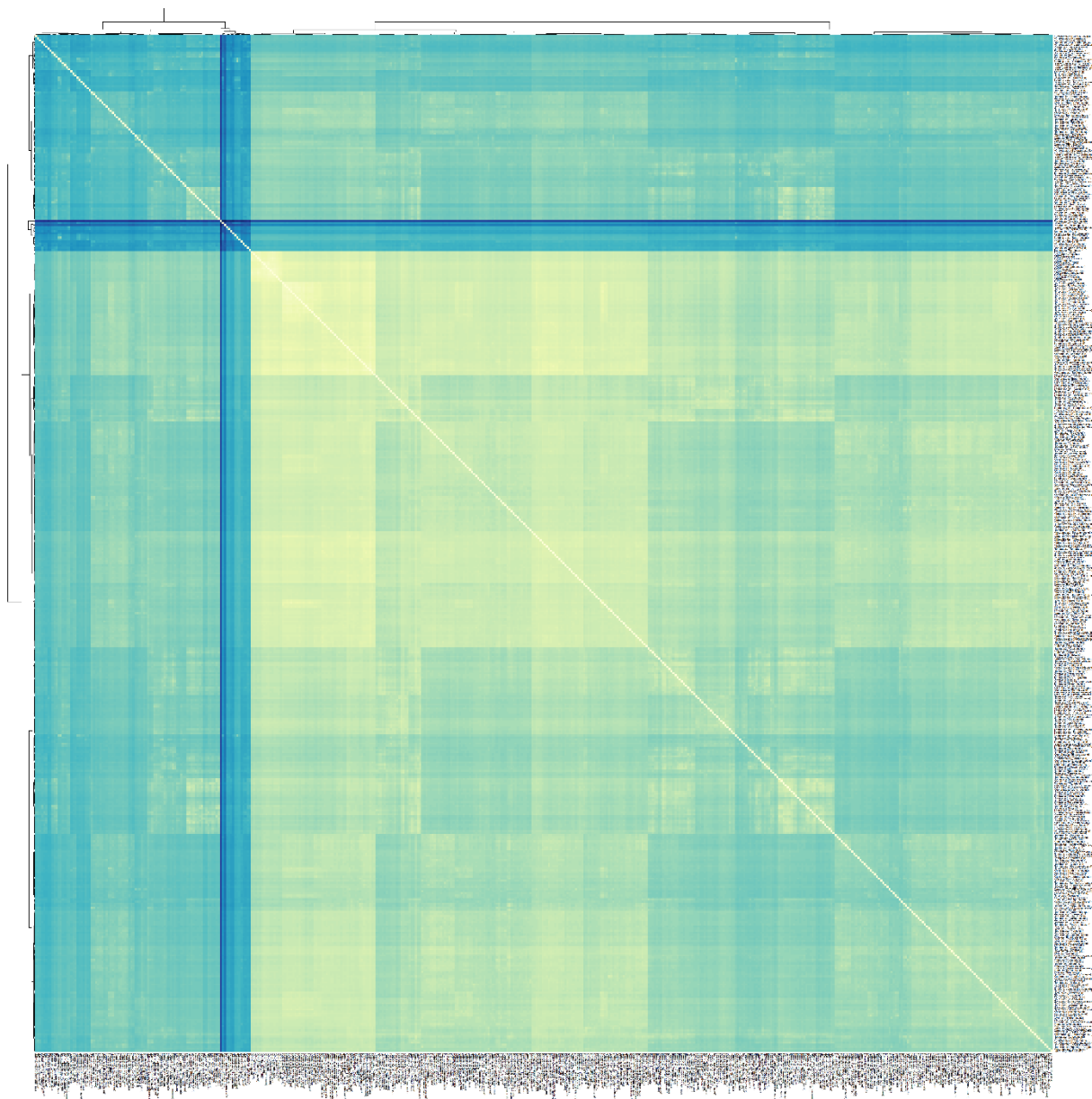


Рис. 3.18. Теплова карта матриці евклідових дистанцій між 490 районами України з відображенням кластерної структури (світлі зони відповідають низьким значенням, темні — високим).

Кластер 4 (15 районів, 3 області) об'єднує переважно поліські території: Рівненську (1 район), Київську (2 райони) та Житомирську (12 з 23 районів) області. Це чітко окреслений поліський кластер із домінуванням Житомирського Полісся. Включення 12 із 23 районів Житомирської області підтверджує виразну поліську ідентичність Житомирщини в межах кластерної структури. Натомість мінімальна присутність Рівненської (1 з 16 районів)

та Київської (2 з 25 районів) областей свідчить про те, що до цього кластера увійшли лише їхні найбільш типові поліські території: Зарічненський район (Рівненська область), а також Поліський і Іванківський райони (Київська область).

Кластер 5 (56 районів, 8 областей) об'єднує переважно подільські та частково центральні області: Чернівецьку (3 з 11 районів), Вінницьку (27 з 27), Хмельницьку (11 з 20), Київську (3 з 25), Кіровоградську (3 з 21), Миколаївську (1 з 19), Черкаську (6 з 20) та Одеську (2 з 26) області. Повна представленість Вінницької області (27 з 27 районів) підтверджує її роль як ядра подільського кластера. Хмельницька область розділена між кластерами 5 (11 районів) і 6 (9 районів), що відображає її транзитне положення між Поділлям і Волинсько-Поліським ареалом. Часткова присутність Київської, Черкаської та Кіровоградської областей засвідчує їхній перехідний статус між центральною та подільською антропонімічними зонами.

Кластер 6 (53 райони, 6 областей) об'єднує волинсько-поліські та частково західні території: Хмельницьку (9 з 20 районів), Житомирську (11 з 23), Рівненську (13 з 16), Тернопільську (3 з 17), Київську (1 з 25) та Волинську (16 з 16) області. Волинська область представлена повністю, що підтверджує її монолітну волинську ідентичність у межах кластерної структури. Рівненська область представлена 13 з 16 районів (81%), що вказує на її чітку належність до волинсько-поліського ареалу, за винятком трьох екстремальних аутлаєрів — Рокитнівського, Дубровицького та Зарічненського районів. Житомирська та Хмельницька області також демонструють поділ між сусідніми кластерами, що підкреслює їхнє прикордонне географічне й антропонімічне положення.

Кластер 7 (57 районів, 5 областей) є чітко окресленим галицько-буковинським кластером. До його складу входять Тернопільська (14 з 17 районів), Львівська (20 з 20), Івано-Франківська (14 з 14), Рівненська (1 з 16) та Чернівецька (8 з 11) області. Повна представленість Львівської та Івано-Франківської областей підтверджує їхню монолітну галицьку

ідентичність. Це єдині дві великі області (поряд із відносно компактними Волинською та Вінницькою), які входять до складу одного кластера повністю. Тернопільська область представлена 14 з 17 районів (82 %), що засвідчує її переважно галицьку природу. Чернівецька область розділена між кластерами 5 (3 райони) і 7 (8 районів), що відображає її подвійне буковинсько-подільське походження.

Кластер 8 (14 районів, 2 області) об'єднує Закарпатську область (13 з 13 районів) і 1 район Рівненської області. Повна ізоляція Закарпаття в окремий кластер підтверджує його унікальність через угорську історію та етнічну специфіку.

Кластер 9 (67 районів, 6 областей) об'єднує переважно південно-східні степові території: Дніпропетровську (19 з 22 районів), Полтавську (1 з 25), Херсонську (17 з 18), Кіровоградську (10 з 21), Запорізьку (19 з 20) та Миколаївську (1 з 19) області. Домінування Дніпропетровської (86 %), Херсонської (94 %), Запорізької (95 %) областей підтверджує формування виразного південно-східного степового макрокластера. Часткова присутність Кіровоградської області відображає її перехідне положення між центральним і степовим ареалами, тоді як поодинокі райони Полтавської та Миколаївської областей репрезентують периферійні включення до цього просторового блоку.

Кластер 10 (38 районів, 4 області) охоплює переважно південні причорноморські території: Одеську (19 з 26 районів), Миколаївську (17 з 19), Херсонську (1 з 18), Кіровоградську (1 з 21) області. Домінування Миколаївської області (89 %) та значна частка Одеської (73 % районів, за винятком специфічного південно-західного анклаву кластера 1) підтверджують формування окремого причорноморського ареалу. Така структура відображає спільну історію колонізації Північного Причорномор'я, портово-торговельну орієнтацію регіону та інтенсивні міграційні процеси XIX–XX століть.

Кластер 11 (15 районів, 1 область) включає виключно всі 15 районів Автономної Республіки Крим, формуючи повністю ізольований монообласний кластер. Така конфігурація засвідчує максимальну генетичну та антропонімічну відокремленість Кримського півострова від решти території України. Ізоляція пояснюється поєднанням географічного чинника (півострівне положення), специфічної етнічної структури (кримськотатарський, російський, український компоненти) та складної історичної траєкторії регіону.

3.4.2. Кластерна структура за дистанцією Ласкера

Кластеризація за дистанцією Ласкера (Рис. 3.19, 3.20) виявила 10 кластерів, що свідчить про дещо спрощену структуру просторової диференціації порівняно з результатами, отриманими за евклідовою дистанцією. Це пов'язано з тим, що дистанція Ласкера враховує ефективні розміри популяцій, що призводить до згладжування дрібних локальних відмінностей.

Кластер 1 (15 районів) повністю відповідає території Автономної Республіки Крим і є ідентичним кластеру, отриманому при використанні евклідової дистанції. Це підтверджує високу генетичну ізольованість кримських популяцій від решти території України.

Кластер 2 (62 райони, 3 області) об'єднує райони Донецької (18), Луганської (18), Харківської (26) областей. Порівняно з евклідовою кластерною структурою (кластер № 2), у кластері за дистанцією Ласкера відсутній один район Сумської області. Таким чином, кластер набуває більш чітко вираженого донбасько-слобожанського характеру, що відображає спільну історію індустріального розвитку цих регіонів.

Кластер 3 (29 районів, 4 області) включає 23 райони Одеської області (з 26), 4 райони Миколаївської області, 1 район Херсонської області та 1 район Кіровоградської області. Цей кластер є новим структурним елементом, який не має прямого аналога в евклідовій кластеризації. Концентрація 23 районів Одеської області (88% області) в одному кластері свідчить

про те, що використання дистанції Ласкера призводить до формування більш компактної та внутрішньо узгодженої структури Одещини.

Кластер 4 (63 райони, 5 областей) формує волинсько-поліський кластер і включає райони Волинської (16 з 16), Житомирської (20 з 23), Рівненської (16 з 16), Тернопільської (3) та Хмельницької (8) областей. Порівняно з відповідним кластером у структурі за евклідовою дистанцією (кластер 6), спостерігається розширення поліської компоненти: Рівненська область представлена повністю повністю (16 районів замість 13), а Житомирська область збільшує свою присутність з 11 до 20 районів. Таким чином, застосування дистанції Ласкера підсилює виразність волинсько-поліського макрокластера, інтегруючи більшу частину поліських територій у спільну кластерну структуру.

Кластер 5 (57 районів, 8 областей) характеризується як подільсько-центральною і включає райони Київської (4), Черкаської (6), Житомирської (1), Одеської (3), Кіровоградської (3), Миколаївської (1), Вінницької (27 з 27) та Хмельницької (12) областей. Порівняно з евклідовим кластером 5, його структура залишається загалом подібною: Вінницька область повністю входить до складу кластера (27 районів), а Хмельницька представлена 12 районами (порівняно з 11 у попередній метриці). Водночас частка інших областей дещо зменшується. Таким чином, Вінницька область зберігає роль ядра подільського кластера незалежно від використаної метрики генетичної дистанції.

Кластер 6 (39 районів, 3 області) формує галицьке ядро і включає райони Тернопільської (14), Львівської (20 з 20) та Івано-Франківської (5 з 14) областей. Порівняно з евклідовим кластером (кластер 7) спостерігається істотне скорочення представництва Івано-Франківської області: з 14 районів до 5, тоді як усі 8 районів Чернівецької області взагалі не входять до цього кластера. Таким чином, використання дистанції Ласкера розділяє раніше єдиний галицько-буковинський кластер, отриманий за евклідовою метрикою, на два окремі структурні блоки: галицьке ядро (кластер 6) та буковинсько-

гірський кластер (кластер 7). Це свідчить про те, що зважування на ефективні розміри популяцій дозволяє виявити глибшу популяційно-генетичну диференціацію між рівнинною Галичиною та гірськими територіями Буковини й Карпатського регіону Івано-Франківщини.

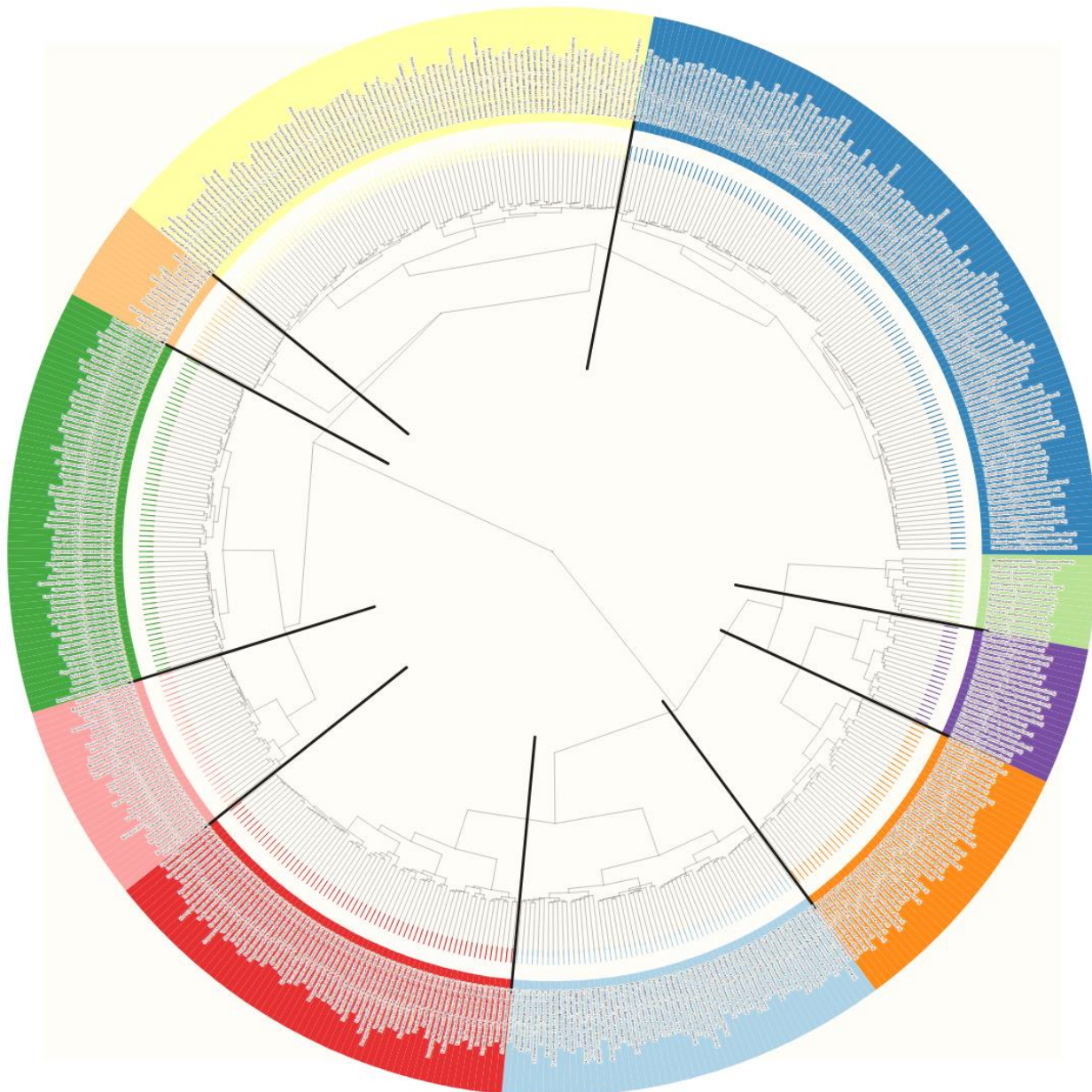


Рис. 3.19. Дендрограма ієрархічної кластеризації 490 районів України, побудована методом Уорда за дистанцією Ласкера. Інтерактивна версія доступна за [ПОСИЛАННЯМ](#).

Кластер 7 (20 районів, 2 області) формує буковинсько-гірський кластер і включає всі 11 районів Чернівецької області та 9 з 14 районів Івано-Франківської області. Цей кластер є новим елементом структури, який не має прямого аналога в евклідовій кластеризації. Він об'єднує території,

що характеризуються переважно гірським рельєфом та історичною ізолюваністю населення. Поділ Івано-Франківської області між кластерами 6 (5 районів) і 7 (9 районів) відображає внутрішньообласну диференціацію між рівнинними та гірськими районами, ймовірно пов'язану з відмінностями у демографічній історії та міграційних процесах.

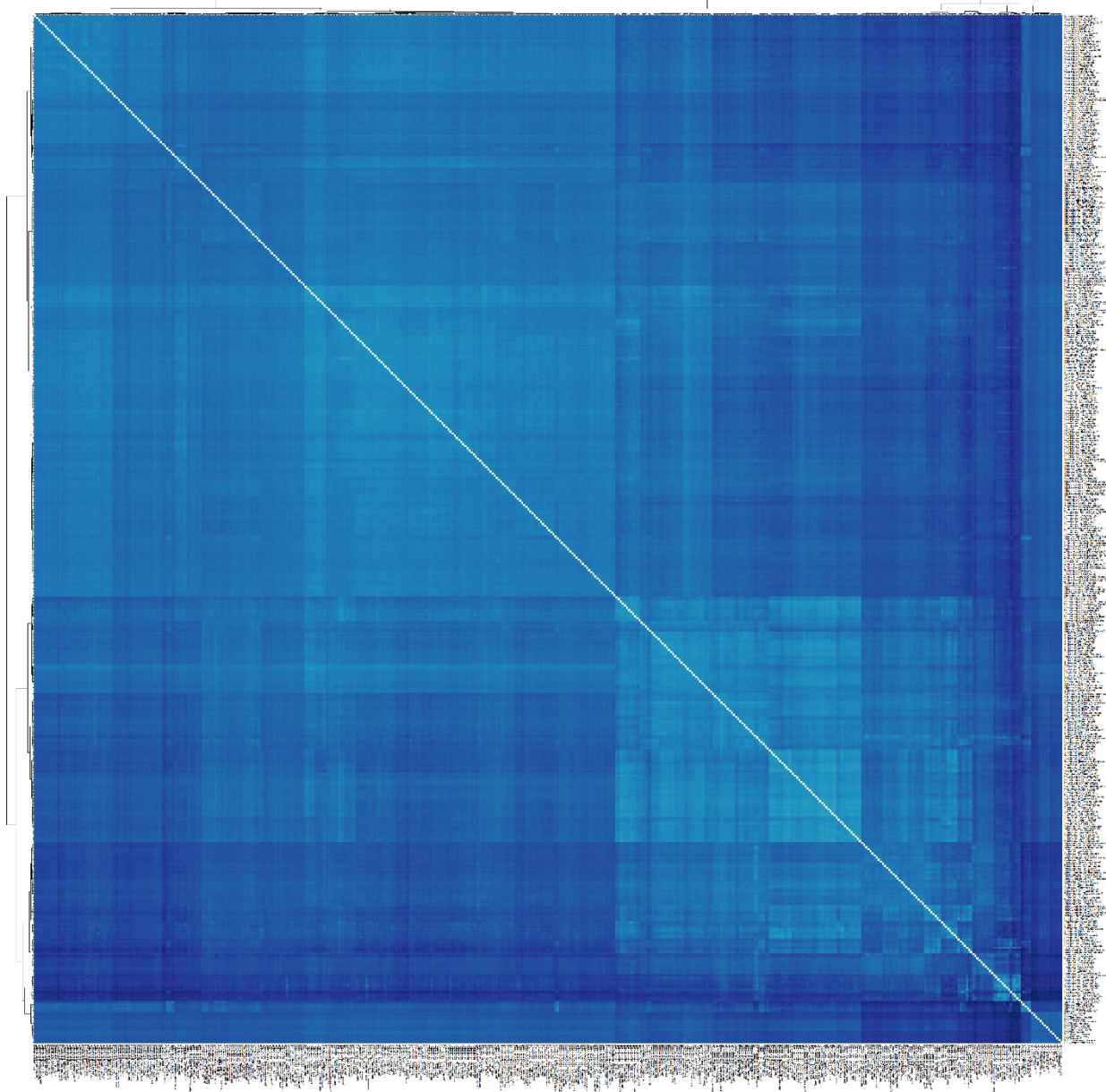


Рис. 3.20. Теплова карта матриці дистанцій Ласкера між 490 районами України з відображенням кластерної структури (світлі зони відповідають низьким значенням показника, темні — високим).

Кластер 8 (13 районів, 1 область) представлений усіма 13 районами Закарпатської області. Його структура практично ідентична результатам кластеризації за евклідовою дистанцією (кластер 8), однак у цьому випадку

до складу не входить жоден район інших областей, що робить кластер повністю монообласним. Таким чином, дистанція Ласкера ще чіткіше підкреслює генетичну та демографічну ізольованість Закарпатського регіону, сформовану його географічною відокремленістю Карпатським гірським масивом та специфічною історією розвитку.

Кластер 9 (108 районів, 7 областей) формує південно-східний кластер і є найбільшим у структурі дистанції Ласкера, охоплюючи близько 22% усіх районів України. До його складу входять усі 22 райони Дніпропетровської області, а також 19 районів Запорізької, 17 — Херсонської, 14 — Миколаївської, 12 — Черкаської, 7 — Київської та 17 — Кіровоградської областей.

Порівняно з кластеризацією за евклідовою дистанцією, цей кластер об'єднує значні частини евклідових кластерів 3, 9 і 10, формуючи великий інтегрований регіональний блок. Така консолідація свідчить про те, що метрика Ласкера, яка враховує ефективні розміри популяцій, згладжує локальні відмінності та підкреслює спільність демографічного формування південно-східних і частини центральних територій України. Історично ця спільність може бути пов'язана з процесами інтенсивної колонізації та господарського освоєння степових територій у XVIII–XX століттях, що супроводжувалися масштабними міграційними потоками та змішуванням населення [46, 82, 93].

Кластер 10 (84 райони, 7 областей) формує лівобережно-центральний кластер, до складу якого входять усі 25 районів Полтавської та всі 18 районів Сумської областей, а також 22 райони Чернігівської області, 14 районів Київської, 2 райони Черкаської, 2 райони Житомирської та 1 район Запорізької області.

Порівняно з відповідним кластером 3 у структурі евклідових дистанцій, цей кластер відображає подібний лівобережно-поліський просторовий блок, однак із більш чіткою територіальною концентрацією. Повна присутність Полтавської та Сумської областей свідчить про їхню високу внутрішню

однорідність та спільність демографічної історії. У результаті застосування дистанції Ласкера лівобережна ідентичність цих регіонів проявляється значно виразніше, ніж у евклідовій моделі, що підкреслює їхню історичну та популяційно-генетичну спорідненість у межах Лівобережної України.

3.4.3. Кластерна структура за дистанцією Нея

Кластеризація за дистанцією Нея повертає структуру з 11 кластерами, аналогічно до результатів, отриманих за евклідовою метрикою, проте з іншою внутрішньою організацією, що відображає більш коректні з філогенетичної точки зору еволюційні взаємовідносини між популяціями (Рис. 3.21, 3.22).

Кластер 1 (15 районів, 1 область) включає всі райони Автономної Республіки Крим. Його склад є повністю ідентичним результатам кластеризації як за евклідовою дистанцією, так і за дистанцією Ласкера. Таким чином, усі три використані метрики демонструють консенсусне виокремлення Криму як абсолютно ізольованого кластера, що свідчить про його специфічну антропонімічну та популяційно-генетичну структуру.

Кластер 2 (70 районів, 4 області) формує розширений східний кластер, до якого входять усі райони Луганської (18 з 18), Харківської (26 з 26) та Донецької (18 з 18) областей, а також 8 районів Одеської області (з 26).

Порівняно з кластеризацією за евклідовою дистанцією, до донбасько-слобожанського ядра було додатково включено частину районів Одеської області. Така конфігурація може відображати спільні історичні демографічні процеси південного сходу України, зокрема міграційні потоки періоду індустріалізації XIX–XX століть та формування багатонаціонального населення степових регіонів. У межах метрики Нея ці території демонструють вищу філогенетичну спорідненість, що свідчить про подібність антропонімічної структури та історії формування населення.

Вісім районів Одеської області формують компактний південно-західний блок регіону: Ренійський, Болградський, Арцизький, Тарутинський і Саратський (ті самі 5 районів, що утворювали ізольований кластер 1 за евклідовою дистанцією), а також Татарбунарський, Ізмаїльський і Кілійський. Усі вони розташовані в межах історико-географічного регіону Придунав'я та Буджаку і характеризуються спільними демографічними та етнокультурними особливостями. Ця територія відзначається значною

часткою болгарського та молдавського населення, сформованою внаслідок колонізаційних процесів початку XIX століття.

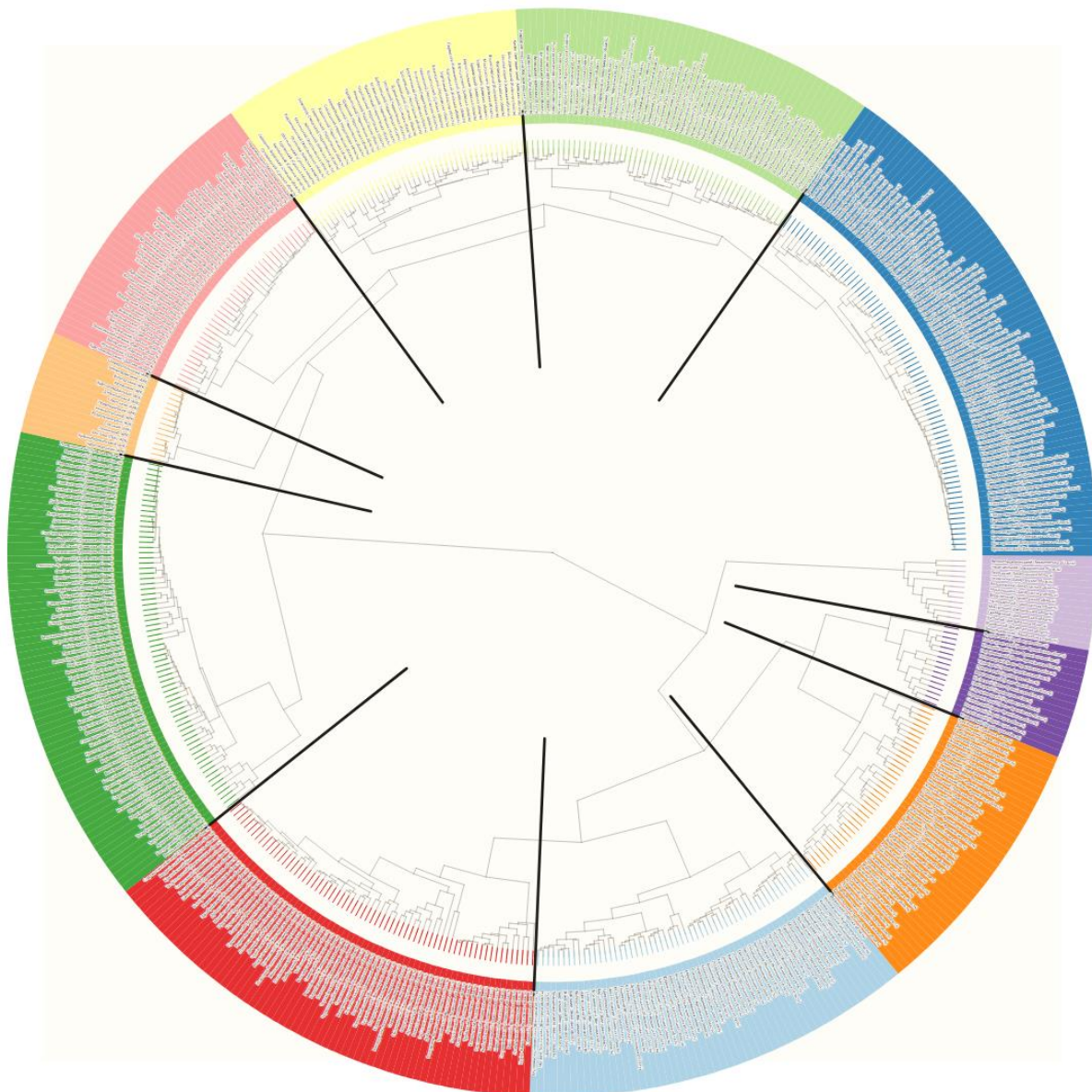


Рис. 3.21. Дендрограма ієрархічної кластеризації 490 районів України, отримана методом Уорда на основі генетичних дистанцій Нея. Інтерактивна версія доступна за [посиланням](#).

Дистанція Нея об'єднує ці райони зі східноукраїнським кластером, що відображає спільність південно-східних територій як колонізаційного фронтиру XVIII–XIX століть, незважаючи на значну географічну відстань між Одещиною та Донбасом. Обидва регіони інтенсивно заселялися в один історичний період після Російсько-турецьких воєн: Придунав'я та Буджак — після приєднання Бессарабії у 1812 році, а Донбас — після ліквідації Кримського ханства у 1783 році [46, 82, 93]. Таким чином, філогенетично

коректна дистанція Нея виявляє спільність демографічних процесів пізньої колонізації, яка в аналізі за евклідовою дистанцією частково маскується географічною відокремленістю та локальною етнічною специфікою цих регіонів.

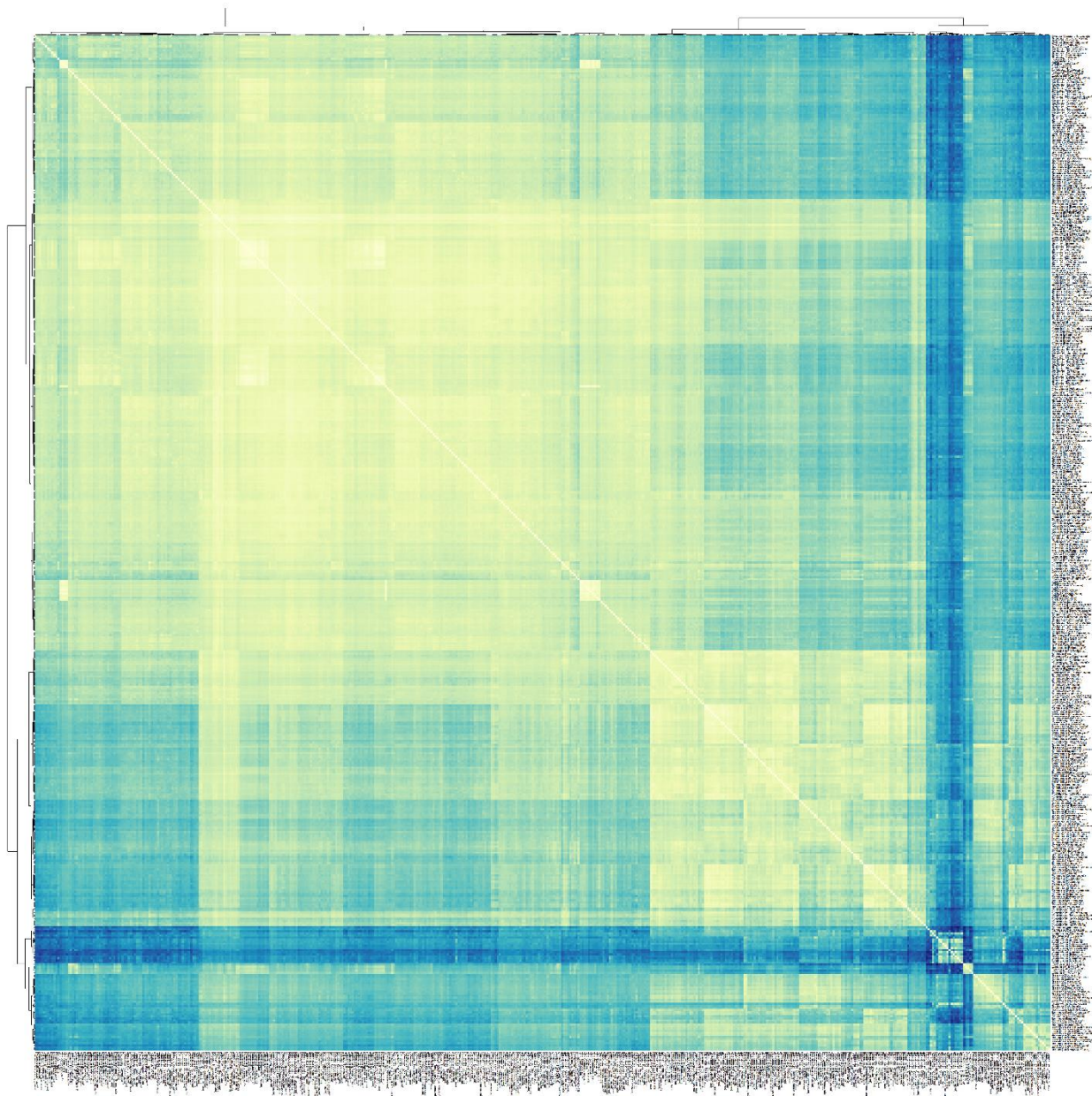


Рис. 3.22. Теплова карта матриці генетичних дистанцій Нея між 490 районами України з відображенням кластерної структури (світлі зони відповідають низьким значенням показника, темні — високим).

Кластер 3 (67 районів, 6 областей) формує волинсько-поліський розширений кластер, до складу якого увійшли районий Волинської області (16 з 16), Рівненської (16 з 16), Житомирської (21 з 23), Хмельницької (8 з 20), Київської (3 з 25), Тернопільської (3 з 17) областей. Порівняно з кластером

за дистанцією Ласкера (кластер 4 — 63 райони) додатково увійшов один район Житомирської області (21 замість 20). Це підтверджує волинсько-поліську монолітність Волині та Рівненщини (обидві повністю) і майже повну належність Житомирщини (91%) до цього ареалу.

Кластер 4 (57 районів, 8 областей) — подільсько-центральный: Вінницька (27 з 27), Одеська (3 з 26), Хмельницька (12 з 20), Чернівецька (4 з 11), Черкаська (6 з 20), Миколаївська (1 з 19), Кіровоградська (3 з 21), Київська (1 з 25). Структура майже ідентична евклідовому кластеру 5 та ласкерівському кластеру 5. Вінницька область консенсусно є ядром подільського кластера у всіх трьох метриках.

Кластер 5 (39 районів, 3 області) — галицьке ядро: Тернопільська (14 з 17), Львівська (20 з 20), Івано-Франківська (5 з 14). Структура цього кластера повністю ідентична ласкерівському кластеру 6. Це підтверджує розділення території Івано-Франківської області між галицьким рівнинним ядром (5 районів) і гірським блоком Карпатського регіону, який формує окремий кластер. Таким чином, кластер відображає компактний просторово-генетичний масив історичної Галичини з центром у Львівській області та периферійним включенням частини Івано-Франківщини.

Кластер 6 (16 районів, 2 області) — буковинсько-гірський: Чернівецька (7 з 11), Івано-Франківська (9 з 14). Його структура майже ідентична ласкерівському кластеру 7, однак відрізняється меншою присутністю Чернівецької області (7 районів замість 11). Решта 4 райони Чернівецької області входять до складу кластера 4 (подільського), що свідчить про просторове тяжіння частини території Буковини до подільського рівнинного масиву.

Основу цього кластера формують гірські та передгірські райони Івано-Франківської області (9 районів), які разом із південно-західною частиною Чернівецької області утворюють єдиний карпатський популяційно-генетичний блок. Це виявляє внутрішню гетерогенність Буковини, де простежується чіткий поділ між гірською частиною (7 районів у кластері 6)

та рівнинною подільською зоною (4 райони у кластері 4). У ширшому інтерпретаційному контексті така структура кластерів відображає диференціацію між гірськими карпатськими популяціями та сусідніми рівнинними територіями Поділля.

Кластер 7 (13 районів, 1 область) формує окремий закарпатський кластер, до складу якого входять усі 13 районів Закарпатської області. Його структура повністю відповідає ласкеровому кластеру 8 і характеризується повною територіальною замкненістю в межах однієї області без залучення районів інших регіонів. Усі три використані метрики генетичної дистанції підтверджують повну відокремленість Закарпаття, що свідчить про стабільну популяційно-генетичну специфіку цього регіону та його чітку просторову диференціацію від інших частин України.

Кластер 8 (75 районів, 7 областей) — південно-східний степовий: Дніпропетровська (21 з 22), Полтавська (1 з 25), Кіровоградська (5 з 21), Херсонська (18 з 18), Запорізька (20 з 20), Миколаївська (2 з 19), Сумська (8 з 18). Порівняно з евклідовою метрикою це відповідає евклідовому кластеру 9, але з додаванням 8 районів Сумської області. Повне включення Запорізької області (20 з 20) і майже повне включення Херсонської області (18 з 18) підтверджує їхню спільну південну степову ідентичність.

Вісім районів Сумської області представляють центральну та південну частини регіону: Сумський, Лебединський, Охтирський, Білопільський, Тростянецький, Краснопільський, Недригайлівський, Великописарівський. Ця група охоплює переважно центральні та південно-східні райони Сумщини, які географічно й історично тяжіють до лівобережного степового ареалу. На відміну від північних районів Сумської області (які належать до поліського кластера), ці райони мають більшу спорідненість із Полтавщиною та південно-східним степовим блоком через спільну історію заселення території Лівобережної Гетьманщини та подальшу інтеграцію до лівобережно-степового демографічного ареалу [46, 82, 93].

Кластер 9 (53 райони, 7 областей) формує центрально-лівобережний кластер, до складу якого входять 24 райони Полтавської області, 14 районів Черкаської, 6 районів Київської, 3 райони Чернігівської, 3 райони Кіровоградської, 2 райони Сумської та 1 район Дніпропетровської області. Це новий кластер у метриці Нея, який не має прямого аналога в інших метриках. Його структура об'єднує переважно центральні райони з домінуванням Полтавської області (24 з 25 — 96 %) і значною часткою Черкаської області (14 з 20 — 70 %). Розподіл Полтавської області між кластерами 8 (1 район) і 9 (24 райони) свідчить про наявність одного південного степового аутлаера, який тяжіє до південно-східного степового просторового блоку.

Кластер 10 (44 райони, 4 області) формує північний поліський кластер, до складу якого входять 19 районів Чернігівської області, 15 районів Київської, 8 районів Сумської та 2 райони Житомирської області. Порівняно з відповідними структурами в інших метриках, цей кластер відображає компактний північний поліський просторовий блок із домінуванням Чернігівської області (19 з 22 — 86 %) та значною часткою Київської області (15 з 25 — 60 %). Розподіл Київської області між кількома кластерами (кластер 3 — 3 райони, кластер 4 — 1 район, кластер 9 — 6 районів, кластер 10 — 15 районів) свідчить про виражену внутрішньообласну просторову гетерогенність столичного регіону. Водночас присутність 8 районів Сумської області (ще 2 входять до кластера 9, ще 8 — до кластера 8) вказує на її поділ між північним, центральним та східним просторовими ареалами.

Кластер 11 (41 район, 3 області) формує південний причорноморський кластер, до складу якого входять 16 районів Миколаївської області, 15 районів Одеської та 10 районів Кіровоградської області. Порівняно з відповідним кластером 10 у структурі евклідових дистанцій, цей кластер загалом зберігає подібну просторову конфігурацію, однак відзначається значно більшою присутністю Кіровоградської області (10 районів замість 1). Такий розподіл свідчить, що в моделі за дистанцією Нея Кіровоградська область розділяється

між кількома регіональними блоками: південним причорноморським (10 районів), подільським (3 райони, кластер 4), південно-східним степовим (5 районів, кластер 8) та центрально-лівобережним (3 райони, кластер 9). Це підтверджує транзитну позицію Кіровоградської області між кількома великими просторовими макрорегіонами України.

3.4.4. Порівняльний аналіз та інтерпретація кластерної структури районного рівня

3.4.4.1. Топологічна стабільність кластерної архітектури

Застосування методу Уорда до матриць генетичних дистанцій для 490 районів України з використанням трьох альтернативних метрик (евклідова дистанція, дистанція Ласкера, дистанція Нея) виявило високу, але не абсолютну топологічну стабільність кластерної структури. Оптимальна кількість кластерів становить 11 для евклідової дистанції та дистанції Нея і 10 для дистанції Ласкера, що вказує на незначну варіацію кластерної архітектури залежно від обраної метрики (Рис. 3.23 – 3.25).

Порівняльний аналіз трьох кластерних структур дав змогу ідентифікувати консенсусні елементи, які зберігаються незалежно від метрики дистанції, а також метрикоспецифічні ефекти, що відображають різні аспекти популяційно-генетичної структури.

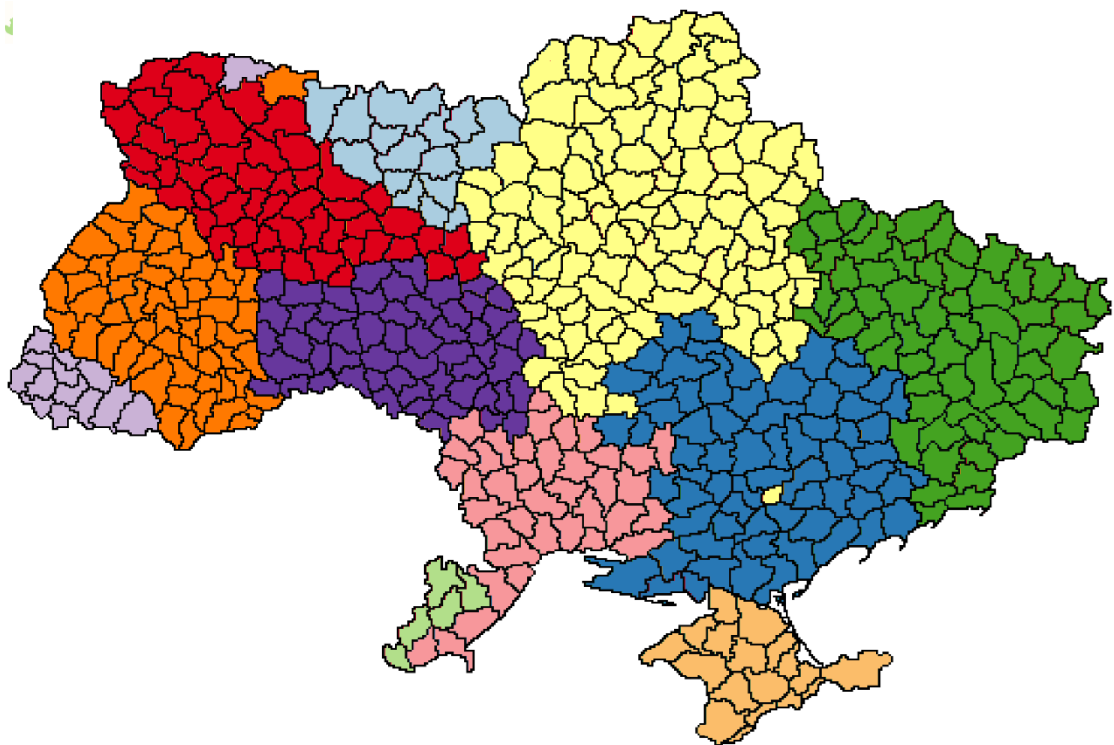


Рис. 3.23. Просторовий розподіл кластерів популяційно-генетичної структури України на районному рівні за евклідовою дистанцією (метод Урда, 11 кластерів). Кольором позначено належність районів до відповідних кластерів. Інтерактивна версія доступна за [ПОСИЛАННЯМ](#).

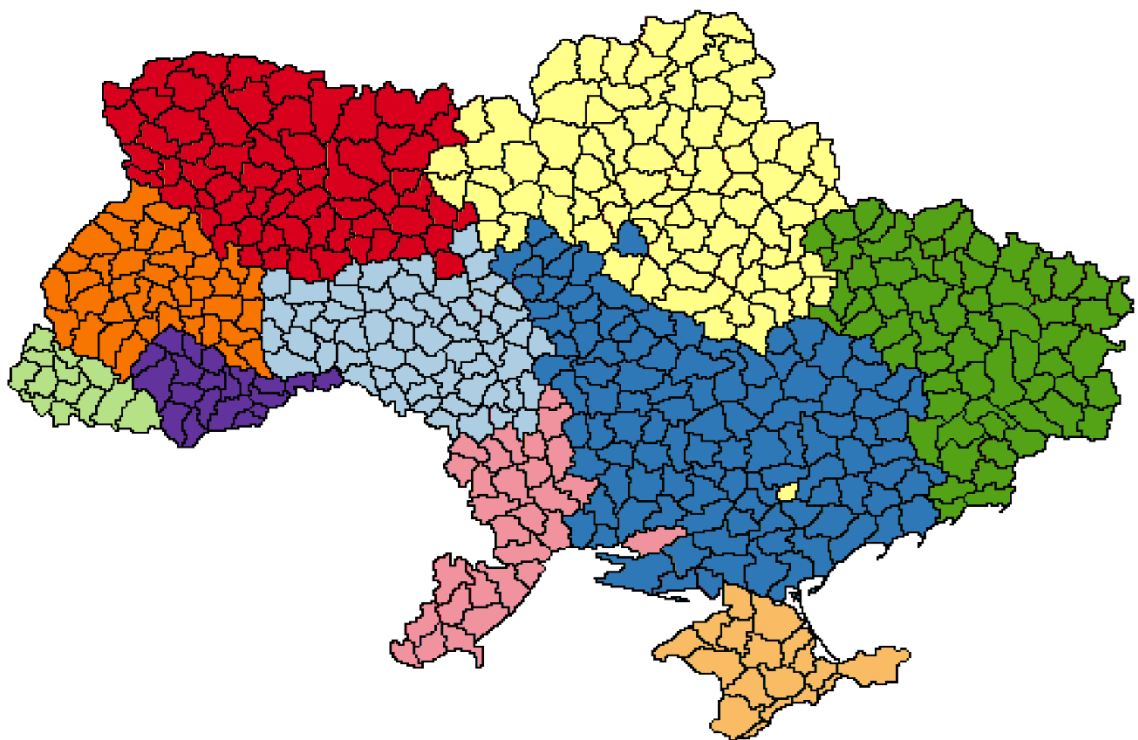


Рис. 3.24. Просторовий розподіл кластерів популяційно-генетичної структури України на районному рівні за дистанцією Ласкера (метод Урда, 10 кластерів). Кольором позначено належність районів до відповідних кластерів. Інтерактивна версія доступна за [ПОСИЛАННЯМ](#).

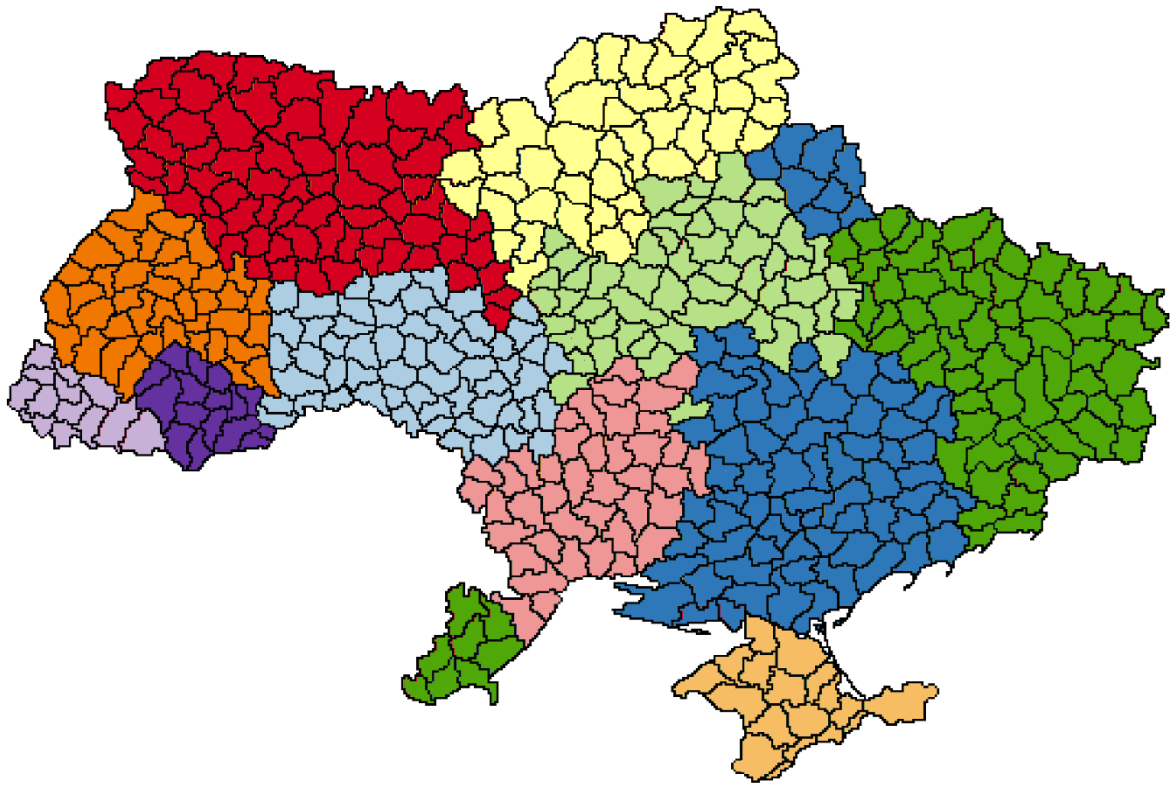


Рис. 3.25. Просторовий розподіл кластерів популяційно-генетичної структури України на районному рівні за дистанцією Нея (метод Уорда, 11 кластерів). Кольором позначено належність районів до відповідних кластерів. Інтерактивна версія доступна за [посиланням](#).

3.4.4.2. Консенсусні філогенетичні одиниці

Аналіз виявив шість абсолютно стабільних популяційно-генетичних одиниць: АР Крим, Закарпатське, Волинське, Вінницьке, Львівське області, а також Східну Україну, які зберігають свою цілісність у всіх трьох метриках.

АР Крим формує абсолютно ізольований монообласний кластер у всіх трьох метриках (кластер 11 евклідової дистанції, кластер 1 дистанції Ласкера та Нея). Усі 15 районів Криму консенсусно відокремлюються від решти України, що є прямим відображенням його унікальної демографічної історії: кримськотатарського коріння, відносно пізнього російського та українського заселення, а також географічної ізоляції півострова [46, 82, 93]. Це єдиний регіон України, який формує повністю монолітний кластер без жодних міжобласних зв'язків, що вказує на максимальну генетичну ізоляцію від континентальної частини країни.

Закарпатська область (13 районів) утворює повністю ізольований монообласний кластер у всіх трьох метриках (кластер 8 евклідової дистанції та дистанції Ласкера, кластер 7 дистанції Нея). Генетична унікальність Закарпаття пояснюється його тривалим історичним перебуванням у складі Угорщини до 1919/1945 років, специфічним етнічним складом населення (угорці, русини, румуни) та географічним бар'єром Карпатського хребта [46, 82, 93, 110]. Ізоляція Закарпаття є виразнішою порівняно із західноукраїнською (галицькою) специфіку, формуючи окремий філогенетичний острів у межах популяційно-генетичної структури України.

Волинська область (16 районів) повністю входить до волинсько-поліського кластера у всіх трьох метриках. Монолітність усіх 16 волинських районів відображає відносно однорідну демографічну історію без значних міграційних хвиль та масштабної індустріалізації. Волинь зберегла автохтонний український субстрат краще, ніж більшість інших регіонів країни, що значною мірою зумовлено переважно аграрним характером її економіки та пізнішим включенням до складу СРСР.

Вінницька область (27 районів) повністю входить до складу подільського кластера в усіх трьох метриках, виступаючи його консенсусним ядром (кластер 5 евклідової дистанції та дистанції Ласкера, кластер 4 дистанції Нея). Відсутність будь-якого розподілу вінницьких районів між різними кластерами свідчить про стабільну демографічну історію подільського регіону з мінімальними зовнішніми впливами після формування його населення у XVI–XVIII століттях [46, 82, 93].

Львівська область (20 районів) також демонструє повну кластену цілісність у всіх трьох (кластер 7 евклідової дистанції, кластер 6 дистанції Ласкера та Нея). Монолітна структура галицького кластера, до якого входять усі 20 районів області, відображає стабільність західноукраїнської регіональної ідентичності. Її формування пов'язане з тривалим історичним розвитком у складі Австро-Угорщини та збереженням специфічних

демографічних характеристик попри радянську уніфікацію другої половини ХХ століття [46, 82, 93].

Донбас (Донецька та Луганська області) утворює стабільну двообласну популяційно-генетичну одиницю, що охоплює 36 районів (по 18 у кожній області) і зберігається в усіх трьох метриках (кластер 2 евклідової дистанції та дистанції Ласкера, частина кластера 2 дистанції Нея). Повна присутність усіх районів обох областей свідчить про надзвичайно високий рівень внутрішньої спорідненості населення Донбасу. Така кластерна цілісність відображає спільну історію формування регіону як великого індустріального центру, що виник у процесі масштабної індустріалізації ХІХ–ХХ століть і супроводжувався інтенсивними міграційними потоками з різних частин Російської імперії та СРСР [46, 82, 93].

Консенсусна ізоляція цих шести територіальних одиниць у цих шести територіальних одиниць у межах усіх трьох метрик є вагомим підтвердженням їхньої популяційно-генетичної стабільності. Стійкість їхнього кластерного положення незалежно від обраної метрики дистанції свідчить про те, що ці регіони не є артефактами статистичної процедури кластеризації, а відображають реальні філогенетичні структури з відносно автономними демографічними історіями та чітко вираженими просторовими межами.

3.4.4.3. Метрикоспецифічні ефекти та їхнє біологічне значення

Попри високу топологічну стабільність консенсусних популяційно-генетичних одиниць, використання трьох різних метрик дистанції дозволяє виявити додаткові аспекти просторової структури населення України. Відмінності між кластерними рішеннями відображають специфічні математичні властивості кожної метрики та пов'язані з ними біологічні припущення щодо природи генетичних дистанцій.

Евклідова дистанція формує найбільш збалансовану кластерну архітектуру з 11 кластерами, розмір яких варіює від 5 до 107 районів. Ця метрика відображає сучасний генетичний стан популяцій без урахування

демографічних параметрів, що робить її особливо придатною для виявлення внутрішньорегіональної гетерогенності. Показовим прикладом є Одеська область, яка в евклідовій моделі розділяється на три структурні компоненти: мінімальний кластер 1 (5 районів), частину подільського кластера 5 (2 райони) та найбільшу частку причорноморського кластера 10 (19 районів). Таке трикомпонентне розщеплення відображає контраст між невеликою групою сильно відокремлених районів, ймовірно пов'язаних із прибережними портовими територіями або південно-західними районами з болгарським і румунсько-молдавським етнічним впливом, невеликою подільською компонентою та основним причорноморським демографічним масивом.

Евклідова дистанція також демонструє найвищий рівень внутрішньообласної диференціації у столичному регіоні. Київська область у цій моделі розподіляється між кількома кластерами: 19 районів входять до кластера 3, 2 — до кластера 4, 3 — до кластера 5 та 1 район — до кластера 6. Така структура відображає складну демографічну історію регіону, де поєднуються поліські, центрально-українські та лівобережні популяційні компоненти, що сформувалися внаслідок різних історичних процесів заселення та міграцій.

Дистанція Ласкера, яка враховує ефективні розміри популяцій, спрощує кластерну структуру до 10 кластерів завдяки злиттю деяких одиниць, виділених за евклідовою метрикою. Найбільш помітним ефектом застосування дистанції Ласкера є формування мегакластера 9 (108 районів — 22 % усіх районів України), який об'єднує південно-східні та частину центральних територій. До його складу входить Дніпропетровська область повністю (22 райони), майже повністю Запорізька (19 з 20) і Херсонська (17 з 18) області, значна частина Миколаївської області (14 з 19), а також суттєві частки Черкаської (12 з 20), Київської (7 з 25) та Кіровоградської (17 з 21) областей.

Такий результат має важливе біологічне та історико-демографічне пояснення. Території пізньої колонізації XVIII–XX століть справді характеризуються спільною демографічною історією, пов'язаною з масовим

переселенням населення з Центральної України та Правобережжя після ліквідації Кримського ханства у 1783 році та укладення Кючук-Кайнарджійського мирного договору. У цьому контексті дистанція Ласкера виявляє їхню генетичну спорідненість як територій колонізаційного фронтиру, попри їх сучасне адміністративне розділення між різними областями [46, 82, 93].

Другим важливим ефектом дистанції Ласкера є розділення єдиного галицько-буковинського блоку, виявленого за евклідовою дистанцією (кластер 7: Львівська — 20, Івано-Франківська — 14, Тернопільська — 14, Чернівецька — 8, Рівненська — 1), на два окремі кластери: галицьке ядро (кластер 6: Львівська — 20, Тернопільська — 14, Івано-Франківська — 5) та буковинсько-гірський блок (кластер 7: Чернівецька — 11, Івано-Франківська — 9). Це розділення відображає екологічні бар'єри між рівнинними та гірськими популяціями: зважування на ефективні розміри популяцій дозволяє виявити глибшу диференціацію між густонаселеною рівнинною Галичиною та розрідженими гірськими районами Буковини і Карпат.

Третім ефектом дистанції Ласкера є компактизація структури Одеської області: 23 з 26 районів (88 %) об'єднуються в один кластер (кластер 3), тоді як евклідова дистанція розділяє їх на три компоненти. Таким чином, дистанція Ласкера «стискає» просторово розпорошену структуру Одещини в більш компактну кластерну одиницю, нівелюючи її внутрішню гетерогенність завдяки врахуванню популяційних розмірів.

Дистанція Нея, як філогенетично найбільш коректна метрика, виявляє найтонші еволюційні взаємозв'язки, які не завжди очевидні з географічної або адміністративної перспективи. Найбільш контрінтуїтивним, але біологічно обґрунтованим ефектом застосування дистанції Нея є об'єднання 8 районів Одеської області (із 26 загалом) зі східним донбасько-слобожанським кластером (кластер 2: Одеська — 8, Луганська — 18, Харківська — 26, Донецька — 18; загалом 70 районів).

Це об'єднання, попри значну географічну відстань між Одещиною та Донбасом, відображає спільність південно-східного колонізаційного фронтиру XVIII–XIX століть: як степова північна Одещина, так і Донбас заселялися приблизно в один історичний період вихідцями з Центральної України та Росії, що сформувало подібний демографічний субстрат [46, 82, 93]. Дистанція Нея, коригуючи оцінки генетичної дистанції з урахуванням еволюційних процесів, виявляє цю глибинну філогенетичну спорідненість, яка в інших метриках частково маскується географічною відокремленістю.

Другим важливим ефектом застосування дистанції Нея є виявлення триплетної структури Сумської області, яка розподіляється між трьома кластерами: північним поліським (8 районів у кластері 10), центральним лівобережним (2 райони у кластері 9) та південним степовим (8 районів у кластері 8). Таке триплетне розщеплення підтверджує транзитний характер Сумщини на перетині Полісся, Лівобережжя та степового Південного Сходу — географічну позицію, яка не проявляється настільки виразно при використанні інших метрик.

3.4.4.4. Области з екстремальною внутрішньообласною гетерогенністю

Рівень внутрішньообласної гетерогенності, визначений як кількість кластерів, між якими розподіляється область, є важливим індикатором складності демографічної історії регіону. Порівняльний аналіз трьох метрик дозволив ідентифікувати чотири області з максимальною внутрішньою гетерогенністю, які стабільно розщеплюються між кількома кластерами.

Одеська область є найгетерогеннішим регіоном України, демонструючи розподіл між трьома або чотирма кластерами залежно від використаної метрики. Евклідова дистанція розділяє 26 районів області на три компоненти (5 + 2 + 19 районів), дистанція Ласкера компактизує більшість у кластер 3 (23 райони), але залишає три райони в інших кластерах, тоді як дистанція Нея знову розподіляє область між трьома блоками (8 районів у східному кластері + 15 районів у причорноморському + 3 райони у подільському).

Така екстремальна гетерогенність відображає унікальну багатосторонню історію Одещини: грецьку та болгарську колонізацію південно-західної частини регіону у XVIII–XIX століттях, молдавсько-румунський вплив Задунав'я, космополітичний характер портового міста Одеси з інтенсивними міграційними процесами, а також степові райони центральної та північної частин області, заселені внаслідок російсько-української колонізації. Жодна інша область України не демонструє такої різноманітності демографічних субстратів у межах однієї адміністративної одиниці.

Київська область демонструє розділення між 4 і 5 кластерами за всіма трьома метриками, що відображає контраст між столичною агломерацією з інтенсивною міграцією XX–XXI століть, приміськими районами, центральними правобережними районами та поліськими північними районами. Евклідова дистанція розподіляє 25 районів Київської області між п'ятьма кластерами, дистанція Ласкера — між чотирма, а дистанція Нея — знову між п'ятьма.

Столичний статус сформував унікальну популяційну мозаїку: райони, прилеглі до Києва, зазнали масової урбанізації та міграції, тоді як периферійні райони зберегли більш традиційну демографічну структуру. Північні поліські райони (ймовірно, Поліський та інші) мають більшу спорідненість із білоруським демографічним субстратом, ніж з центральноукраїнським, що додатково ускладнює структуру області.

Кіровоградська область розподіляється між чотирма кластерами (за дистанцією Нея: 10 районів причорноморських, 5 — південно-східного степового, 3 — подільського та 3 — центрального кластерів), що підтверджує її розташування на перетині чотирьох макрорегіонів України. Така структура вказує на транзитний характер регіону, який не має чітко вираженого власного популяційно-генетичного ядра. Північні райони області тяжіють до подільського ареалу (Вінницької області), західні — до центральноукраїнського масиву Черкащини, південні — до причорноморського регіону (Миколаївської області), тоді як східні —

до південно-східного степового блоку (Дніпропетровської області). Таким чином, Кіровоградська область виступає своєрідним демографічним «перехрестям» України, що й зумовлює її високий рівень внутрішньообласної гетерогенності.

Івано-Франківська область розподіляється між двома кластерами — рівнинним галицьким (5 районів) та гірським гуцульським (9 районів) — у моделях дистанцій Ласкера та Нея, тоді як евклідова дистанція об'єднує всі 14 районів області в єдиний галицько-буковинський блок. Розділення, виявлене за дистанціями Ласкера та Нея, відображає екологічний бар'єр між рівнинною Прикарпатською Галичиною та гірськими районами Карпат. Гуцульські райони (зокрема Верховинський, Косівський, Надвірнянський та інші високогірні території) демонструють більшу популяційно-генетичну спорідненість із гірською частиною Буковини (Чернівецькою областю), ніж із рівнинними територіями Львівської чи Тернопільської областей.

Гірська ізоляція сприяла збереженню низки архаїчних популяційно-генетичних характеристик гуцульського населення, які певною мірою відрізняють його від рівнинних галичан, попри спільну західноукраїнську культурно-історичну ідентичність. Урахування ефективних розмірів популяцій (у дистанції Ласкера) та еволюційних відстаней (у дистанції Нея) дозволяє виявити цю диференціацію, яка частково маскується при використанні евклідової дистанції.

3.4.4.5. Ізоляція за відстанню

Для перевірки кореляції між географічними та генетичними відстанями було застосовано тест Мантеля. Аналіз проводили між матрицею географічних відстаней (розрахованою за формулою гаверсинуса (Haversine) між центрами 490 адміністративних районів) та трьома матрицями генетичних дистанцій (евклідовою, Ласкера та Нея).

Тест Мантеля виявив статистично значущу позитивну кореляцію між географічною та генетичною відстанню, розрахованою за однією із трьох метрик:

- за евклідовою дистанцією: $r = 0,581$, $p < 0,001$;
- за дистанцією Ласкера: $r = 0,691$, $p < 0,001$;
- за дистанцією Нея: $r = 0,246$, $p < 0,001$.

Найвищий коефіцієнт кореляції спостерігається для дистанції Ласкера ($r = 0,691$), що є очікуваним, оскільки цей показник чутливий до спільних рідкісних прізвищ і краще відображає історичні міграційні зв'язки. Помірна кореляція за евклідовою дистанцією ($r = 0,581$) та слабша, але все ще значуща кореляція за дистанцією Нея ($r = 0,246$) свідчать про те, що географічна відстань залишається важливим фактором формування популяційної структури, хоча її вплив проявляється по-різному залежно від обраної метрики.

Отримані результати підтверджують дію моделі ізоляції за відстанню в українських популяціях навіть після масштабних демографічних потрясінь ХХ століття (Голодомор, Друга світова війна, масові міграції та індустріалізація). Це вказує на відносну стабільність історично сформованої просторової структури генофонду, яка не була повністю зруйнована процесами урбанізації та міграції радянського періоду. Найсильніша кореляція характерна для західних і центральних регіонів, тоді як у східних і південних областях вплив географічної відстані дещо послаблюється через інтенсивний приплив мігрантів у ХІХ–ХХ століттях.

3.4.4.6. Історико-демографічна інтерпретація макрорегіональної структури

Кластерний аналіз, проведений із використанням трьох метрик генетичної дистанції, дозволяє реконструювати чотири основні макрорегіональні блоки, що відображають різні траєкторії демографічної історії території України.

Західноукраїнський блок (Волинська, Рівненська, Львівська, Тернопільська, Івано-Франківська, Чернівецька, Закарпатська області) формує найбільш філогенетично диференційований регіональний комплекс із чітко вираженою внутрішньою структурою. Для цього блоку характерне збереження автохтонних українських популяційних субстратів, що значною мірою

пов'язано з пізнім входженням відповідних територій до складу СРСР (1939–1945 роки для Волині, Галичини та Буковини; 1945 рік для Закарпаття). Це суттєво обмежило вплив масштабних процесів радянської індустріалізації та внутрішньосоюзних міграцій, які значно трансформували демографічну структуру інших регіонів країни [46, 82, 93].

У межах західноукраїнського блоку виокремлюються чотири основні субодиноці. Перша — волинське ядро, яке охоплює всю Волинську область, більшість районів Рівненської області та частину Житомирської області і характеризується максимальною збереженістю автохтонного українського демографічного субстрату. Друга — галицьке ядро, до складу якого входять усі райони Львівської області, більшість районів Тернопільської області та рівнинна частина Івано-Франківської області; його формування пов'язане з тривалим історичним розвитком у межах Австро-Угорської монархії. Третю субодиноцю становить буковинсько-карпатський блок (Чернівецька область разом із гірською частиною Івано-Франківської області), де гірська ізоляція сприяла збереженню низки архаїчних популяційно-генетичних характеристик. Четверта субодиноця — закарпатський ізольований блок (13 районів), який відзначається угорсько-русинським демографічним субстратом і демонструє виразну генетичну відокремленість від решти території України.

Східний індустріальний блок (Донецька, Луганська, Харківська області, 62–70 районів залежно від метрики) формує один із найбільш гомогенних макрорегіонів України. Його формування пов'язане з масштабною індустріалізацією XIX–XX століть, яка супроводжувалася масовим припливом мігрантів із різних частин Російської імперії, а згодом і СРСР [46, 82, 93]. Високий рівень урбанізації (понад 80 % населення) сприяв розмиванню локальних сільських популяційних структур та формуванню відносно гомогенного індустріального регіону, де інтенсивне змішування різних етнічних груп відбувалося переважно в межах великих промислових центрів.

Повне включення Харківської області (26 районів) до складу цього макрорегіонального блоку підтверджує, що навіть її особливий історичний статус — зокрема функціонування як першої столиці Української РСР та важливого освітньо-культурного центру — не перешкодив її популяційно-генетичній інтеграції до східного індустриального ареалу. Слобожанська регіональна ідентичність Харківщини, попри певні історико-культурні відмінності від Донбасу, значною мірою нівелювалася під впливом спільних процесів індустріалізації та урбанізації ХХ століття [46, 82, 93].

Центральний та лівобережний блок (Полтавська, Сумська, Чернігівська, Черкаська, частина Київської області) формує складний мозаїчний макрорегіон із помірним рівнем внутрішньої гетерогенності. Ці території мають давню історію безперервного українського заселення, що сягає періоду формування Лівобережної Гетьманщини у XV–XVII століттях, однак упродовж ХХ століття зазнали значних міграційних впливів, пов'язаних із процесами індустріалізації та урбанізації [46, 82, 93]. Полтавська область (23 – 25 районів залежно від метрики) виступає ядром історичного лівобережного козацького регіону, зберігаючи відносну популяційно-генетичну однорідність попри процеси радянської соціально-економічної уніфікації.

Сумська область демонструє виразну триpletну структуру, розподіляючись між трьома популяційно-генетичними ареалами: північним поліським, центральним лівобережним та південним степовим. Така структура відображає її транзитне географічне положення на межі трьох макрорегіонів. Чернігівська область, своєю чергою, переважно належить до північного поліського кластера (19–23 райони залежно від метрики), що може свідчити про вплив спільного поліського демографічного субстрату, сформованого в умовах близькості до сучасного білоруського прикордоння та подібних природно-екологічних умов Полісся.

Південний степовий пояс (Херсонська, Запорізька, Дніпропетровська, Миколаївська, Одеська області) демонструє найбільш складну внутрішню

структуру, що пов'язано з пізньою колонізацією цих територій у XVIII–XX століттях після ліквідації Кримського ханства у 1783 році. Формування цього макрорегіону супроводжувалося поєднанням кількох колонізаційних потоків: державних переселень селян із Центральної України та Правобережжя, вільної міграції населення з Росії, іноземної колонізації (греки, болгары, німці, молдаване), а також пізнішої індустріалізації кінця XIX – XX століття [46, 82, 93].

У межах південного поясу виокремлюються два основні субблоки. Перший — південно-східний степовий, який охоплює всю Запорізьку область, майже всю Херсонську область та більшу частину Дніпропетровської області і характеризується спільною історією степової колонізації та подальшої індустріалізації. Другий — причорноморський субблок, до якого належить більшість Миколаївської області та частина Одеської області; його специфіка пов'язана з портовою економікою та більш інтенсивними міграційними процесами, характерними для приморських територій. Одеська область займає проміжне положення в межах цього макрорегіону, розподіляючись між південно-східним степовим, причорноморським та подільським популяційно-генетичними ареалами залежно від конкретних районів.

3.4.4.7. Чотири філогенетичні осі популяційної структури України

Інтеграція результатів кластерного аналізу, отриманих із використанням трьох метрик генетичної дистанції, дозволяє виокремити чотири основні філогенетичні осі, що визначають популяційно-генетичну структуру України на районному рівні.

Вісь Схід–Захід є найпотужнішою і пояснює найбільшу частку міжкластерної варіації. Вона відображає фундаментальний контраст між індустріалізованим і високоурбанізованим сходом країни, де відбувалася інтенсивна міграційна гомогенізація населення, та переважно аграрним заходом із краще збереженими локальними автохтонними популяційними субстратами. Східні області (Донецька, Луганська, Харківська) формують відносно генетично гомогенний регіональний блок, що є наслідком

масштабних міграційних процесів XIX–XX століть. Натомість західні області (Волинська, Львівська, Тернопільська) зберегли значно більшу внутрішню диференціацію, що пов'язано з нижчим рівнем індустріалізації та пізнішим входженням цих територій до складу СРСР.

Ця вісь має глибоке історико-демографічне підґрунтя [46, 82, 93]. Схід України формувався як колонізаційний фронтір Російської імперії з інтенсивним змішуванням різних етнічних груп, тоді як західні регіони тривалий час розвивалися в межах політичних структур Речі Посполитої та Австро-Угорщини, де зберігалася відносно вища етнічна однорідність сільського населення. Подальша радянська індустріалізація XX століття лише посилила цей контраст, перетворивши східні території на високоурбанізований індустріальний регіон, тоді як західна частина країни зберегла переважно аграрний характер соціально-економічного розвитку.

Вісь Північ–Південь відображає контраст між стабільними північними територіями з давньою історією безперервного заселення (Полісся та Лівобережжя, заселені з XV–XVII століть) і південними степовими регіонами пізньої колонізації XVIII–XX століть, для яких характерне інтенсивне змішування населення [46, 82, 93]. Північні області (Чернігівська, Сумська, північна частина Житомирської області) зберегли відносно стабільну популяційну структуру, тоді як південні регіони (Херсонська, Запорізька, Миколаївська області) сформувалися внаслідок поєднання кількох колонізаційних потоків.

Ця вісь має виразне екологічне підґрунтя. Поліські ліси та лісостепові райони Півночі забезпечували стабільні умови для традиційного сільського господарства і тривалого осілого заселення, тоді як південні степи до кінця XVIII століття залишалися прикордонною зоною з високою мобільністю населення. Пізня колонізація південних територій після ліквідації Кримського ханства сформувала демографічно молодий регіон із підвищеним рівнем генетичного змішування населення [46, 82, 93].

Вісь рівнина–гори проявляється у розділенні Івано-Франківської та Чернівецької областей між рівнинними та гірськими кластерами в моделях дистанцій Ласкера та Нея. Гірська ізоляція сприяла збереженню архаїчних популяційно-генетичних характеристик населення гуцулів та буковинців, які певною мірою відрізняються від рівнинних галичан, попри спільну західноукраїнську культурно-історичну ідентичність. Гірські території загалом характеризуються меншими ефективними розмірами популяцій, вищим рівнем ізоляції та більшим збереженням локальних генетичних варіантів.

Ця вісь має як екологічне, так і історико-демографічне підґрунтя. Карпатський гірський масив тривалий час виступав природним бар'єром для міграційних потоків, що сприяло формуванню та збереженню генетичної своєрідності гірських популяцій. Натомість рівнинні території Західної України (передусім Львівська та Тернопільська області) були значно більш інтегрованими в систему торговельних і міграційних шляхів, що сприяло інтенсивнішому міжрегіональному демографічному обміну.

Вісь індустріалізації відображає вплив урбанізації та індустріалізації на демографічну структуру регіонів. Високоіндустріалізовані регіони (Донбас, Дніпропетровська, Запорізька, Харківська області) формують відносно генетично гомогенний блок унаслідок масових міграційних потоків і інтенсивного змішування в промислових містах. Натомість переважно аграрні регіони (Волинь, Вінниччина, Тернопільщина) зберігають більшу внутрішню популяційно-генетичну різноманітність, що пов'язано з тривалим домінуванням традиційної сільської моделі розселення

Ця вісь є найбільш динамічною: індустріалізація ХХ століття істотно трансформувала популяційну структуру східних і південно-східних областей, сформувавши нові урбанізовані популяції з високим рівнем генетичного змішування [46, 82, 93]. Водночас західні аграрні регіони, які значною мірою уникнули масштабної індустріалізації, зберегли традиційну сільську популяційну структуру з виразнішими локальними генетичними особливостями.

3.4.4.8. Методологічне значення порівняння трьох метрик дистанцій

Порівняльний аналіз кластерних структур, отриманих із використанням трьох різних метрик дистанцій, має важливе методологічне значення для популяційно-генетичних досліджень. Кожна з метрик відображає різні аспекти популяційної структури відповідно до своїх математичних властивостей і біологічних припущень.

Евклідова дистанція є метрикою загального призначення, яка трактує алельні частоти як координати в багатовимірному просторі без додаткових корекцій. Вона є оптимальною для відображення сучасного генетичного стану популяцій і найкраще розрізняє внутрішньообласну гетерогенність. найбільш чутливою до виявлення внутрішньообласної гетерогенності. У результаті застосування цієї метрики сформувалася найбільш збалансована 11-кластерна структура з діапазоном розмірів кластерів від 5 до 107 районів, що відображає реальну складність популяційної структури України без значних узагальнень або спрощень.

Перевагою евклідової дистанції є її висока чутливість до локальних відмінностей. Саме вона дозволила виявити триpletне розщеплення Одеської області (5 + 2 + 19 районів), а також максимальну внутрішньообласну гетерогенність Київської області, райони якої розподіляються між п'ятьма різними кластерами. Водночас обмеженням цієї метрики є відсутність корекції на демографічні процеси, що потенційно може призводити до переоцінки ролі генетичного дрейфу в малих або ізольованих популяціях

Дистанція Ласкера, коригуючи значення генетичних відстаней з урахуванням ефективних розмірів популяцій, дозволяє виявляти структури, що відображають баланс між генетичним дрейфом і потоком генів. Ця метрика є особливо придатною для аналізу великомасштабних демографічних процесів, таких як колонізація та міжрегіональні міграції. Формування мегакластера 9 (108 районів — 22 % усіх районів України) у цьому випадку не є методологічним артефактом, а відображає реальну історичну спільність

південно-східних і частини центральних територій як регіонів пізньої колонізації з подібними джерелами міграційних потоків.

Перевагою дистанції Ласкера є її здатність виявляти глибші демографічні взаємозв'язки між регіонами. Зокрема, розділення галицько-буковинського блоку на рівнинне та гірське ядра відображає реальний екологічний бар'єр, який не настільки чітко проявляється при використанні евклідової дистанції. Аналогічно, компактизація Одеської області (23 з 26 районів у межах одного кластера) свідчить про наявність спільної популяційної історії більшості території області, попри локальні відмінності між окремими районами. Водночас обмеженням цієї метрики є тенденція до спрощення кластерної структури через злиття дрібніших кластерів, що інколи може маскувати реальну локальну популяційну диференціацію.

Дистанція Нея є філогенетично однією з найкоректніших метрик генетичної дистанції, оскільки враховує внутрішньопопуляційну гетерозиготність і є оптимальною для реконструкції еволюційних взаємовідносин між популяціями. Завдяки цьому вона здатна виявляти глибинні філогенетичні зв'язки, які не завжди очевидні з географічної або адміністративної перспективи. Зокрема, об'єднання восьми районів Одеської області зі східним кластером (70 районів загалом) відображає спільність південно-східного колонізаційного фронтиру XVIII–XIX століть — демографічного процесу, який маскується географічною відокремленістю в інших метриках.

Перевагою дистанції Нея є її здатність виявляти тонкі еволюційні структури популяцій. Зокрема, триpletне розщеплення Сумської області між північним, центральним і південним популяційно-генетичними ареалами підтверджує її транзитний характер на перетині трьох макрорегіонів. Аналогічно, виявлення максимальної внутрішньообласної гетерогенності Кіровоградської області (розподіл між чотирма кластерами) відображає її позицію як демографічного «перехрестя» України. Водночас обмеженням цієї

метрики є складність інтерпретації деяких кластерних зв'язків, які можуть суперечити географічній інтуїції, як, наприклад, виявлена спорідненість між частиною Одеської області та східним індустріальним регіоном.

Консенсусна ізоляція шести популяційно-генетичних одиниць (Крим, Закарпаття, Волинь, Вінниччина, Львівщина, Донбас) у всіх трьох метриках є найпереконливішим свідченням їхньої філогенетичної реальності. Топологічна стабільність цих кластерів указує на те, що вони відображають реальні популяційно-генетичні структури, а не є артефактами конкретної метрики дистанції або методу кластеризації.

3.4.4.9. Внутрішньообласна гетерогенність як індикатор демографічної складності

Систематичний аналіз розподілу районів кожної області між кластерами дозволяє встановити ієрархію областей за рівнем внутрішньої популяційно-генетичної гетерогенності. Цей рівень є прямим індикатором складності демографічної історії регіону: області з простою лінійною історією формують монолітні кластери, тоді як області з багат шаровою історією розщеплюються між множинними кластерами.

Області з мінімальною гетерогенністю (1 кластер) включають сім адміністративних одиниць, які консенсусно належать до одного кластера у всіх трьох метриках: АР Крим (15 районів), Закарпатська (13), Волинська (16), Вінницька (27), Львівська (20), Донецька (18), Луганська (18). Ці області представляють три різні типи демографічної історії.

Перший тип — екстремальна географічна або політична ізоляція (Крим, Закарпаття). Крим відокремлений від континентальної України Перекопським перешийком і має унікальну історію з кримськотатарським субстратом, пізньою колонізацією та багатоетнічним складом. Закарпаття ізольоване Карпатським хребтом і має угорську історію до 1945 року. Обидва регіони зберегли популяційно-генетичну унікальність через природні та політичні бар'єри.

Другий тип — збережені автохтонні субстрати аграрних регіонів (Волинь, Вінниччина, Львівщина). Ці області уникнули масової індустріалізації та міграції, зберігши традиційну сільську популяційну структуру. Волинь і Львівщина додатково захищені пізнім входженням до СРСР (1939), що мінімізувало вплив радянської уніфікації. Вінниччина, попри раніше входження до складу СРСР, зберегла однорідність через відсутність великих міст і промисловості [46, 82, 93].

Третій тип — міграційна гомогенізація індустріальних регіонів (Донецька, Луганська). Донбас сформувався як генетично однорідний індустріальний регіон через масову міграцію XIX–XX століть, яка «стерла» локальні відмінності і створила нову урбанізовану популяцію з високим рівнем змішування [46, 82, 93]. Парадоксально, екстремальна індустріалізація призвела до такої ж монолітності, як і аграрна збереженість західних областей, але через протилежні механізми.

Області з проміжною гетерогенністю (2–3 кластери) включають Хмельницьку (розділена між подільським і волинсько-поліським кластерами), Житомирську (між поліським і волинсько-поліським), Чернівецьку (між буковинсько-гірським і подільським), Сумську (між північним, центральним і південним ареалами), Рівненську (переважно волинсько-поліська з поліськими та галицькими аутлаєрами). Ці області займають транзитні позиції на межах макрорегіонів або мають помірну внутрішню екологічну диференціацію.

Хмельницька область (20 районів) розділяється між подільським кластером (11–12 районів) і волинсько-поліським (8–9 районів), що відображає її позицію на межі Поділля та Волинського Полісся. Північні райони області мають більшу спорідненість з Волиню, тоді як центральні та південні належать до подільського ареалу Вінниччини.

Житомирська область (23 райони) розділяється між поліським кластером (12 районів за евклідовою дистанцією, 20–21 за іншими метриками) і волинсько-поліським, що відображає градієнт від екстремального

Житомирського Полісся на півночі до перехідних зон на півдні. Дистанція Ласкера та Нея об'єднують більшість Житомирщини в один поліський блок, тоді як евклідова дистанція виявляє більшу диференціацію.

Чернівецька область (11 районів) розділяється між гірським буковинським кластером (7–11 районів) і подільським (3–4 райони), що відображає контраст між гірською Буковиною та рівнинними районами на межі з Вінниччиною. Дистанція Ласкера об'єднує всю Буковину в один гірський кластер, тоді як евклідова дистанція та дистанція Нея виявляють подільську компоненту.

Сумська область (18 районів) демонструє найскладнішу триплетну структуру: 8 районів у північному поліському кластері, 8 у південному степовому, 2 у центральному лівобережному (за дистанцією Нея). Ця структура підтверджує транзитну природу Сумщини на перетині Полісся, Лівобережжя та степового Півдня. Різні частини області мають різну демографічну історію: північні райони тяжіють до Чернігівщини та білоруського субстрату, центральні — до Полтавщини, південні — до степового Сходу.

Області з максимальною гетерогенністю (4–5 кластерів) включають чотири адміністративні одиниці з найскладнішою демографічною історією: Одеська (розділена між 3 – 4 кластерами залежно від метрики), Київська (між 4 – 5 кластерами), Кіровоградська (між 4 кластерами), Івано-Франківська (між 2 кластерами, але з чітким рівнинно-гірським розділенням).

Одеська область є абсолютним лідером за внутрішньою гетерогенністю. Евклідова дистанція виявляє триплетну структуру: 5 районів у мінімальному кластері 1 (Саратський, Ренійський, Болградський, Арцизький, Тарутинський), 2 райони в подільському кластері 5, 19 районів у причорноморському кластері 10.

П'ять районів кластера 1 формують компактний південно-західний блок Одещини з унікальною демографічною історією. Ренійський район є частиною Придунав'я з інтенсивним молдавсько-румунським впливом

та унікальним етнічним складом, сформованим через близькість до дунайської дельти. Болградський, Арцизький і Тарутинський райони належать до зони болгарської колонізації початку XIX століття: після Російсько-турецької війни 1806-1812 років болгари масово переселялися до Бессарабії, формуючи компактні поселення на півдні Одещини. Саратський район також має значну болгарську присутність і розташований у південно-західній периферії області.

Генетична ізоляція цих п'яти районів від решти Одещини відображає три фактори: етнічну специфіку (болгарсько-молдавський субстрат замість українсько-російського), географічну периферійність (південно-західна окраїна області на кордоні з Молдовою та Румунією), історичну відокремленість (Придунав'я та Буджак мали окрему адміністративну історію у складі Бессарабської губернії). Цей кластер є найяскравішим прикладом збереження локального етнічного субстрату в межах однієї області через компактність розселення та географічну ізоляцію від обласного центру. Ця екстремальна гетерогенність є прямим наслідком унікальної багатоетнічної історії Одещини. Південний захід області (Болградський, Ізмаїльський, Татарбунарський райони) заселений греками та болгарами у XVIII–XIX століттях. Задунайські райони (Ізмаїльський, Кілійський, Рені) мають молдавсько-румунський вплив. Прибережні райони з портовою Одесою характеризуються космополітичним складом з інтенсивною міграцією. Степові райони центральної та північної Одещини заселені переважно українцями та росіянами у XIX столітті. Жодна інша область не має такої різноманітності накладених один на одного демографічних субстратів.

Київська область (25 районів) розділяється між 4–5 кластерами: 19 районів у центральному кластері 3 (за евклідовою дистанцією), 2 райони у поліському кластері 4, 3 райони в подільському кластері 5, 1 район у волинсько-поліському кластері 6. Дистанція Ласкера та Нея дають подібні розщеплення з перерозподілом районів між кластерами.

Гетерогенність Київської області пояснюється контрастом між різними частинами регіону. Північні поліські райони (ймовірно Поліський,

Іванківський) мають більшу спорідненість з Житомирським Поліссям і навіть білоруським субстратом. Райони, прилеглі до столичної агломерації (Бориспільський, Броварський, Бучанський, Вишгородський), зазнали екстремальної урбанізації та міграції у ХХ–ХХІ століттях, що радикально змінило їхню популяційну структуру. Південні та південно-західні райони зберегли більш традиційну центральноукраїнську демографічну структуру. Столичний статус створив унікальну популяційну мозаїку з контрастами, які не спостерігаються в інших областях.

Кіровоградська область (21 район) розщеплена між чотирма кластерами за дистанцією Нея: 10 районів у причорноморському кластері 11, 5 районів у південно-східному степовому кластері 8, 3 райони в подільському кластері 4, 3 райони у центральному кластері 9. Ця структура підтверджує позицію Кіровоградщини як демографічного "перехрестя" України без власного чіткого популяційно-генетичного ядра.

Північні райони Кіровоградської області тяжіють до Вінниччини (подільський ареал), західні — до Черкащини (центрального ареал), південні — до Миколаївщини (причорноморський ареал), східні — до Дніпропетровщини (південно-східний степовий ареал). Область розташована на перетині чотирьох макрорегіональних блоків і не має власної домінуючої популяційно-генетичної ідентичності. Це типовий приклад транзитної території, яка сформувалася як адміністративна одиниця без попереднього історико-демографічного ядра.

Івано-Франківська область (14 районів) демонструє чітке рівнинно-гірське розділення у дистанції Ласкера та Нея: 5 районів у галицькому кластері (ймовірно рівнинні Прикарпатські райони), 9 районів у буковинсько-гірському кластері (ймовірно Верховинський, Косівський, Надвірнянський та інші Карпатські райони). Евклідова дистанція об'єднує всі 14 районів у єдиний галицько-буковинський блок, не виявляючи цієї диференціації.

Розділення відображає екологічний бар'єр між рівниною та горами: гірські гуцульські райони мають меншу густоту населення, вищий рівень

ізоляції та зберегли архаїчні популяційно-генетичні риси. Рівнинні райони більш інтегровані в галицький культурний та демографічний простір. Зважування на ефективні розміри популяцій (Ласкер) та філогенетичні відстані (Ней) виявляє цю диференціацію, яка маскується у простій евклідовій дистанції.

3.4.4.10. Висновки

Застосування методу Уорда до матриць генетичних дистанцій 490 районів України з використанням трьох альтернативних метрик (евклідова дистанція, дистанція Ласкера, дистанція Ней) виявило складну ієрархічну популяційно-генетичну структуру з високою топологічною стабільністю ключових елементів.

Консенсусна ізоляція шести популяційно-генетичних одиниць (АР Крим, Закарпатська, Волинська, Вінницька, Львівська області, Східна Україна) у всіх трьох метриках є найсильнішим доказом їхньої філогенетичної реальності. Ці регіони не є артефактами методу кластеризації, а відображають справжні демографічні одиниці з окремими історіями формування та чіткими генетичними межами.

Метрикоспецифічні ефекти відображають різні аспекти популяційної структури: евклідова дистанція оптимальна для виявлення сучасного генетичного стану та внутрішньообласної гетерогенності; дистанція Ласкера виявляє великомасштабні демографічні процеси через зважування на ефективні розміри популяцій; дистанція Ней реконструює найтонші еволюційні взаємовідносини, виявляючи глибинні філогенетичні зв'язки.

Сучасна генетична структура української популяції є результатом складної багат шарової демографічної історії з чотирма основними формувальними процесами: збереження автохтонних субстратів на Заході через пізнє входження до ЄС та мінімальну індустріалізацію; міграційна гомогенізація Сходу через масову індустріалізацію XIX–XX століть; пізня колонізація Півдня з змішуванням різних потоків після ліквідації Кримського

ханства; географічна ізоляція периферійних регіонів (Крим, Закарпаття) через природні та історико-політичні бар'єри.

Чотири основні філогенетичні осі (Схід–Захід, Північ–Південь, рівнина–гори, вісь індустріалізації) визначають популяційно-генетичну структуру України на районному рівні, пояснюючи максимальну частку міжкластерної варіації. Вісь Схід–Захід є найпотужнішою і відображає фундаментальний контраст між індустріалізованим урбанізованим Сходом і аграрним Заходом зі збереженими автохтонними субстратами.

Рівень внутрішньообласної гетерогенності є потужним індикатором складності демографічної історії: області з простою лінійною історією формують монолітні кластери (Волинь, Вінниччина, Львівщина, Східна Україна), тоді як області з багат шаровою історією розщеплюються між множинними кластерами (Одеська, Київська, Кіровоградська). Одеська область є абсолютним лідером за внутрішньою гетерогенністю через унікальну багатоетнічну історію з накладенням грецької, болгарської, молдавської, української та російської колонізацій.

Метод Уорда виявився оптимальним інструментом для виявлення складної ієрархічної структури українських популяцій, забезпечуючи високу роздільну здатність як на макрорегіональному, так і на внутрішньообласному рівнях. Топологічна стабільність ключових кластерів у трьох різних метриках підтверджує надійність виявленої структури та її відповідність реальним філогенетичним взаємовідносинам субпопуляцій України.

УЗАГАЛЬНЕННЯ ТА ВИСНОВКИ ДО РОЗДІЛУ 3

Комплексний аналіз популяційно-генетичної структури України на основі розподілу 697 147 унікальних прізвищ виявив складну багат шарову організацію з чіткими регіональними закономірностями. Загальноукраїнський індекс ізонімії становить $2,1 \times 10^{-4}$, коефіцієнт інбридингу — $5,3 \times 10^{-5}$, індекс різноманітності — $47,0 \times 10^2$, що вказує на відносно низький рівень генетичної ізоляції та високу гетерогенність [101]. Виявлено південно-східний —

північно-західний градієнт: ізонімія зростає від $2,2 \times 10^{-4}$ до $3,4 \times 10^{-4}$, різноманітність знижується від $44,8 \times 10^2$ до $29,4 \times 10^2$.

Застосування методу CRSED дозволило класифікувати регіони на три категорії: регіони імміграції (східні та південні області, сформовані індустріалізацією XIX–XX ст.), рекультивовані регіони (центральні області після демографічних потрясінь 1930–1940-х років) та регіони еміграції (західні та північні області з історичним відтоком населення). Районний аналіз 490 одиниць виявив екстремальну варіабельність: ізонімія від $1,6 \times 10^{-4}$ до $98,5 \times 10^{-4}$ (61-кратна різниця), інбридинг від $3,4 \times 10^{-5}$ до $246,2 \times 10^{-5}$ (72-кратна різниця).

Кластерний аналіз трьох метрик дистанцій виявив 10–11 популяційних кластерів з шістьма консенсусними філогенетичними одиницями (АР Крим, Закарпаття, Волинь, Вінниччина, Львівщина, Східна Україна) та чотирма основними осями структури (Схід–Захід, Північ–Південь, рівнина–гори, індустріалізація). Одеська область є найгетерогеннішою (розподіл між 3–4 кластерами) через багатоетнічну історію, тоді як шість областей демонструють монолітну структуру.

Сучасна структура сформована чотирма процесами: збереженням автохтонних субстратів на Заході, міграційною гомогенізацією Сходу через індустріалізацію, пізньою колонізацією Півдня після 1783 року та географічною ізоляцією периферійних регіонів, що обумовило формування одного з найрізноманітніших генофондів серед європейських популяцій.

НАУКОВЕ ТА ПРАКТИЧНЕ ЗНАЧЕННЯ РОБОТИ

Наукове значення дисертації полягає у створенні першої комплексної характеристики популяційно-генетичної структури України на двох ієрархічних рівнях з використанням найповнішого масиву даних (46,4 млн носіїв 697 тис. прізвищ), що формує базову референтну точку для майбутніх досліджень демографічних змін. Вперше для українських популяцій застосовано метод CRSED, що дозволило класифікувати регіони за типом демографічної історії, та розроблено багатометричний підхід до кластерного

аналізу з порівнянням трьох типів генетичних дистанцій, що забезпечує валідацію популяційної структури. Ідентифіковано шість консенсусних популяційно-генетичних одиниць (АР Крим, Закарпаття, Волинь, Вінниччина, Львівщина, Східну Україну) та чотири основні філогенетичні осі (Схід–Захід, Північ–Південь, рівнина–гори, індустріалізація), що відображають фундаментальні закономірності формування генетичного різноманіття України.

Практичне значення роботи визначається можливістю застосування результатів у медичній генетиці для ідентифікації зон підвищеного ризику спадкових захворювань на основі виявлених регіональних відмінностей коефіцієнта інбридингу (від $4,0 \times 10^{-5}$ у Криму до $17,9 \times 10^{-5}$ на Волині), що дозволяє оптимізувати медико-генетичне консультування та скринінгові програми у регіонах з підвищеним інбридингом. Детальна реконструкція міграційних маршрутів та популяційних взаємозв'язків сприяє розумінню процесів формування українського етносу в антропології, а кількісні оцінки міграційної інтенсивності (індекс міграції від $0,02 \times 10^{-4}$ до $1168,6 \times 10^{-4}$ на районному рівні) дозволяють реконструювати історичну соціальну мобільність та прогнозувати сучасні демографічні процеси. Отримані результати є референтною точкою для моніторингу наслідків сучасних демографічних потрясінь (події 2014 року на Донбасі, масові міграції з 2022 року), оскільки швидкий порівняльний аналіз майбутніх даних прізвищ дозволить виявити регіони з найбільшими змінами популяційної структури без необхідності дорогого молекулярно-генетичного тестування. Виявлення областей з екстремальною внутрішньою гетерогенністю (Одеська — розподіл між 3–4 кластерами, Київська — між 4–5 кластерами) вказує на необхідність диференційованого підходу до організації медичних служб з урахуванням різної генетичної структури субрегіонів, а регіональна специфіка розподілу прізвищ може використовуватися як допоміжний статистичний інструмент для оцінки ймовірного географічного походження. Робота закладає методологічну основу для часового аналізу динаміки популяційної структури, інтеграції

з молекулярно-генетичними даними та детального мікрорегіонального аналізу, що визначає перспективи подальших міждисциплінарних досліджень у генетиці, демографії, антропології, історії та медицині.

ВИСНОВКИ

Комплексний аналіз розподілу українських прізвищ по всіх областях України виявив складну багаторівневу популяційно-генетичну структуру; розраховані показники ізонімії, інбридингу та різноманітності прізвищ узгоджуються з демографічними наслідками міграційних, соціально-економічних та історико-політичних процесів.

1. Порівняння України з понад двадцятьма країнами світу за основними популяційно-генетичними показниками, зокрема ізонімією та різноманітністю прізвищ, показало, що Україна належить до популяцій з високою різноманітністю та низькою ізонімією поряд із західноєвропейськими країнами, маючи найвищий показник різноманітності прізвищ серед досліджених європейських популяцій.

2. Розраховані для 490 адміністративних районів України популяційно-генетичні показники — показник внутрішньогрупової ізонімії (I^{jj} : $1,6 \times 10^{-4} - 98,5 \times 10^{-4}$), коефіцієнт інбридингу (F : $3,4 \times 10^{-5} - 246,2 \times 10^{-5}$) та індекс різноманітності Фішера (α : $2,49 \times 10^2 - 63,18 \times 10^2$) — виявили значну міжрайонну варіабельність, що відображає демографічні наслідки індустріалізації XIX–XX ст., колонізації Півдня після 1783 р. та переважання відтоку населення на Заході.

3. Типологія регіонів за демографічною історією визначена на основі показника CRSED, на підставі чого виділено три категорії:

— регіони імміграції ($CRSED \leq 0,01$): АР Крим, Донецька, Луганська, Харківська, Запорізька, Дніпропетровська області — найнижча ізонімія ($1,61 - 2,92 \times 10^{-4}$), зумовлена інтенсивним міграційним припливом у XIX–XX ст.;

— рекультивовані регіони ($0,01 < CRSED \leq 0,05$): Полтавська, Черкаська, Хмельницька, Вінницька, Тернопільська, Сумська, Житомирська області — наслідок масових демографічних потрясінь 1930–1940-х рр.;

— регіони еміграції ($CRSED > 0,05$): Львівська, Волинська, Рівненська, Чернівецька, Чернігівська області — підвищена ізонімія ($4,57 - 7,14 \times 10^{-4}$), зумовлена переважанням відтоку населення.

4. Субрегіональна кластерна структура населення виявлена за результатами аналізу 490 адміністративних районів методом Уорда з використанням трьох метрик генетичних дистанцій (евклідової, Ласкера, Нея). Визначено 10–11 основних популяційних кластерів; шість стабільних філогенетичних одиниць незалежно від метрики (АР Крим, Закарпатська, Волинська, Вінницька, Львівська обл., Східна Україна), з Одеською обл. як найгетерогеннішим регіоном (розподіл між 3–4 кластерами). Кластерний аналіз виявив чотири основні філогенетичні осі популяційної структури: Схід–Захід, Північ–Південь, рівнина–гори та індустріалізація.

5. Тест Мантеля виявив статистично значущу кореляцію між генетичними та географічними відстанями ($p < 0,001$): коефіцієнти кореляції становлять 0,581 (евклідова дистанція), 0,691 (дистанція Ласкера) та 0,246 (дистанція Нея), що підтверджує наявність моделі ізоляції за відстанню в українських популяціях.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Allentoft ME, Sikora M, Sjögren KG, Rasmussen S, Rasmussen M, Stenderup J, et al. Population genomics of bronze age Eurasia. *Nature*. 2015;522(7555):167–72. DOI: [10.1038/nature14507](https://doi.org/10.1038/nature14507)
2. Ayala FJ, Kiger JA Jr. *Modern genetics*. Menlo Park (CA): Benjamin/Cummings; 1980. 844 p. ISBN 0-8053-0312-X
3. Baek SK, Kiet HAT, Kim BJ. Family name distributions: Master equation approach. *Phys Rev E*. 2007;76(4):046113. DOI: [10.1103/PhysRevE.76.046113](https://doi.org/10.1103/PhysRevE.76.046113)
4. Balanovskaya E, Balanovsky O. *The gene pool of Russians on Russian plain*. M: Luch; 2007. 416 p. ISBN 978-5-87140-267-2
5. Balanovskaya EV, Romanov AG, Balanovsky OP. Namesakes or relatives? Approaches to investigating the relationship between Y chromosome haplogroups and surnames. *Mol Biol*. 2011;45(3):473–85. DOI: [10.1134/s0026893311030022](https://doi.org/10.1134/s0026893311030022)
6. Balanovsky O, Pshenichnov F, Rootsi S, Churnosov M, Atramentova L, Tegako L, et al. Intra-ethnic variation of the Y chromosome in European countries: a comparative study. In: *The Fifth ISABS Conference in Forensic Genetics and Molecular Anthropology*; 2007; Split, Croatia. p. 121.
7. Barraï I, Rodríguez-Larralde A, Mamolini E, Manni F, Scapoli C. Elements of the surname structure of Austria. *Ann Hum Biol*. 2000;27(6):607–22. DOI: [10.1080/03014460050178696](https://doi.org/10.1080/03014460050178696)
8. Barraï I, Rodriguez-Larralde A, Manni F, Ruggiero V, Tartari D, Scapoli C. Isolation by language and distance in Belgium. *Ann Hum Genet*. 2004;68(1):1–16. DOI: [10.1046/j.1529-8817.2003.00044.x](https://doi.org/10.1046/j.1529-8817.2003.00044.x)
9. Barraï I, Rodriguez-Larralde A, Mamolini E, Scapoli C. Isonymy and isolation by distance in Italy. *Hum Biol*. 1999;71(6):947–61.
10. Barraï I, Rodriguez-Larralde A, Manni F, Scapoli C. Isonymy and isolation by distance in the Netherlands. *Hum Biol*. 2002;74(2):263–83. DOI: [10.1353/hub.2002.0016](https://doi.org/10.1353/hub.2002.0016)

11. Barraï I, Scapoli C, Beretta M, Nesti C, Mamolini E, Rodríguez-Larralde A. Isonymy and the genetic structure of Switzerland I. The distributions of surnames. *Ann Hum Biol.* 1996;23(6):431–55. DOI: [10.1080/03014469600004672](https://doi.org/10.1080/03014469600004672)
12. Barraï I, Rodríguez-Larralde A, Mamolini E, Manni F, Scapoli C. Isonymy structure of USA population. *Am J Phys Anthropol.* 2001;114(2):109–23. DOI: [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)
13. Barraï I, Formica G, Scapoli C, Beretta M, Mamolini E, Volinia S, et al. Microevolution in Ferrara: isonymy 1890–1990. *Ann Hum Biol.* 1992;19(4):371–85. DOI: [10.1080/03014469200002242](https://doi.org/10.1080/03014469200002242)
14. Barraï I, Rodríguez-Larralde A, Dipierri J, Alfaro E, Acevedo N, Mamolini E, et al. Surnames in Chile: a study of the population of Chile through isonymy. *Am J Phys Anthropol.* 2012;147(3):380–88. DOI: [10.1002/ajpa.22000](https://doi.org/10.1002/ajpa.22000)
15. Branco CC, Mota-Vieira L. Surnames in the Azores: analysis of the isonymy structure. *Hum Biol.* 2005;77(1):37–44. DOI: [10.1353/hub.2005.0026](https://doi.org/10.1353/hub.2005.0026)
16. Cabello PH, Krieger H. Note on estimates of the inbreeding coefficient through study of pedigrees and isonymous marriages. *Hum Biol.* 1991;63(5):719–23.
17. Cai H. A society without fathers or husbands: the Na of China. New York: Zone Books; 2001. 505 p. ISBN: 1-890951-12-9
18. Carrieri A, Sans M, Dipierri JE, Alfaro E, Mamolini E, Sandri M, et al. The structure and migration patterns of the population of Uruguay through isonymy. *J Biosoc Sci.* 2020;52(2):300–14. DOI: [10.1017/S0021932019000476](https://doi.org/10.1017/S0021932019000476)
19. Cavalli-Sforza LL, Edwards AW. Phylogenetic analysis. Models and estimation procedures. *Am J Hum Genet.* 1967;19(3 Pt 1):233–57. DOI: [10.2307/2406616](https://doi.org/10.2307/2406616)
20. Chen J, Chen L, Liu Y, Li X, Yuan Y, Wang Y. An index of Chinese surname distribution and its implications for population dynamics. *Am J Phys Anthropol.* 2019;169(4):608–18. DOI: [10.1002/ajpa.23863](https://doi.org/10.1002/ajpa.23863)
21. Crow JF, Mange AP. Measurement of inbreeding from the frequency of marriages between persons of the same surname. *Eugen Q.* 1965;12(4):199–203. DOI: [10.1080/19485565.1965.9987630](https://doi.org/10.1080/19485565.1965.9987630)

22. Crow JF. The estimation of inbreeding from isonymy. *Hum Biol.* 1989;61(1):85–92.
23. Darlu P, Ruffié J. Relationships between consanguinity and migration rate from surname distributions and isonymy in France. *Ann Hum Biol.* 1992;19(2):133–37. DOI: [10.1080/03014469200002012](https://doi.org/10.1080/03014469200002012)
24. Dipierri J, Rodriguez-Larralde A, Alfaro E, Scapoli C, Mamolini E, Salvatorelli G, et al. A study of the population of Paraguay through isonymy. *Ann Hum Genet.* 2011;75(6):678–87. DOI: [10.1111/j.1469-1809.2011.00676.x](https://doi.org/10.1111/j.1469-1809.2011.00676.x)
25. Dipierri J, Rodríguez-Larralde A, Barraí I, Camelo JL, Redomero EG, Rodríguez CA, et al. Random inbreeding, isonymy, and population isolates in Argentina. *J Community Genet.* 2014;5(3):241–48. DOI: [10.1007/s12687-013-0181-x](https://doi.org/10.1007/s12687-013-0181-x)
26. Elzhov TV, Mullen KM, Spiess AN, Bolker B. minpack.lm: R Interface to the Levenberg-Marquardt Nonlinear Least-Squares Algorithm Found in MINPACK, Plus Support for Bounds [Internet]. Version 1.2–4; 2016. Available from: <https://cran.r-project.org/web/packages/minpack.lm/index.html>
27. Fisher RA, Corbet AS, Williams CB. The relation between the number of species and the number of individuals in a random sample of an animal population. *J Anim Ecol.* 1943;12(1):42–58. DOI: [10.2307/1411](https://doi.org/10.2307/1411)
28. Fort J. Demic and cultural diffusion propagated the Neolithic transition across different regions of Europe. *J R Soc Interface.* 2015;12(106). DOI: [10.1098/rsif.2015.0166](https://doi.org/10.1098/rsif.2015.0166)
29. Fox WR, Lasker GW. The distribution of surname frequencies. *Int Stat Rev.* 1983;51(1):81–87. DOI: [10.2307/1402733](https://doi.org/10.2307/1402733)
30. Gower JC. Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika.* 1966;53(3–4):325–38. DOI: [10.2307/2333639](https://doi.org/10.2307/2333639)
31. Haak W, Lazaridis I, Patterson N, Rohland N, Mallick S, Llamas B, et al. Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature.* 2015;522(7555):207–11. DOI: [10.1038/nature14317](https://doi.org/10.1038/nature14317)

32. Herrera Paz EF, Scapoli C, Mamolini E, Sandri M, Carrieri A, Rodriguez-Larralde A, et al. Surnames in Honduras: a study of the population of Honduras through isonymy. *Ann Hum Genet.* 2014;78(3). DOI: [10.1111/ahg.12057](https://doi.org/10.1111/ahg.12057)
33. Hill MO. Diversity and evenness: a unifying notation and its consequences. *Ecology.* 1973;54(2):427–32. DOI: [10.2307/1934352](https://doi.org/10.2307/1934352)
34. Jobling MA. In the name of the father: surnames and genetics. *Trends Genet.* 2001;17(6):353–57. DOI: [10.1016/s0168-9525\(01\)02284-3](https://doi.org/10.1016/s0168-9525(01)02284-3)
35. Kamel C, Saliba-Serre B, Lizee MH, Signoli M, Costedoat C. Surnames in south-eastern France: structure of the rural population during the 19th century through isonymy. *J Biosoc Sci.* 2023;55(1):174–89. DOI: [10.1017/S0021932021000699](https://doi.org/10.1017/S0021932021000699)
36. Karlin S, McGregor J. The number of mutant forms maintained in a population. In: *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*; 1967; Berkeley (CA). Vol. 4. Berkeley: University of California Press; 1967. p. 415–38.
37. Kimura M. *Outline of population genetics.* Tokyo: Baifukan; 1960.
38. Kruskal JB. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika.* 1964;29(1):1–27. DOI: [10.1007/BF02289565](https://doi.org/10.1007/BF02289565)
39. Kruskal JB. Nonmetric multidimensional scaling: a numerical method. *Psychometrika.* 1964;29(2):115–29. DOI: [10.1007/BF02289694](https://doi.org/10.1007/BF02289694)
40. Kushniarevich, A, Utevska O, Chuhryaeva M, Agdzhoyan A, Dibirova K, Uktveryte I, et al. Genetic heritage of the Balto-Slavic speaking populations: a synthesis of autosomal, mitochondrial and Y-chromosomal data. *PLoS One.* 2015;10(9):e0135820. DOI: [10.1371/journal.pone.0135820](https://doi.org/10.1371/journal.pone.0135820)
41. Laherrere J, Sornette D. Stretched exponential distributions in nature and economy: "fat tails" with characteristic scales. *Eur Phys J B.* 1998;2(4):525–39. DOI: [10.1007/s100510050276](https://doi.org/10.1007/s100510050276)

42. Larralde AR, Scapoli C, Mamolini E, Barraï I. Surnames in Texas: a population study through isonymy. *Hum Biol.* 2007;79(2):215–39. DOI: [10.1353/hub.2007.0034](https://doi.org/10.1353/hub.2007.0034)
43. Lasker GW. A coefficient of relationship by isonymy: a method for estimating the genetic relationship between populations. *Hum Biol.* 1977;49(3):489–93. Available from: <https://digitalcommons.wayne.edu/humbiol/vol49/iss3/21/>
44. Lasker GW. Surnames and genetic structure. Cambridge: Cambridge University Press; 1985. ISBN: 0521302854, 9780521302852.
45. Liu Y, Chen L, Yuan Y, Chen J. A study of surnames in China through isonymy. *Am J Phys Anthropol.* 2012;148(3):341–50. DOI: [10.1002/ajpa.22055](https://doi.org/10.1002/ajpa.22055)
46. Magocsi PR. A history of Ukraine: the land and its peoples. Toronto: University of Toronto Press; 2010. ISBN: 978-1442610217
47. Malécot G. The decrease of relationship with distance. *Cold Spring Harb Symp Quant Biol.* 1955;20:52–53.
48. Mantel N. The detection of disease clustering and a generalized regression approach. *Cancer Res.* 1967;27(2 Pt 1):209–20.
49. Matuzevičiūtė GM, Telizhenko S. The first farmers of Ukraine: an archaeobotanical investigation and AMS dating of wheat grains from the Ratniv-2 site. *Archaeol Lituana.* 2016;17:100–11. DOI: [10.15388/ArchLit.2016.17.10685](https://doi.org/10.15388/ArchLit.2016.17.10685)
50. Menozzi P, Piazza A, Cavalli-Sforza L. Synthetic maps of human gene frequencies in Europeans. *Science.* 1978;201(4358):786–92. DOI: [10.1126/SCIENCE.356262](https://doi.org/10.1126/SCIENCE.356262)
51. Mikerezi I, Xhina E, Scapoli C, Barbujani G, Mamolini E, Sandri M, et al. Surnames in Albania: a study of the population of Albania through isonymy. *Ann Hum Genet.* 2013;77(3):232–43. DOI: [10.1111/ahg.12015](https://doi.org/10.1111/ahg.12015)
52. Miyazima S, Lee Y, Nagamine T, Miyajima H. Power-law distribution of family names in Japanese societies. *Physica A.* 2000;278(1–2):282–88. DOI: [10.1016/S0378-4371\(99\)00546-4](https://doi.org/10.1016/S0378-4371(99)00546-4)

53. Moré JJ. The Levenberg-Marquardt algorithm: implementation and theory. In: Watson GA, editor. Numerical analysis: proceedings of the biennial Conference held at Dundee, June 28–July 1, 1977. Berlin: Springer; 1978. p. 105–16. DOI: [10.1007/bfb0067700](https://doi.org/10.1007/bfb0067700)
54. Murtagh F, Legendre P. Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion? *J Classif.* 2014;31(3):274–95. DOI: [10.1007/s00357-014-9161-z](https://doi.org/10.1007/s00357-014-9161-z)
55. Nei M. Genetic distance between populations. *Am Nat.* 1972;106(949):283–92. DOI: [10.1086/282771](https://doi.org/10.1086/282771)
56. Nei M. Molecular population genetics and evolution. Amsterdam; New York: North-Holland; 1975. 288 p. ISBN: 0444107517
57. Nei M. The theory and estimation of genetic distance. In: Morton NE, editor. Genetic structure of populations. Honolulu: University of Hawaii Press; 1973. p. 45–54.
58. Nikitin AG, Lazaridis I, Patterson N, Ivanova S, Videiko M, Dergachev V, et al. A genomic history of the North Pontic Region from the Neolithic to the Bronze Age. *Nature.* 2025;639(8053):1–8. DOI: [10.1038/s41586-024-08372-2](https://doi.org/10.1038/s41586-024-08372-2)
59. Patterson N, Price AL, Reich D. Population structure and eigenanalysis. *PLoS Genet.* 2006;2(12):e190. DOI: [10.1371/journal.pgen.0020190](https://doi.org/10.1371/journal.pgen.0020190)
60. Piazza A, Rendine S, Zei G, Moroni A, Cavalli-Sforza LL. Migration rates of human populations from surname distributions. *Nature.* 1987;329(6141):714–16. DOI: [10.1038/329714a0](https://doi.org/10.1038/329714a0)
61. Preda-Bălănică B, Frînculeasa A, Heyd V. The Yamnaya impact north of the Lower Danube. *Bull Soc Préhist Fr.* 2020;117(1):85–102. DOI: [10.3406/bspf.2020.15079](https://doi.org/10.3406/bspf.2020.15079)
62. R Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vienna: R Foundation for Statistical Computing; 2024. Available from: <https://www.R-project.org/>
63. Relethford JH. Estimation of kinship and genetic distance from surnames. *Hum Biol.* 1988;60(3):475–92.

64. Rode ZR. The origin of Jewish family names. *Names*. 1976;24(3):165–79. DOI: [10.1179/nam.1976.24.3.165](https://doi.org/10.1179/nam.1976.24.3.165)
65. Rodriguez-Larralde A, Barraí I, Nesti C, Mamolini E, Scapoli C. Isonymy and isolation by distance in Germany. *Hum Biol*. 1998;70(6):1041–56. Available from: <https://digitalcommons.wayne.edu/humbiol/vol70/iss6/6/>
66. Rodriguez-Larralde A, Scapoli C, Beretta M, Nesti C, Mamolini E, Barraí I. Isonymy and the genetic structure of Switzerland. II. Isolation by distance. *Ann Hum Biol*. 1998;25(6):533–40. DOI: [10.1080/03014469800006772](https://doi.org/10.1080/03014469800006772)
67. Rodríguez-Larralde A, Morales J, Barraí I. Surname frequency and the isonymy structure of Venezuela. *Am J Hum Biol*. 2000;12(3):352–62. DOI: [10.1002/\(SICI\)1520-6300\(200005/06\)12:3<352::AID-AJHB5>3.0.CO;2-S](https://doi.org/10.1002/(SICI)1520-6300(200005/06)12:3<352::AID-AJHB5>3.0.CO;2-S)
68. Rodriguez-Larralde A, Dipierri J, Gomez EA, Scapoli C, Mamolini E, Salvatorelli G, et al. Surnames in Bolivia: a study of the population of Bolivia through isonymy. *Am J Phys Anthropol*. 2011;144(2):177–84. DOI: [10.1002/ajpa.21379](https://doi.org/10.1002/ajpa.21379)
69. Rodriguez-Larralde A, Gonzales-Martin A, Scapoli C, Barraí I. The names of Spain: a study of the isonymy structure of Spain. *Am J Phys Anthropol*. 2003;121(3):280–92. DOI: [10.1002/ajpa.10209](https://doi.org/10.1002/ajpa.10209)
70. Rossi P. Surname distribution in population genetics and in statistical physics. *Phys Life Rev*. 2013;10(4):395–415. DOI: [10.1016/j.pprev.2013.06.005](https://doi.org/10.1016/j.pprev.2013.06.005)
71. Saag L, Utevska O, Zadnikov S, Shramko I, Gorbenko K, Bandrivskyi M, et al. North Pontic crossroads: mobility in Ukraine from the Bronze Age to the early modern period. *Sci Adv*. 2025;11(2):eadr0695. DOI: [10.1126/sciadv.adr0695](https://doi.org/10.1126/sciadv.adr0695)
72. Sanday PR. *Women at the center: life in a modern matriarchy*. Ithaca: Cornell University Press; 2002. 288 p. ISBN: 9780801489068
73. Scapoli C, Goebel H, Sobota S, Mamolini E, Rodriguez-Larralde A, Barraí I. Surnames and dialects in France: population structure and cultural evolution. *J Theor Biol*. 2005;237(1):75–86. DOI: [10.1016/j.jtbi.2005.03.035](https://doi.org/10.1016/j.jtbi.2005.03.035)
74. Schimmel AM. *Islamic names: an introduction*. Edinburgh: Edinburgh University Press; 1990. 137 p. ISBN: [0-85224-612-9](https://doi.org/10.1017/CBO9780511524444)

75. Schneider DM, Gough K, editors. Matrilineal kinship. Berkeley; Los Angeles: University of California Press; 1961. 776 p. [ISBN: 0-520-02529-6](#)
76. Shambadal P. Development and application of the concept of entropy. Moscow: Nauka; 1967. 279 p.
77. Shuba V, Yasinskyi Y, Atramentova L. The structure of the Ukrainian population based on surname distribution. *Am J Hum Biol.* 2025;37(7):e70097. DOI: [10.1002/ajhb.70097](#)
78. Simpson EH. Measurement of diversity. *Nature.* 1949;163(4148):688. DOI: [10.1038/163688a0](#)
79. Sinnott RW. Virtues of the Haversine. *Sky Telesc.* 1984;68(2):158.
80. Smouse PE, Long JC, Sokal RR. Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Syst Zool.* 1986;35(4):627–32. DOI: [10.2307/2413122](#)
81. Speicher MR, Motulsky AG, Antonarakis SE, editors. Vogel and Motulsky's human genetics. Berlin; Heidelberg: Springer; 2010. DOI: [10.1007/978-3-540-37654-5](#)
82. Subtelny O. Ukraine: a history. Toronto: University of Toronto Press; 2009. ISBN: 1442640162
83. Tarskaia L, El'Chinova GI, Scapoli C, Mamolini E, Carrieri A, Rodriguez-Larralde A, et al. Surnames in Siberia: a study of the population of Yakutia through isonymy. *Am J Phys Anthropol.* 2009;138(2):190–98. DOI: [10.1002/ajpa.20918](#)
84. Tay JSH, Yip WCL. The estimation of inbreeding from isonymy: relationship to the average inbreeding coefficient. *Ann Hum Genet.* 1984;48(2):185–94. DOI: [10.1111/j.1469-1809.1984.tb01013.x](#)
85. Ward JH Jr. Hierarchical grouping to optimize an objective function. *J Am Stat Assoc.* 1963;58(301):236–44. DOI: [10.2307/2282967](#)
86. Wickham H. Data analysis. In: *ggplot2: elegant graphics for data analysis.* Cham: Springer; 2016. p. 189–201. DOI: [10.1007/978-3-319-24277-4_9](#)

87. Wickham H, François R, Henry L, Müller K, Vaughan D. dplyr: A Grammar of Data Manipulation [Internet]. Version 1.1.2; 2023. Available from: <https://CRAN.R-project.org/package=dplyr/>
88. Yasuda N. Studies of isonymy and inbreeding in Japan. *Hum Biol.* 1983;55(2):263–76.
89. Yasuda N, Morton NE. Studies on human population structure. In: *Third International Congress of Human Genetics*; 1966; Chicago. Baltimore: Johns Hopkins Press; 1967. p. 249–65.
90. Zei G, Lisa A, Fiorani O, Magri C, Quintana-Murci L, Semino O, et al. From surnames to the history of Y chromosomes: the Sardinian population as a paradigm. *Eur J Hum Genet.* 2003;11(10):802–07. DOI: [10.1038/sj.ejhg.5201040](https://doi.org/10.1038/sj.ejhg.5201040)
91. Zhang J, Zhu Y, Xiao H. The distribution of surnames in Xiantao City through isonymy. *J Geod Geoinform Sci.* 2024;7(2). DOI: [10.11947/j.JGGS.2024.0204](https://doi.org/10.11947/j.JGGS.2024.0204)
92. Атраментова ЛА, Горпинченко МЮ. Про можливість використання українських прізвищ для оцінки рівня інбридингу. Фактори експериментальної еволюції організмів. 2014;14:182–86.
93. Атраментова ЛО, Шуба ВВ, Горпинченко МЮ. Прізвища українського населення як інструмент популяційно-генетичних досліджень. *Вісн Харків нац ун-ту ім В Н Каразіна. Серія «Біологія».* 2024;43:44–49. DOI: [10.26565/2075-5457-2024-43-4](https://doi.org/10.26565/2075-5457-2024-43-4)
94. Бойко ОД. *Історія України.* Київ: Академвидав; 2018.
95. Горпинченко МЮ, Атраментова ЛО. Популяційно-генетичний аналіз населення Закарпатської області за даними прізвищ. Фактори експериментальної еволюції організмів. 2019;25:44–48. DOI: [10.7124/FEEEO.v25.1137](https://doi.org/10.7124/FEEEO.v25.1137)
96. Горпинченко МЮ, Атраментова ЛО. Розподіл індексу ізонімії на території України. Фактори експериментальної еволюції організмів. 2020;26:51–55. DOI: [10.7124/FEEEO.v26.1240](https://doi.org/10.7124/FEEEO.v26.1240)

97. Горпинченко МЮ, Утєвська ОМ, Атраментова ЛО. Значення показника індекс місця прізвища при вивченні українських популяцій. Вісн Львів ун-ту. Серія біологічна. 2015;70:74–81.
98. Горпинченко МЮ, Атраментова ЛО. Популяційно-генетичні характеристики населення України, отримані з використанням прізвищ. Вісн Київ нац ун-ту ім Тараса Шевченка. Серія: Біологія. 2015;1:68–72.
99. Горпинченко МЮ, Атраментова ЛО. Здатність показника І_p розрізняти українські популяції регіонального рівня. Фактори експериментальної еволюції організмів. 2015;16:192–96.
100. Горпинченко МЮ, Утєвська ОМ, Атраментова ЛО. Міграційна структура населення Валківського району Харківської області за даними про квазігенетичні маркери. Фактори експериментальної еволюції організмів. 2013;13:296–99.
101. Горпинченко МЮ. Можливість використання українських прізвищ в якості квазігенетичних маркерів в популяційно-генетичному аналізі населення України [дисертація в Інтернеті]. 2020. Available from: uacademic.info/ua/document/0421U100030
102. Горпинченко МЮ, Атраментова ЛО. Показник «індекс місця» для локальних популяцій районного рівня (на прикладі Одеської області). Фактори експериментальної еволюції організмів. 2016;18:227–30.
103. Горпинченко МЮ, Атраментова ЛО. Розмаїття українських прізвищ як інструмент популяційно-генетичних досліджень. Вісн Харків нац ун-ту ім В Н Каразіна. Серія: Біологія. 2014;22:54–60.
104. Горпинченко МЮ, Буркова ВВ, Утєвська ОМ, Атраментова ЛО. Українські прізвища як квазігенетичні маркери. Асоціація з групами крові та гаплогрупами Y-хромосоми. Вісн Харків нац ун-ту ім В Н Каразіна. Серія: Біологія. 2015;25:101–09.
105. Горпинченко МЮ, Атраментова ЛО. Характеристики Чернігівської популяції за даними прізвищ. Фактори експериментальної еволюції організмів. 2018;22:40–45. DOI: [10.7124/FEEO.v22.921](https://doi.org/10.7124/FEEO.v22.921)

106. Державний комітет статистики України. Підсумки Всеукраїнського перепису населення 2001 року. Чисельність і склад населення України [Інтернет]. 2001. Available from: <http://2001.ukrcensus.gov.ua/results/general/nationality>
107. Дністрянський МС. Етногеографія України. Львів: Видавничий центр ЛНУ імені Івана Франка; 2008. 154 с.
108. Редько ЮК. Довідник українських прізвищ. Варченко І, редактор. Київ: Радянська школа; 1968. 253 с.
109. Редько ЮК. Сучасні українські прізвища. Київ: Наукова думка; 1966. 214 с.
110. Сегеда СП. Антропология: навч. посіб. Київ: Либідь; 2001. 336 с. ISBN 966-06-0165-4
111. Сорока Ю. Насильницьке виселення населення західноукраїнських земель у 1940–1950-х роках. Вісн Київ нац ун-ту ім Тараса Шевченка. Історія. 2007;91–93:122–26.
112. Утєвська ОМ, Агджоян АТ, Балановська ОВ, Атраментова ЛО, Балановський ОП. Джерела формування українського генофонду за даними про Y-хромосому. Вісн Харків нац ун-ту ім В Н Каразіна. Серія «Біологія». 2013;1079:87–98 URL: http://seriesbiology.univer.kharkov.ua/vol_1079.html
113. Дібірова ХД, Балановський ОП, Кузнєцова МО, Фролова СА, Балаганська ОА, et al. Генетичний профіль армян по поліморфізму Y-хромосоми. В: З'їзд етнографів та антропологів Росії: матеріали ІХ наукової конференції; 2011 січ 4–8; Петрозаводськ. 2011. с. 545.
114. Утєвська ОМ. Генофонд українського народу в контексті навколишніх регіонів (за даними про поліморфізм Y хромосоми). В: Проблеми популяційної та загальної генетики: матеріали Міжнародної наукової конференції; 2011 лист 14–15; Москва. 2011. с. 150–51.
115. Утєвська ОМ. Гомогенність генофонду українців: аналіз історико-територіальних об'єднань. В: Наукова спадщина М І Вавилова та сучасність: матеріали наукової конференції; 2012 лют 2. 2012. с. 95.

116. Утєвська ОМ, Дібірова ХД, Чухряєва МІ, Агджоян АТ, Атраментова ЛО, Балановська ОВ, et al. Західноєвропейські та західноазіатські варіанти гаплогрупи R1b в українських популяціях. Фактори експериментальної еволюції організмів. 2015;16:255–59. Available from: <https://utgis.org.ua/journals/index.php/Factory/article/view/455>
117. Утєвська ОМ, Атраментова ЛО, Балановська ОВ, Балановський ОП. Мінливість STR локусів Y-хромосоми в популяціях українців. Вісн Київ нац ун-ту ім Тараса Шевченка. Серія: Біологія. 2015;2:55–61. DOI: [10.17721/1728_2748.2015.70.55-61](https://doi.org/10.17721/1728_2748.2015.70.55-61)
118. Утєвська ОМ, Чухряєва МІ, Схаляхо РО, Дібірова ХД, Теучеж ІЕ, Агджоян АТ, et al. Походження основних груп козацтва за даними про поліморфізм Y-хромосоми. Вісн Одес нац ун-ту. Серія: Біологія. 2015;20(37):61–69. DOI: [10.18524/2077-1746.2015.2\(37\).54983](https://doi.org/10.18524/2077-1746.2015.2(37).54983)
119. Утєвська ОМ, Пшеничнов ОС, Балановський ОП, Фролова СО, Кузнєцова МО, Романов ОГ, et al. Географічні та історичні закономірності у мінливості генофонду населення України та півдня Росії (аналіз ДНК-маркерів). Фактори експериментальної еволюції організмів. 2011;11:541–45. Available from: https://utgis.org.ua/images/pdf/factory/tom_11.pdf
120. Утєвська ОМ, Чухряєва МІ, Агджоян АТ, Атраментова ЛО, Балановська ОВ, Балановський ОП. Популяції Закарпаття та Буковини на генетичному фоні навколишніх територій. Вісн Дніпропетровського ун-ту. Біологія, медицина. 2015;6:133–40. DOI: [10.15421/021524](https://doi.org/10.15421/021524)
121. Цапкова ЛО, Чурносів МІ, Іщук МА. Аналіз алельних частот STR локусу DYS393 Y-хромосоми у популяціях українців. Медична генетика. 2005;4(6):285–86.
122. Шуба ВВ, Горпинченко МЮ, Атраментова ЛО. Вплив господарської діяльності на структуру популяцій на прикладі Луганської області. Фактори експериментальної еволюції організмів. 2023;33:114–17. DOI: [10.7124/FEEO.v33.1577](https://doi.org/10.7124/FEEO.v33.1577)

123. Шуба ВВ, Горпинченко МЮ, Атраментова ЛО. Популяційний аналіз населення Волинської області за даними прізвищ. Вісн Львів ун-ту. Серія біологічна. 2024;94:37–43. DOI: [10.30970/VLUBS.2025.94.04](https://doi.org/10.30970/VLUBS.2025.94.04)
124. Шуба ВВ, Атраментова ЛО. Популяційно-генетичний аналіз населення Запорізької області за даними прізвищ. Вісн Львів ун-ту. Серія біологічна. 2024;93:3–9. DOI: [10.30970/vlubs.2024.93.01](https://doi.org/10.30970/vlubs.2024.93.01)
125. Шуба ВВ, Горпинченко МЮ, Атраментова ЛО. Популяційно-генетичний аналіз населення Львівської області за даними прізвищ. Фактори експериментальної еволюції організмів. 2024;34:44–48. DOI: [10.7124/FEEO.v34.1615](https://doi.org/10.7124/FEEO.v34.1615)
126. Шуба ВВ, Горпинченко МЮ, Атраментова ЛО. Характеристика популяцій Київської області за даними прізвищ. Фактори експериментальної еволюції організмів. 2024;36:28–32. DOI: [10.7124/FEEO.v36.1708](https://doi.org/10.7124/FEEO.v36.1708)

СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА ЗА ТЕМОЮ ДИСЕРТАЦІЇ

Наукові праці, у яких опубліковані основні наукові результати дисертації (публікації у фахових виданнях України):

1. **Шуба В. В.**, Горпинченко М. Ю., Атраментова Л. О. (2023). Вплив господарської діяльності на структуру популяцій на прикладі Луганської області. *Фактори експериментальної еволюції організмів*, (33), 114–117. [DOI: 10.7124/FEEO.v33.1577](https://doi.org/10.7124/FEEO.v33.1577) (Особистий внесок здобувача: проведено пошук наукової літератури, виконано розрахунки, здійснено їх аналіз та узагальнення, написано текст; особистий внесок Атраментової Л. О.: концепція роботи, написання тексту; особистий внесок Горпинченка М. Ю.: підготовка наукової ілюстрації, написання та форматування тексту).
2. **Шуба В. В.**, Горпинченко М. Ю., Атраментова Л. О. (2024). Популяційно-генетичний аналіз населення Львівської області за даними прізвищ. *Фактори експериментальної еволюції організмів*, (34), 44–48. [DOI: 10.7124/FEEO.v34.1615](https://doi.org/10.7124/FEEO.v34.1615) (Особистий внесок здобувача: проведено пошук наукової літератури, виконано розрахунки, здійснено їх аналіз та узагальнення, написано текст; особистий внесок Атраментової Л. О.: концепція роботи, написання тексту; особистий внесок Горпинченка М. Ю.: підготовка наукової ілюстрації, написання та форматування тексту).
3. Атраментова Л. О., **Шуба В. В.**, Горпинченко М. Ю. (2024). Прізвища українського населення як інструмент популяційно-генетичних досліджень *Вісник Харківського національного університету імені В. Н. Каразіна. Серія «Біологія»*, (43), 44–49. [DOI: 10.26565/2075-5457-2024-43-4](https://doi.org/10.26565/2075-5457-2024-43-4) (Особистий внесок здобувача: проведено пошук наукової літератури, виконано частину розрахунків, написано текст; особистий внесок Атраментової Л. О.: концепція роботи, написання тексту; особистий

внесок Горпинченка М. Ю.: виконано частину розрахунків, підготовлено наукової ілюстрації, написано текст).

4. **Шуба В. В.**, Атраментова Л. О. (2024). Популяційно-генетичний аналіз населення Запорізької області за даними прізвищ. *Вісник Львівського університету. Серія біологічна*, (93), 3–9. [DOI: 10.30970/vlubs.2024.93.01](https://doi.org/10.30970/vlubs.2024.93.01) (*Особистий внесок здобувача: проведено пошук наукової літератури, виконано розрахунки, здійснено їх аналіз та узагальнення, написано текст; особистий внесок Атраментової Л. О.: концепція роботи, написання тексту*).
5. **Шуба В. В.**, Горпинченко М. Ю., Атраментова Л. О. (2025). Популяційний аналіз населення Волинської області за даними прізвищ. *Вісник Львівського університету. Серія біологічна*, (94), 37–43. [DOI: 10.30970/VLUBS.2025.94.04](https://doi.org/10.30970/VLUBS.2025.94.04) (*Особистий внесок здобувача: проведено пошук наукової літератури, виконано розрахунки, здійснено їх аналіз та узагальнення, написано текст; особистий внесок Атраментової Л. О.: концепція роботи, написання тексту; особистий внесок Горпинченка М. Ю.: підготовка наукової ілюстрації, написання та форматування тексту*).
6. **Шуба В. В.**, Горпинченко М. Ю., Атраментова Л. О. (2024). Характеристика популяцій Київської області за даними прізвищ. *Фактори експериментальної еволюції організмів*, (36), 28–32. [DOI: 10.7124/FEEEO.v36.1708](https://doi.org/10.7124/FEEEO.v36.1708) (*Особистий внесок здобувача: проведено пошук наукової літератури, виконано розрахунки, здійснено їх аналіз та узагальнення, написано текст; особистий внесок Атраментової Л. О.: концепція роботи, написання тексту; особистий внесок Горпинченка М. Ю.: підготовка наукової ілюстрації, написання і форматування тексту*).

Наукові праці, в яких опубліковані основні наукові результати дисертації у періодичних наукових виданнях що входять до міжнародних наукометричних баз Scopus або Web of Science:

1. **Shuba V.**, Yasinskyi Y., Atramentova L. (2025). The Structure of the Ukrainian Population Based on Surname Distribution. *American Journal of Human Biology*. (37,7), e70097. [DOI: 10.1002/ajhb.70097](https://doi.org/10.1002/ajhb.70097) (Особистий внесок здобувача: проведено пошук та аналіз наукової літератури, виконано розрахунки, здійснено їх аналіз та узагальнення, написано текст; особистий внесок Атраментової Л. О.: концепція роботи, написання тексту; особистий внесок Ясінського Я.С.: аналіз наукової літератури, виконано розрахунки, формальний аналіз).

Наукова праця, яка засвідчує апробацію матеріалів дисертації

1. **Шуба В. В.**, Горпинченко М. Ю., Атраментова Л. О. (2023). Вплив господарської діяльності на структуру популяцій на прикладі Луганської області. XVIII-а Міжнародна наукова конференція «Фактори експериментальної еволюції організмів», 26–29 вересня 2023 р., м. Київ, 2023.
2. **Шуба В. В.**, Горпинченко М. Ю., Атраментова Л. О. (2024). Популяційно-генетичний аналіз населення Львівської області за даними прізвищ. XIX-а Міжнародна наукова конференція «Фактори експериментальної еволюції організмів», 23–26 вересня 2024 р., м. Тернопіль, 2024.
3. **Шуба В. В.**, Горпинченко М. Ю., & Атраментова Л. О. (2025). Характеристика популяцій Київської області за даними прізвищ. // XX-а Міжнародна наукова конференція «Фактори експериментальної еволюції організмів», 22–26 вересня 2025 р., м. Київ, 2025.

Популяційні показники Запорізької області [124]

Популяції	N	S	$I^{jj} \times 10^{-4}$	$F \times 10^{-5}$	$v \times 10^{-4}$	$a \times 10^2$	H	R	I_{20}
Запорізька область	1748151	144325	2,53	6,34	22,56	39,53	14,39	30,62	19
місто Запоріжжя	879395	89101	2,65	6,64	42,83	37,83	14,11	28,54	13
Райони									
Бердянський	130845	22954	3,55	8,87	215,45	28,81	12,79	23,69	22
Василівський	69076	13529	6,79	16,98	212,99	15,03	12,18	24,21	92
Великобілозерський	9824	2333	19,02	47,56	534,15	5,54	10,10	23,83	633
Веселівський	23319	5944	5,81	14,52	737,93	18,58	11,64	19,75	560
Вільнянський	43298	10399	4,84	12,10	477,12	21,69	12,24	20,54	33
Гуляйпільський	36210	5777	12,71	31,79	216,93	8,03	10,99	27,43	390
Запорізький	46799	11477	5,35	13,38	399,18	19,46	12,28	20,84	40
Кам'янсько-Дніпровський	86577	19148	3,81	9,53	302,88	27,04	12,87	21,54	129
Куйбишевський	28738	5047	9,39	23,47	370,35	11,05	11,08	25,22	227
Мелітопольський	45430	11140	4,20	10,50	523,64	25,10	12,41	19,82	103
Михайлівський	30303	6874	6,83	17,07	483,01	15,38	11,61	21,99	262
Новомиколаївський	20327	4906	7,63	19,08	644,20	14,00	11,30	21,02	835
Оріхівський	53257	9324	6,75	16,87	278,14	15,24	11,79	24,90	208
Пологівський	45758	7363	9,91	24,76	220,41	10,31	11,28	27,15	1951
Приазовський	34393	6664	6,52	16,29	445,83	16,05	11,58	23,17	370
Приморський	38647	7580	6,79	16,98	380,81	15,30	11,68	23,35	563
Розівський	11976	3847	7,14	17,85	1168,63	15,85	11,21	17,23	1016
Токмацький	58637	11471	6,22	15,56	273,80	16,51	12,08	23,74	197
Чернігівський	23224	4501	12,47	31,18	344,81	8,29	10,82	25,41	1101
Якимівський	32117	8413	4,62	11,55	673,93	23,21	12,10	19,16	179

Примітки. N — кількість носіїв прізвищ, S — кількість прізвищ, I^{jj} — показник ізонімії, F — коефіцієнт інбридингу, v — індекс міграції, a — показник різноманітності прізвищ, H — показник ентропії розподілу прізвищ, R — показник надмірності розподілу прізвищ, I_{20} — індекс місця.

Популяційні показники Харківської області

Популяції	N	S	$I^{jj} \times 10^{-4}$	$F \times 10^{-5}$	$v \times 10^{-4}$	$a \times 10^2$	H	R	I_{20}
Харківська область	2818552	182765	2,53	6,34	22,56	39,53	14,39	30,62	18
місто Харків	1293863	93072	3,3	8,25	30,36	13,75	32,27	30,36	54
Райони									
Балаклійський	94073	11619	7,47	18,68	142,2	13,57	11,8	28,59	342
Барвінківський	35816	5583	9,13	22,83	305,47	11,29	11,19	26,02	418
Близнюківський	25714	4866	8,35	20,87	465,42	12,55	11,21	23,5	171
Богодухівський	55612	6894	12,32	30,8	145,8	8,23	11,03	30,02	75
Борівський	21172	3767	16,09	40,23	293,03	6,39	10,55	26,58	1059
Валківський	48429	6688	11,98	29,95	172,16	8,48	11,04	29,1	609
Великобурлуцький	29867	4351	17,73	44,33	188,49	5,74	10,51	29,3	585
Вовчанський	60279	7605	10,39	25,98	159,47	9,77	11,31	28,76	627
Дворічанський	25732	4075	14,15	35,38	274,24	7,26	10,68	27,11	4839
Дергачівський	109163	13293	8,52	21,29	107,47	11,86	11,76	29,76	199
Зачепилівський	20267	3530	13,77	34,43	357,82	7,52	10,6	25,89	1451
Зміївський	85806	10832	7,98	19,96	145,85	12,7	11,72	28,5	248
Золочівський	36787	5688	13,75	34,37	197,44	7,41	10,88	28,29	1174
Ізюмський	82615	10122	8,03	20,09	150,54	12,63	11,72	28,22	275
Кегичівський	23048	4594	9,73	24,32	445,6	10,75	11,13	23,19	299
Красноградський	51415	8095	8,31	20,77	233,94	12,32	11,63	25,68	496
Краснокутський	38759	5399	12,79	31,98	201,45	7,97	10,87	28,66	383
Куп'янський	91933	11975	8,81	22,02	123,39	11,48	11,82	28,31	309
Лозівський	107797	14380	5,25	13,12	176,67	19,39	12,37	25,98	37
Нововодолазький	44739	6313	12,97	32,43	172,09	7,83	11,01	28,74	614
Первомайський	54770	8966	6,77	16,93	269,42	15,16	11,80	25,01	575
Печенізький	13831	2154	25,56	63,90	282,19	4,02	9,74	29,21	663
Сахновщинський	28712	5409	8,19	20,47	425,07	12,75	11,33	23,50	241
Харківський	219612	25704	5,52	13,80	82,43	18,25	12,63	28,84	65
Чугуївський	92208	11898	6,74	16,84	160,91	15,08	11,92	27,73	1014
Шевченківський	26533	4298	11,93	29,83	315,54	8,65	10,83	26,30	523

Примітки. N — кількість носіїв прізвищ, S — кількість прізвищ, I^{jj} — показник ізонімії, F — коефіцієнт інбридингу, v — індекс міграції, a — показник різноманітності прізвищ, H — показник ентропії розподілу прізвищ, R — показник надмірності розподілу прізвищ, I_{20} — індекс місця.

Популяційні показники Закарпатської області

Популяції	N	S	$I^{jj} \times 10^{-4}$	$F \times 10^{-5}$	$v \times 10^{-4}$	$a \times 10^2$	H	R	I_{20}
Закарпатська область	1283620	64796	6,09	15,23	1,25	1,26	12,78	16,42	1238
місто Ужгород	115187	21306	3,58	8,94	242,59	28,64	13,01	22,60	61
Райони									
Берегівський	80726	8851	27,73	69,32	44,55	3,61	10,79	33,81	106
Великобerezнянський	29604	2615	21,64	54,10	155,75	4,68	9,73	34,52	566
Виноградівський	109474	8160	17,87	44,69	51,01	5,61	10,69	36,16	170
Воловецький	24878	2260	39,87	99,66	100,43	2,52	9,09	37,75	494
Іршавський	96093	5461	23,11	57,78	44,92	4,34	9,97	39,77	234
Міжгірський	53799	2931	30,02	75,05	61,73	3,34	9,35	40,51	199
Мукачівський	199214	18930	10,43	26,07	48,09	9,63	11,78	33,09	119
Перечинський	34394	3617	23,22	58,04	124,94	4,35	9,99	33,70	362
Рахівський	93175	6131	22,03	55,07	48,62	4,55	10,18	38,31	195
Свалявський	55559	5337	47,89	119,73	37,40	2,09	9,76	38,07	373
Тячівський	177557	8863	27,42	68,55	20,48	3,64	9,98	42,78	52
Ужгородський	79852	9425	16,90	42,25	73,97	5,95	11,24	30,97	169
Хустський	134108	8697	22,35	55,88	33,28	4,48	10,41	38,90	170

Примітки. N — кількість носіїв прізвищ, S — кількість прізвищ, I^{jj} — показник ізонімії, F — коефіцієнт інбридингу, v — індекс міграції, a — показник різноманітності прізвищ, H — показник ентропії розподілу прізвищ, R — показник надмірності розподілу прізвищ, I_{20} — індекс місця.

Популяційні показники Івано-Франківської області

Популяції	N	S	$I^{jj} \times 10$	$F \times 10^{-5}$	$v \times 10^{-4}$	$a \times 10^2$	H	R	I_{20}
Івано-Франківська область	1432000	66602	4,62	11,55	15,10	21,66	13,12	35,83	204
місто Івано-Франківськ	352358	38796	3,67	9,17	77,35	27,47	13,32	27,71	13
Райони									
Богородчанський	64128	4956	19,26	48,16	80,80	5,22	10,15	36,46	202
Верховинський	39457	2379	36,31	90,79	69,54	2,76	9,14	40,14	1893
Галицький	67429	8108	9,68	24,19	153,11	10,48	11,33	29,37	190
Городенківський	63538	5878	11,57	28,93	135,87	8,75	10,82	32,19	173
Долинський	95563	8994	11,38	28,44	91,87	8,86	11,14	32,66	164
Калуський	60113	4685	21,42	53,54	77,51	4,70	10,22	35,64	70
Коломийський	110646	7486	15,07	37,68	59,87	6,66	10,70	36,11	49
Косівський	89098	6399	17,72	44,31	63,21	5,67	10,35	37,03	1828
Надвірнянський	130963	10155	15,32	38,29	49,78	6,55	10,81	36,40	80
Рогатинський	54372	6047	10,01	25,03	183,49	10,16	11,04	29,81	493
Рожнятівський	80251	4748	22,04	55,09	56,43	4,55	10,04	38,37	165
Снятинський	75450	6585	15,91	39,78	83,17	6,33	10,62	34,47	146
Тисменицький	92542	8760	10,28	25,70	105,01	9,82	11,23	31,96	122
Тлумацький	56092	5294	14,37	35,91	123,92	7,04	10,66	32,41	169

Примітки. N — кількість носіїв прізвищ, S — кількість прізвищ, I^{jj} — показник ізонімії, F — коефіцієнт інбридингу, v — індекс міграції, a — показник різноманітності прізвищ, H — показник ентропії розподілу прізвищ, R — показник надмірності розподілу прізвищ, I_{20} — індекс місця.

Популяційні показники Львівської області [125]

Популяції	N	S	$I^{jj} \times 10^{-4}$	$F \times 10^{-5}$	$\nu \times 10^{-4}$	$a \times 10^2$	H	R	I_{20}
Львівська область	2675504	132189	2,30	5,75	16,25	43,55	14,14	33,79	97
місто Львів	705136	62225	2,27	5,69	62,33	44,23	13,90	28,43	13
Райони									
Бродівський	66274	7282	10,40	25,99	144,98	9,75	11,21	27,71	1003
Буський	53205	5618	10,84	27,09	173,26	9,38	11,00	29,91	427
Городоцький	80907	8870	7,46	18,65	165,53	13,62	11,56	29,10	424
Дрогобицький	248294	20967	4,69	11,71	85,91	21,52	12,43	30,64	103
Жидачівський	84905	8370	7,75	19,38	151,78	13,09	11,44	30,10	1020
Жовківський	107254	10789	8,23	20,59	113,13	12,27	11,57	30,74	316
Золочівський	80644	8449	7,29	18,23	169,89	13,94	11,54	29,18	506
Кам'янка-Бузький	63840	7198	10,87	27,17	143,97	9,33	11,18	29,96	575
Миколаївський	98110	9510	8,22	20,56	123,86	12,30	11,52	30,51	350
Мостиський	69304	6261	9,14	22,85	157,73	11,11	11,06	31,23	457
Перемишлянський	48016	5697	9,77	24,44	212,86	10,44	11,04	28,99	269
Пустомитівський	131194	16542	3,94	9,84	193,58	25,90	12,61	25,80	29
Радехівський	56872	5476	11,68	29,19	150,41	8,68	10,89	31,07	276
Самбірський	120622	12139	5,89	14,72	140,75	17,22	11,95	29,21	472
Сколівський	56914	4481	12,30	30,76	142,62	8,23	10,54	33,31	988
Сокальський	178803	18115	4,59	11,49	121,68	22,02	12,51	28,32	71
Старосамбірський	85945	7536	8,49	21,22	136,94	11,93	11,30	31,08	413
Стрийський	139328	14465	6,45	16,12	111,21	15,67	12,08	29,38	490
Турківський	53288	3672	26,45	66,13	70,76	3,80	10,01	36,24	870
Яворівський	146649	12752	7,16	17,91	95,10	14,08	11,72	31,70	209

Примітки. N — кількість носіїв прізвищ, S — кількість прізвищ, I^{jj} — показник ізонімії, F — коефіцієнт інбридингу, ν — індекс міграції, a — показник різноманітності прізвищ, H — показник ентропії розподілу прізвищ, R — показник надмірності розподілу прізвищ, I_{20} — індекс місця.

Популяційні показники Одеської області

Популяції	N	S	$I^{jj} \times 10^{-4}$	$F \times 10^{-5}$	$\nu \times 10^{-4}$	$a \times 10^2$	H	R	I_{20}
Одеська область	2357113	179591	2,10	5,24	20,71	47,82	14,78	30,08	84
місто Одеса	932488	102593	2,07	5,17	51,82	48,58	14,56	26,60	15
Райони									
Ананівський	38556	6746	9,42	23,56	275,01	10,90	11,31	25,75	204
Арцизький	51613	7295	18,61	46,52	103,94	5,42	10,85	30,68	629
Балтський	53558	9975	7,39	18,46	252,62	13,88	11,82	24,74	193
Березівський	39097	8651	5,05	12,62	506,38	20,85	12,05	21,03	368
Білгород-Дністровський	120824	20270	5,48	13,69	151,06	18,53	12,58	25,51	152
Біляївський	92478	20893	3,72	9,30	290,54	27,67	13,01	21,13	123
Болградський	77107	7762	16,55	41,38	78,23	6,08	10,65	34,41	987
Великомихайлівський	25944	6289	6,08	15,19	633,82	17,56	11,67	20,43	627
Іванівський	27301	6756	6,09	15,23	600,93	17,45	11,74	20,35	380
Ізмаїльський	124275	19290	5,37	13,43	149,73	18,89	12,45	26,46	311
м. Іллічівськ	53527	15175	2,84	7,10	657,24	37,66	13,00	17,26	32
Кілійський	65145	8320	10,40	26,01	147,38	9,74	11,26	29,62	431
Кодимський	39897	6483	12,09	30,23	207,05	8,44	11,08	27,52	441
Комінтернівський	74659	19876	2,71	6,78	494,07	38,80	13,22	18,33	21
Котовський	75179	13142	6,09	15,21	218,45	16,79	12,22	24,54	246
Красноокнянський	20747	4912	9,00	22,50	535,06	11,73	11,22	21,78	739
Любашівський	35171	6928	7,97	19,92	356,48	13,00	11,50	23,85	193
Миколаївський	24013	5595	6,98	17,44	596,54	15,23	11,49	21,05	304
Овідіопольський	101915	20642	3,07	7,69	319,00	33,58	13,11	21,17	48
Ренійський	35247	5209	15,18	37,95	186,60	6,70	10,73	28,96	1176
Роздільнянський	46350	11115	4,19	10,47	514,81	25,16	12,41	19,93	55
Савранський	25414	3742	17,57	43,92	223,58	5,81	10,37	29,12	813
Саратський	50776	6059	16,74	41,84	117,48	6,04	10,69	31,59	572
Тарутинський	38226	5290	16,76	41,91	155,81	6,05	10,70	29,74	792
Татарбунарський	41830	5355	13,94	34,84	171,31	7,29	10,72	30,15	224
Фрунзівський	19124	4270	9,33	23,32	559,97	11,34	11,06	22,24	478
Ширяївський	26652	5728	7,24	18,10	517,85	14,56	11,46	22,04	401

Примітки. N — кількість носіїв прізвищ, S — кількість прізвищ, I^{jj} — показник ізонімії, F — коефіцієнт інбридингу, ν — індекс міграції, a — показник різноманітності прізвищ, H — показник ентропії розподілу прізвищ, R — показник надмірності розподілу прізвищ, I_{20} — індекс місця. Дані повністю збігаються з роботами М.Ю. Горпинченка [101]

Популяційні показники Полтавської області

Популяції	N	S	$I^{jj} \times 10^{-4}$	$F \times 10^{-5}$	$v \times 10^{-4}$	$a \times 10^2$	H	R	I_{20}
Полтавська область	1548788	105340	4,09	10,23	15,77	24,46	13,58	33,97	23
місто Полтава	274343	36847	3,93	9,82	92,78	25,96	13,17	27,12	18
Район									
Великобагачанський	29592	4520	14,59	36,46	231,36	7,01	10,61	28,59	395
Гадяцький	43590	7805	9,74	24,36	235,25	10,50	11,22	27,23	317
Глобинський	56020	8938	8,21	20,53	217,23	12,44	11,56	26,68	213
Гребінківський	26161	4782	13,71	34,28	278,37	7,49	10,80	26,40	248
Диканський	21485	4313	12,36	30,89	376,19	8,40	10,81	24,88	559
Зіньківський	42992	5465	10,89	27,23	213,31	9,37	10,90	29,21	181
Карлівський	44564	7484	8,30	20,75	270,17	12,37	11,47	25,72	455
Кобеляцький	54895	8704	7,40	18,50	246,01	13,85	11,57	26,53	76
Козельщинський	24649	4220	13,48	33,70	300,55	7,64	10,67	26,85	583
Котелевський	22303	3748	18,12	45,29	247,06	5,65	10,33	28,51	558
Кременчуцький	302795	40203	4,04	10,10	81,75	24,96	13,26	27,17	24
Лохвицький	49777	7598	10,10	25,24	198,80	10,10	11,29	27,65	81
Лубенський	99486	13325	8,00	19,99	125,59	12,65	11,82	28,81	42
Машівський	20864	4370	10,82	27,05	442,47	9,66	10,94	23,79	164
Миргородський	81796	11503	9,01	22,52	135,62	11,25	11,63	28,75	804
Новосанжарський	39478	6128	11,77	29,44	214,88	8,67	10,96	28,24	290
Оржицький	29523	4252	16,83	42,07	200,93	6,05	10,51	29,25	522
Пирятинський	35146	6765	9,42	23,56	301,62	10,93	11,29	25,23	235
Полтавський	64101	11881	5,90	14,74	264,39	17,41	12,09	24,31	85
Решетилівський	29729	5277	11,08	27,71	303,14	9,29	11,02	25,85	280
Семенівський	31736	5025	11,11	27,77	283,39	9,26	10,92	26,97	81
Хорольський	41306	5778	12,90	32,25	187,43	7,89	10,85	29,25	159
Чорнухинський	16167	2399	21,18	52,96	291,40	4,85	9,86	29,48	325
Чутівський	25184	5168	10,26	25,66	386,51	10,13	11,12	23,92	131
Шишацький	22815	3954	12,86	32,14	340,47	8,04	10,63	26,57	519

Примітки. N — кількість носіїв прізвищ, S — кількість прізвищ, I^{jj} — показник ізонімії, F — коефіцієнт інбридингу, v — індекс міграції, a — показник різноманітності прізвищ, H — показник ентропії розподілу прізвищ, R — показник надмірності розподілу прізвищ, I_{20} — індекс місця. Дані частково збігаються з роботами М.Ю. Горпинченка [101]

Популяційні показники Житомирської області

Популяції	N	S	$I^{jj} \times 10^{-4}$	$F \times 10^{-5}$	$v \times 10^{-4}$	$a \times 10^2$	H	R	I_{20}
Житомирська область	1376812	88715	5,34	13,34	13,58	18,76	13,26	34,98	108
місто Житомир	257745	34537	4,46	11,16	86,88	22,59	13,11	27,09	17
Район									
Андрушівський	39278	6254	11,85	29,63	214,58	8,61	11,18	26,72	86
Баранівський	48209	5901	17,60	44,00	117,66	5,74	10,61	31,79	135
Бердичівський	121696	18369	6,82	17,05	120,40	14,83	12,39	26,65	27
Брусилівський	19346	3727	12,26	30,66	421,00	8,50	10,68	25,02	472
Володарсько-Волинський	44194	5717	18,15	45,37	124,45	5,57	10,58	31,41	682
Ємільчинський	41693	4610	28,33	70,83	84,42	3,55	9,93	35,31	499
Житомирський	60656	11363	8,28	20,71	198,83	12,31	11,91	25,03	34
Коростенський	91865	12152	15,55	38,87	69,90	6,47	11,24	31,80	227
Коростишівський	45970	6609	11,50	28,74	188,99	8,86	11,12	28,23	150
Лугинський	22758	2710	35,41	88,53	123,65	2,85	9,52	34,26	398
Любарський	31315	4072	13,93	34,83	228,93	7,34	10,59	29,10	685
Малинський	52807	6731	18,52	46,30	102,07	5,45	10,72	31,64	639
Народицький	11838	2030	26,67	66,68	315,91	3,86	9,65	28,70	329
Новоград-Волинський	112424	13989	12,27	30,67	72,42	8,20	11,57	31,04	38
Овруцький	67520	6611	25,91	64,77	57,02	3,87	10,30	35,82	191
Олевський	46740	4247	38,18	95,46	55,82	2,62	9,54	38,49	246
Попільнянський	36024	7058	9,11	22,78	281,00	11,28	11,40	25,25	109
Радомишльський	43177	5865	13,90	34,76	166,34	7,30	10,81	29,81	182
Романівський	32959	5150	12,47	31,17	243,04	8,21	10,90	27,39	123
Ружинський	36291	5164	13,49	33,71	204,06	7,56	10,80	28,68	159
Червоноармійський	28768	3589	29,57	73,93	117,20	3,41	9,95	32,80	161
Черняхівський	36739	4574	16,78	41,96	161,91	6,05	10,46	31,03	197
Чуднівський	46800	6574	13,58	33,94	157,17	7,47	11,02	28,97	98

Примітки. N — кількість носіїв прізвищ, S — кількість прізвищ, I^{jj} — показник ізонімії, F — коефіцієнт інбридингу, v — індекс міграції, a — показник різноманітності прізвищ, H — показник ентропії розподілу прізвищ, R — показник надмірності розподілу прізвищ, I_{20} — індекс місця.

Популяційні показники Київської області [126]

Популяції	N	S	$I^{jj} \times 10^{-4}$	$F \times 10^{-5}$	$v \times 10^{-4}$	$a \times 10^2$	H	R	I_{20}
Київська область	4480704	225045	3,51	8,77	6,33	28,51	14,28	35,41	14
місто Київ	2601732	136187	3,24	8,1	11,87	30,93	14,1	34,1	11
Райони									
Баришівський	68068	10373	8,27	20,66	177,60	12,31	11,66	27,38	228
Білоцерківський	265530	32719	5,52	13,81	68,15	18,22	12,84	28,72	21
Богуславський	47605	7107	13,14	32,85	159,66	7,72	11,05	28,92	384
Бориспільський	115145	18044	6,05	15,13	143,39	16,75	12,31	26,79	235
Бородянський	58657	10089	9,17	22,93	185,73	11,10	11,62	26,66	421
Броварський	165442	23260	6,66	16,65	90,70	15,14	12,43	28,30	150
Васильківський	113338	17267	8,07	20,17	109,28	12,52	12,08	28,03	27
Вишгородський	87182	14487	8,80	22,00	130,22	11,50	11,94	27,24	258
Володарський	25973	4955	11,54	28,85	333,28	8,95	11,01	24,94	369
Згурівський	32438	6883	9,49	23,72	324,64	10,88	11,35	24,28	1283
Іванківський	40893	5547	20,42	51,05	119,51	4,95	10,38	32,26	103
Кагарлицький	48406	7375	11,42	28,55	180,68	8,91	11,16	28,32	396
Києво-Святошинський	224961	35496	4,53	11,32	98,12	22,29	13,19	25,82	17
Макарівський	51265	8011	11,28	28,20	172,75	9,01	11,26	28,06	93
Миронівський	47200	7314	10,53	26,33	200,98	9,68	11,23	27,70	300
Обухівський	71281	12610	7,29	18,21	192,43	13,99	12,00	25,56	164
Переяслав-Хмельницький	68077	9341	10,94	27,36	134,08	9,25	11,30	29,60	232
Поліський	9239	1904	36,40	91,01	296,27	2,82	9,41	28,58	890
Рокитнянський	39061	5952	14,49	36,24	176,38	7,01	10,83	28,98	436
Сквирський	51079	7787	7,75	19,38	252,32	13,22	11,51	26,38	76
Ставищенський	29368	4427	15,13	37,81	224,79	6,75	10,58	28,70	485
Таращанський	39661	5059	15,70	39,26	160,31	6,46	10,67	30,18	643
Тетіївський	41642	6716	11,11	27,77	215,96	9,19	11,28	26,51	430
Фастівський	91410	15106	6,38	15,95	171,35	15,94	12,22	25,84	208
Яготинський	46051	7214	10,90	27,25	199,02	9,35	11,26	27,30	389

Примітки. N — кількість носіїв прізвищ, S — кількість прізвищ, I^{jj} — показник ізонімії, F — коефіцієнт інбридингу, v — індекс міграції, a — показник різноманітності прізвищ, H — показник ентропії розподілу прізвищ, R — показник надмірності розподілу прізвищ, I_{20} — індекс місця.

Популяційні показники Кіровоградської області

Популяції	N	S	$I^{jj} \times 10^{-4}$	$F \times 10^{-5}$	$v \times 10^{-4}$	$a \times 10^2$	H	R	I_{20}
Кіровоградська область	1152353	97390	3,59	8,97	24,18	27,94	13,96	30,65	23
місто Кіровоград	264926	37668	3,58	9,63	97,96	26,21	13,40	25,61	15
Райони									
Бобринецький	33716	7019	5,87	14,68	504,81	17,92	11,75	21,88	367
Вільшанський	17837	3512	13,49	33,73	415,03	7,72	10,62	24,80	608
Гайворонський	44593	7435	10,68	26,70	209,74	9,55	11,37	26,41	208
Голованівський	39211	7062	10,41	26,03	244,69	9,84	11,32	25,81	221
Добровеличківський	39926	8587	6,99	17,47	358,27	14,84	11,86	22,42	146
Долинський	40062	8873	5,39	13,47	463,16	19,46	12,03	21,31	70
Знам'янський	70793	13223	5,49	13,37	257,03	18,68	12,31	23,59	91
Кіровоградський	41732	9740	5,65	14,12	424,04	18,48	12,12	21,03	90
Компаніївський	19401	4557	8,40	21,00	613,26	12,68	11,22	21,22	429
Маловисківський	47282	9331	7,39	18,46	286,18	13,93	11,84	23,75	154
Новгородківський	22010	4408	9,25	23,12	490,76	11,36	11,08	23,17	193
Новоархангельський	32028	5668	8,99	22,47	347,13	11,52	11,26	24,77	294
Новомиргородський	38018	6856	8,78	21,96	299,20	11,73	11,45	24,72	288
Новоукраїнський	43685	8650	7,06	17,65	323,96	14,63	11,78	23,56	275
Олександрівський	36707	6085	11,08	27,71	245,50	9,24	11,10	26,83	56
Олександрійський	149468	22150	4,90	12,25	136,52	20,69	12,78	25,68	231
Онуфріївський	22453	4439	10,30	25,76	431,87	10,13	10,95	24,27	433
Петрівський	28288	6081	8,08	20,21	436,99	12,93	11,46	22,52	313
Світловодський	73814	13228	6,06	15,14	223,55	16,88	12,26	24,18	90
Ульяновський	30389	5326	11,57	28,93	284,07	8,89	10,98	26,24	325
Устинівський	16014	4055	8,34	20,85	748,33	12,95	11,14	20,24	616

Примітки. N — кількість носіїв прізвищ, S — кількість прізвищ, I^{jj} — показник ізонімії, F — коефіцієнт інбридингу, v — індекс міграції, a — показник різноманітності прізвищ, H — показник ентропії розподілу прізвищ, R — показник надмірності розподілу прізвищ, I_{20} — індекс місця.

Популяційні показники Миколаївської області

Популяції	N	S	$I^{jj} \times 10^{-4}$	$F \times 10^{-5}$	$\nu \times 10^{-4}$	$a \times 10^2$	H	R	I_{20}
Миколаївська область	1163805	122422	2,41	6,03	35,64	41,63	14,52	27,95	15
місто Миколаїв	438230	63050	2,60	6,50	87,73	38,79	11,04	25,10	28
Райони									
Арбузинський	26655	6403	8,42	21,05	445,30	12,42	11,50	21,75	275
Баштанський	41059	9144	5,44	13,59	447,85	19,25	12,03	21,47	96
Березанський	24496	6466	6,48	16,20	629,57	16,46	11,70	19,72	656
Березнегуватський	23116	5108	9,34	23,35	462,81	11,22	11,22	26,61	488
Братський	22049	5330	7,94	19,85	570,80	13,35	11,39	21,06	575
Веселинівський	25074	7178	4,67	11,68	853,46	23,40	12,00	17,92	129
Вознесенський	113333	23300	3,33	8,34	264,53	30,79	13,15	21,69	17
Врадіївський	20763	4226	11,01	27,54	436,79	9,48	10,92	23,83	660
Доманівський	27557	6435	6,06	15,14	598,69	17,55	11,68	20,79	312
Єланецький	17555	4406	7,41	18,52	768,37	14,61	11,25	20,24	624
Жовтневий	49962	12399	3,62	9,04	553,24	29,26	12,59	19,35	186
Казанківський	23765	5801	6,20	15,49	678,69	17,30	11,56	20,48	1093
Кривоозерський	29210	5689	8,53	21,32	401,08	12,21	11,32	23,71	313
Миколаївський	37930	10399	4,37	10,93	602,91	24,34	12,38	18,62	237
Новобузький	32725	8120	5,53	13,82	552,64	19,14	11,95	20,31	189
Новоодеський	35962	8190	5,76	14,41	482,12	18,22	11,94	21,13	108
Очаківський	32490	9141	4,99	12,48	616,30	21,34	12,17	18,78	159
Первомайський	100594	19688	4,12	10,30	241,13	24,86	12,84	22,72	34
Снігурівський	41280	10052	4,57	11,42	530,16	23,11	12,28	19,91	304

Примітки. N — кількість носіїв прізвищ, S — кількість прізвищ, I^{jj} — показник ізонімії, F — коефіцієнт інбридингу, ν — індекс міграції, a — показник різноманітності прізвищ, H — показник ентропії розподілу прізвищ, R — показник надмірності розподілу прізвищ, I_{20} — індекс місця.

Популяційні показники Рівненської області

Популяції	N	S	$I^{jj} \times 10^{-4}$	$F \times 10^{-5}$	$v \times 10^{-4}$	$a \times 10^2$	H	R	I_{20}
Рівненська область	1089343	60456	6,77	16,91	13,12	14,79	12,82	36,23	178
місто Рівне	211380	24083	6,32	15,80	74,81	15,93	12,68	28,32	15
Райони									
Березнівський	56381	3956	28,94	72,35	61,11	3,47	9,66	38,81	189
Володимирецький	77108	6398	18,03	45,07	71,81	5,58	10,39	36,01	154
Гоцанський	38800	3892	19,12	47,79	134,57	5,29	10,22	32,97	48
Демидівський	17893	2689	21,09	52,73	264,42	4,86	10,04	28,95	257
Дубенський	55029	8141	11,51	28,78	157,70	8,82	11,38	27,77	29
Дубровицький	56412	4406	27,55	68,87	64,17	3,64	9,81	37,83	262
Зарічненський	31048	1920	98,48	246,19	32,39	1,01	8,72	41,59	411
Здолбунівський	58535	7241	12,90	32,24	132,31	7,85	11,14	29,68	22
Корецький	37478	3296	19,58	48,96	135,98	5,17	10,03	34,00	113
Костопільський	76338	5977	17,81	44,52	73,43	5,65	10,52	35,12	124
Млинівський	40399	5043	13,49	33,73	183,22	7,54	10,87	28,95	161
Острозький	40332	5120	16,68	41,70	148,40	6,08	10,59	30,81	57
Радивилівський	39488	4148	16,97	42,42	148,99	5,97	10,41	31,84	113
Рівненський	89397	9505	9,66	24,14	115,72	10,47	11,55	29,77	92
Рокитнівський	44996	2690	62,17	155,43	35,53	1,60	8,68	43,88	206
Сарненський	118329	6723	18,62	46,54	45,31	5,39	10,36	38,54	414

Примітки. N — кількість носіїв прізвищ, S — кількість прізвищ, I^{jj} — показник ізонімії, F — коефіцієнт інбридингу, v — індекс міграції, a — показник різноманітності прізвищ, H — показник ентропії розподілу прізвищ, R — показник надмірності розподілу прізвищ, I_{20} — індекс місця.

Популяційні показники Сумської області

Популяції	N	S	$I^{jj} \times 10^{-4}$	$F \times 10^{-5}$	$\nu \times 10^{-4}$	$a \times 10^2$	H	R	I_{20}
Сумська область	1007814	83304	4,47	11,17	17,55	22,41	13,42	33,82	23
місто Суми	171471	26800	5,45	13,62	106,98	18,54	12,77	26,56	13
Райони									
Білопільський	56312	7106	11,38	28,46	155,83	8,91	11,04	30,03	123
Буринський	35982	4562	12,70	31,76	218,47	8,04	10,67	29,48	117
Великописарівський	30228	3688	16,85	42,12	196,05	6,04	10,27	30,96	423
Глухівський	48637	8024	7,84	19,61	261,93	13,08	11,47	26,31	162
Конотопський	92325	13857	8,27	20,67	130,91	12,25	11,83	28,28	62
Краснопільський	31320	5193	11,22	28,06	284,14	9,16	10,94	26,75	170
Кролевецький	47355	6082	11,87	29,67	177,72	8,57	10,90	29,85	219
Лебединський	44039	7423	10,36	25,91	218,90	9,86	11,24	27,12	124
Липоводолинський	25821	2997	19,66	49,16	196,59	5,18	10,05	31,41	168
Недригайлівський	32990	3619	19,26	48,16	157,06	5,26	10,18	32,17	236
Охтирський	58402	9039	9,58	23,95	178,58	10,62	11,40	27,99	171
Путівльський	33536	6174	7,38	18,45	403,80	14,11	11,43	24,00	601
Роменський	69547	9644	8,48	21,20	169,44	11,99	11,51	28,44	54
Середино-Бурдський	20807	4097	12,87	32,18	372,86	8,06	10,76	24,98	849
Сумський	63415	9269	9,08	22,70	173,48	11,20	11,52	27,77	31
Тростянецький	39754	6418	11,15	27,87	225,37	9,17	11,10	27,34	167
Шосткинський	77296	12807	6,82	17,05	189,62	14,94	11,96	26,36	217
Ямпільський	28577	5372	9,77	24,42	357,95	10,61	11,09	25,09	532

Примітки. N — кількість носіїв прізвищ, S — кількість прізвищ, I^{jj} — показник ізонімії, F — коефіцієнт інбридингу, ν — індекс міграції, a — показник різноманітності прізвищ, H — показник ентропії розподілу прізвищ, R — показник надмірності розподілу прізвищ, I_{20} — індекс місця.

Популяційні показники Тернопільської області

Популяції	N	S	$I^{jj} \times 10^{-4}$	$F \times 10^{-5}$	$\nu \times 10^{-4}$	$a \times 10^2$	H	R	I_{20}
Тернопільська область	1119051	63089	3,20	7,99	27,95	31,37	13,57	32,46	73
місто Тернопіль	234818	29697	3,02	7,55	140,99	33,58	13,31	25,38	12
Райони									
Бережанський	41461	5558	9,63	24,06	250,35	10,65	11,11	27,54	238
Борщівський	69544	8922	6,72	16,81	213,73	15,19	11,70	27,24	127
Бучацький	65274	6273	9,20	22,99	166,43	11,05	11,11	30,56	141
Гусятинський	73847	8667	7,08	17,70	191,08	14,39	11,65	27,99	2296
Заліщицький	52682	6633	9,12	22,81	207,90	11,19	11,24	28,32	140
Збаразький	62980	6869	7,87	19,68	201,52	12,95	11,40	28,47	98
Зборівський	44964	5677	9,55	23,88	232,57	10,71	11,13	27,97	129
Козівський	37331	5046	9,98	24,95	268,18	10,29	11,05	27,23	265
Кременецький	58771	7766	11,58	28,94	146,79	8,76	11,17	29,51	132
Лановецький	31671	3962	16,44	41,10	191,74	6,19	10,46	30,02	165
Монастирський	33502	4198	12,28	30,69	242,85	8,34	10,73	28,60	698
Підволочиський	50684	6433	8,84	22,09	223,11	11,57	11,29	27,76	130
Підгаєцький	20967	3052	11,73	29,33	406,10	8,88	10,54	26,57	897
Теребовлянський	66292	8367	5,99	14,97	251,82	17,12	11,76	26,60	96
Тернопільський	61224	9324	4,98	12,44	328,11	20,77	12,03	24,36	189
Чортківський	78133	10511	5,97	14,93	214,15	17,10	11,94	26,51	122
Шумський	34906	3744	21,44	53,61	133,31	4,72	10,16	32,64	143

Примітки. N — кількість носіїв прізвищ, S — кількість прізвищ, I^{jj} — показник ізонімії, F — коефіцієнт інбридингу, ν — індекс міграції, a — показник різноманітності прізвищ, H — показник ентропії розподілу прізвищ, R — показник надмірності розподілу прізвищ, I_{20} — індекс місця.

Популяційні показники Херсонської області

Популяції	N	S	$I^{jj} \times 10^{-4}$	$F \times 10^{-5}$	$\nu \times 10^{-4}$	$a \times 10^2$	H	R	I_{20}
Херсонська область	1144565	117012	2,57	6,41	34,05	39,11	14,37	28,62	36
місто Херсон	304026	50850	2,76	6,90	119,13	36,66	13,85	23,95	18
Райони									
Бериславський	52791	11882	4,41	11,02	429,68	23,70	12,42	20,82	154
Білозерський	90888	18691	3,84	9,61	286,14	26,77	12,87	21,85	90
Великопетиський	21245	4665	8,61	21,53	546,03	12,27	11,16	22,38	718
Великоолександрівський	33905	6975	6,34	15,86	464,75	16,53	11,69	22,35	173
Верхньорогачицький	15378	3380	12,36	30,90	525,49	8,53	10,65	23,44	428
Високопільський	19795	4606	7,61	19,03	663,02	14,06	11,27	21,02	722
Генічеський	60589	13681	3,29	8,22	501,63	32,00	12,69	20,13	138
Голопристанський	60360	11837	6,21	15,52	266,65	16,54	12,12	23,69	92
Горностаївський	24247	5150	6,94	17,35	593,87	15,31	11,36	22,00	418
Іванівський	18039	4992	8,68	21,70	638,18	12,30	11,20	20,78	581
Каланчацький	27488	6396	10,41	26,03	349,11	9,94	11,48	22,15	230
Каховський	150651	26570	3,38	8,45	196,23	30,15	13,21	23,18	42
Нижньосірогозький	21208	4890	7,86	19,65	599,39	13,52	11,30	21,38	807
Нововоронцовський	29040	5854	7,57	18,93	454,42	13,82	11,44	22,84	593
Новотроїцький	45348	9658	4,96	12,41	443,96	21,07	12,16	21,41	129
Скадовський	53362	10634	7,15	17,88	261,89	14,35	12,13	22,78	130
Цюрупинський	72615	13957	6,54	16,36	210,29	15,60	12,22	24,30	90
Чаплинський	43590	9813	6,26	15,64	366,52	16,58	12,41	21,24	232

Примітки. N — кількість носіїв прізвищ, S — кількість прізвищ, I^{jj} — показник ізонімії, F — коефіцієнт інбридингу, ν — індекс міграції, a — показник різноманітності прізвищ, H — показник ентропії розподілу прізвищ, R — показник надмірності розподілу прізвищ, I_{20} — індекс місця.

Популяційні показники Хмельницької області

Популяції	N	S	$I^{jj} \times 10^{-4}$	$F \times 10^{-5}$	$\nu \times 10^{-4}$	$a \times 10^2$	H	R	I_{20}
Хмельницька область	1524052	98551	5,05	12,62	12,99	19,83	13,67	33,43	44
місто Хмельницький	235004	34309	4,49	11,23	94,72	22,47	13,25	25,75	13
Район									
Білогірський	32913	4023	16,31	40,77	186,01	6,24	10,44	30,45	210
Віньковецький	34340	4462	14,27	35,67	203,80	7,14	10,63	29,45	157
Волочиський	67812	8706	8,46	21,16	174,06	12,01	11,55	28,04	89
Городоцький	64084	8308	7,16	17,90	217,73	14,26	11,61	27,28	88
Деражнянський	42511	7385	9,67	24,19	242,91	10,58	11,49	25,25	541
Дунаєвецький	84038	10511	8,18	20,44	145,44	12,40	11,75	28,16	62
Ізяславський	95963	13572	10,75	26,89	96,79	9,38	11,72	29,19	26
Кам'янець-Подільський	194772	25912	5,73	14,33	89,52	17,59	12,87	26,76	15
Красилівський	63016	7637	13,36	33,41	118,59	7,56	11,05	30,71	80
Летичівський	41955	6144	10,08	25,21	236,16	10,15	11,19	27,11	169
Новоушицький	41728	5623	12,69	31,73	188,58	8,02	10,94	28,73	162
Полонський	51626	6521	11,69	29,22	165,54	8,69	11,02	29,64	251
Славутський	76422	9971	17,53	43,82	74,52	5,74	11,07	31,74	84
Старокостянтинівський	74590	10417	10,93	27,33	122,50	9,25	11,51	28,91	23
Старосинявський	24095	3608	14,11	35,28	293,69	7,29	10,49	27,97	433
Теофіпольський	33754	4192	16,62	41,56	177,94	6,12	10,44	30,58	67
Хмельницький	59006	9908	8,17	20,44	207,16	12,48	11,74	25,93	58
Чемеровецький	60350	7689	9,63	24,08	171,85	10,55	11,36	28,47	164
Шепетівський	95510	12238	10,42	26,06	100,36	9,68	11,68	29,42	46
Ярмолинецький	50563	7804	10,30	25,75	191,81	9,89	11,41	26,95	106

Примітки. N — кількість носіїв прізвищ, S — кількість прізвищ, I^{jj} — показник ізонімії, F — коефіцієнт інбридингу, ν — індекс міграції, a — показник різноманітності прізвищ, H — показник ентропії розподілу прізвищ, R — показник надмірності розподілу прізвищ, I_{20} — індекс місця.

Популяційні показники Черкаської області

Популяції	N	S	$I^{jj} \times 10^{-4}$	$F \times 10^{-5}$	$v \times 10^{-4}$	$a \times 10^2$	H	R	I_{20}
Черкаська область	1494074	99066	5,08	12,71	13,16	19,69	13,48	34,29	20
місто Черкаси	308932	42149	4,16	10,40	77,75	24,21	13,37	26,67	19
Райони									
Городищенський	56607	7076	14,11	35,28	125,00	7,17	10,90	30,96	83
Драбівський	46311	5373	13,40	33,49	160,98	7,58	10,69	31,04	233
Жашківський	50853	6628	11,88	29,71	165,30	8,55	11,05	29,32	106
Звенигородський	96456	12296	9,13	22,82	113,50	11,07	11,76	28,95	65
Золотоніський	81366	10896	8,44	21,09	145,55	12,02	11,59	28,95	106
Кам'янський	40942	6014	14,37	35,92	169,77	7,07	10,83	29,35	86
Канівський	57530	9200	9,26	23,14	187,59	11,00	11,55	26,98	72
Катеринопільський	32474	5055	16,12	40,29	190,77	6,32	10,78	28,10	122
Корсунь-Шевченківський	53928	6946	14,79	36,97	125,20	6,84	10,84	31,02	108
Лисянський	32019	4563	14,21	35,52	219,54	7,19	10,67	28,68	219
Маньківський	35726	5273	12,33	30,83	226,74	8,29	10,90	27,93	159
Монастирищенський	42258	5623	12,48	31,20	189,36	8,16	10,94	28,81	343
Смілянський	111871	16484	8,29	20,73	107,69	12,18	12,11	27,80	25
Тальнівський	46731	7019	10,80	26,99	197,99	9,44	11,23	27,59	61
Уманський	131466	18929	7,37	18,41	103,19	13,71	12,34	27,42	39
Христинівський	39920	6599	10,70	26,76	233,77	9,56	11,22	26,62	153
Черкаський	79224	11016	11,53	28,83	109,32	8,76	11,33	30,37	92
Чигиринський	39342	5733	13,31	33,27	190,76	7,65	10,83	29,06	90
Чорнобаївський	55365	6454	11,28	28,20	159,92	9,00	10,93	30,61	140
Шполянський	54753	7896	11,85	29,61	154,00	8,56	11,21	28,77	174

Примітки. N — кількість носіїв прізвищ, S — кількість прізвищ, I^{jj} — показник ізонімії, F — коефіцієнт інбридингу, v — індекс міграції, a — показник різноманітності прізвищ, H — показник ентропії розподілу прізвищ, R — показник надмірності розподілу прізвищ, I_{20} — індекс місця.

Популяційні показники Чернігівської області

Популяції	N	S	$I^{jj} \times 10^{-4}$	$F \times 10^{-5}$	$v \times 10^{-4}$	$a \times 10^2$	H	R	I_{20}
Чернігівська область	1074996	76610	4,29	10,72	26,80	23,38	13,25	32,82	20
місто Чернігів	62239	8586	8,87	22,17	181,03	11,40	11,40	28,39	14
Райони									
Бахмацький	59703	7267	11,52	28,20	145,25	8,80	11,05	30,36	201
Бобровицький	42236	5998	10,39	25,98	227,62	9,84	10,97	28,58	235
Борзнянський	50899	5387	11,87	29,67	165,35	8,65	10,77	31,15	78
Варвинський	21618	3244	17,18	42,96	268,73	5,97	10,26	28,73	230
Городнянський	37787	5480	10,88	27,19	243,04	9,41	10,95	28,01	186
Ічнянський	44025	5243	12,55	31,39	180,70	8,10	10,77	30,16	63
Козелецький	61702	7677	11,07	27,67	146,29	9,16	11,07	30,46	207
Коропський	33930	4180	13,51	33,77	217,88	7,56	10,53	30,02	424
Корюківський	35216	4161	17,25	43,14	164,30	5,88	10,31	31,72	211
Куликівський	22523	3115	18,06	45,14	245,47	5,67	10,11	30,08	429
Менський	53132	6041	10,84	27,11	173,37	9,37	10,98	30,08	88
Ніжинський	120626	12793	8,30	20,75	99,80	12,16	11,66	30,91	72
Новгород-Сіверський	39084	5602	9,82	24,54	260,41	10,45	11,02	27,78	355
Носівський	40896	4986	12,86	32,15	189,91	7,92	10,70	30,18	119
Прилуцький	111918	14204	7,99	19,97	111,75	12,65	11,88	29,19	126
Ріпкинський	57441	9705	8,58	21,44	202,84	11,89	11,63	26,42	117
Семенівський	24005	3398	16,62	41,54	250,30	6,16	10,33	29,02	555
Сосницький	30000	3135	17,32	43,29	192,18	5,88	10,14	31,79	167
Срібнянський	14071	2215	22,38	55,96	316,81	4,60	9,83	28,66	365
Талалаївський	17345	2605	19,06	47,65	301,90	5,40	10,05	28,64	657
Чернігівський	62239	8586	8,87	22,17	181,03	11,47	11,40	28,39	153
Щорський	32361	4441	16,20	40,51	190,40	6,28	10,52	29,77	279

Примітки. N — кількість носіїв прізвищ, S — кількість прізвищ, I^{jj} — показник ізонімії, F — коефіцієнт інбридингу, v — індекс міграції, a — показник різноманітності прізвищ, H — показник ентропії розподілу прізвищ, R — показник надмірності розподілу прізвищ, I_{20} — індекс місця. Дані частково збігаються з роботами М.Ю. Горпинченка [101]

Популяційні показники Донецької області

Популяції	N	S	$I^{jj} \times 10^{-4}$	$F \times 10^{-5}$	$v \times 10^{-4}$	$a \times 10^2$	H	R	I_{20}
Донецька область	4575766	282661	2,04	5,11	10,93	48,97	14,99	32,17	38
місто Донецьк-Макіївка	1279979	97128	2,81	7,02	27,83	35,72	14,05	30,74	86
Райони									
Амвросіївський	52873	8729	7,93	19,84	238,19	12,90	11,69	25,52	433
Бахмутський	537590	53668	3,28	8,20	56,66	30,63	13,58	28,68	44
Великоновосілівський	54628	7471	7,82	19,56	233,84	13,08	11,50	26,95	1508
Волноваський	118518	15319	4,72	11,79	178,81	21,58	12,41	26,36	189
Нікольський	29479	6001	7,30	18,25	464,41	14,36	11,44	22,93	2289
Добропільський	102437	15975	4,07	10,17	239,99	25,19	12,64	24,09	165
Костянтинівський	193934	25292	3,79	9,49	135,84	26,71	13,01	25,95	73
Покровський	260371	30370	3,54	8,85	108,43	28,54	13,11	27,11	130
Лиманський	59555	7709	10,29	25,72	163,06	9,87	11,24	29,11	894
Мар'їнський	114990	15420	4,93	12,33	176,28	20,63	12,40	26,26	51
Новоазовський	52472	9268	6,00	14,99	317,69	17,22	11,89	24,20	421
Олександрівський	25934	4691	9,03	22,57	426,79	11,56	11,18	23,75	723
Мангушський	442711	48025	2,88	7,19	78,53	35,04	13,62	27,38	50
Слов'янський	437456	41133	4,40	11,01	51,90	22,82	13,16	29,79	61
Старобешівський	60103	10040	9,00	22,51	184,64	11,31	11,81	25,60	931
Бойківський	39213	6504	7,67	19,18	332,23	13,48	11,50	24,64	1532
Шахтарський	389507	41986	3,55	8,87	72,30	28,37	13,43	27,68	325
Ясинуватський	224016	30585	3,32	8,30	134,40	30,52	13,24	25,52	79

Примітки. N — кількість носіїв прізвищ, S — кількість прізвищ, I^{jj} — показник ізонімії, F — коефіцієнт інбридингу, v — індекс міграції, a — показник різноманітності прізвищ, H — показник ентропії розподілу прізвищ, R — показник надмірності розподілу прізвищ, I_{20} — індекс місця.

Популяційні показники Луганської області [122]

Популяції	N	S	$I^{jj} \times 10^{-4}$	$F \times 10^{-5}$	$v \times 10^{-4}$	$a \times 10^2$	H	R	I_{20}
Луганська область	2229101	173457	2,40	6,01	18,65	41,65	14,54	31,02	43
місто Луганськ	375608	41897	3,84	9,60	69,33	26,22	13,30	28,16	14
Райони									
Антрацитівський	259329	30771	3,89	9,72	99,19	25,98	13,11	27,13	16
Біловодський	25905	3697	16,70	41,74	230,84	6,12	10,42	28,92	435
Білокуракинський	23284	3398	19,99	49,97	214,46	5,10	10,21	29,59	586
Довжанський	119819	17726	5,02	12,55	166,21	20,25	12,61	25,26	86
Сорокинський	118676	19011	4,20	10,50	200,47	24,28	12,80	24,09	149
Кремінський	118451	14735	6,42	16,04	131,49	15,78	12,12	28,10	245
Лутугинський	144073	20584	4,71	11,78	147,20	21,52	12,72	25,77	67
Марківський	17826	2571	23,74	59,34	235,78	4,30	9,88	30,02	640
Міловський	13325	2509	23,19	57,98	322,88	4,45	9,96	27,31	661
Новоайдарський	139516	19231	5,39	13,48	132,88	18,79	12,52	26,73	163
Новоопсковський	34098	4268	16,41	41,04	178,38	6,19	10,51	30,23	378
Перевальський	181243	24927	3,67	9,18	150,18	27,63	13,05	25,30	18
Попаснянський	386714	41614	3,47	8,68	74,46	29,01	13,36	28,02	17
Сватівський	43610	6336	9,83	24,57	233,14	10,41	11,20	27,33	117
Слов'яносербський	97428	15626	4,39	10,99	233,45	23,29	12,53	24,37	67
Станично-Луганський	50375	8881	8,20	20,50	241,87	12,49	11,70	25,07	510
Старобільський	54925	7714	11,66	29,16	155,93	8,70	11,29	28,28	252
Троїцький	24896	3777	13,35	33,38	300,41	7,71	10,59	27,46	994

Примітки. N — кількість носіїв прізвищ, S — кількість прізвищ, I^{jj} — показник ізонімії, F — коефіцієнт інбридингу, v — індекс міграції, a — показник різноманітності прізвищ, H — показник ентропії розподілу прізвищ, R — показник надмірності розподілу прізвищ, I_{20} — індекс місця.

Популяційні показники Чернівецької області

Популяції	N	S	$I^{jj} \times 10^{-4}$	$F \times 10^{-5}$	$v \times 10^{-4}$	$a \times 10^2$	H	R	I_{20}
Чернівецька область	870210	62819	4,57	11,43	17,37	21,89	13,24	34,67	342
місто Чернівці	234598	31215	3,36	8,40	126,82	30,13	13,26	25,65	29
Райони									
Вижницький	51294	4904	23,75	59,37	81,89	4,24	10,11	35,40	78
Герцаївський	30203	3125	19,59	48,98	168,67	5,18	10,02	32,64	371
Глибоцький	60679	5807	15,44	38,60	106,58	6,54	10,30	35,15	141
Заставнівський	62156	5980	10,71	26,78	150,03	9,47	10,90	31,54	146
Кельменецький	50816	4766	18,03	45,07	108,97	5,60	10,40	33,50	146
Кіцманський	72451	6962	14,57	36,42	94,60	6,92	10,83	32,91	70
Новоселицький	78981	6925	11,73	29,34	107,78	8,60	10,85	33,31	233
Путильський	23888	2125	36,23	90,57	115,14	2,78	9,23	36,51	255
Сокирянський	55195	5905	17,11	42,79	105,68	5,90	10,62	32,56	91
Сторожинецький	76171	7291	16,17	40,43	81,05	6,22	10,76	33,68	117
Хотинський	73778	6635	14,95	37,38	90,51	6,74	10,70	33,81	114

Примітки. N — кількість носіїв прізвищ, S — кількість прізвищ, I^{jj} — показник ізонімії, F — коефіцієнт інбридингу, v — індекс міграції, a — показник різноманітності прізвищ, H — показник ентропії розподілу прізвищ, R — показник надмірності розподілу прізвищ, I_{20} — індекс місця.

Популяційні показники Вінницької області

Популяції	N	S	$I^{jj} \times 10^{-4}$	$F \times 10^{-5}$	$v \times 10^{-4}$	$a \times 10^2$	H	R	I_{20}
Вінницька область	1905140	112371	5,32	13,30	9,86	18,81	13,68	34,42	59
місто Вінниця	402505	41521	4,74	11,85	52,40	21,20	13,24	28,86	11
Райони									
Барський	73119	8717	9,40	23,50	145,36	10,79	11,47	29,01	102
Бершадський	73509	7551	10,53	26,32	129,08	9,61	11,15	31,00	59
Вінницький	70948	9836	9,69	24,23	145,31	10,46	11,58	28,16	47
Гайсинський	70402	8159	9,91	24,78	143,16	10,23	11,37	29,39	88
Жмеринський	46239	7010	10,60	26,51	203,73	9,62	11,33	26,87	204
Іллінецький	51800	6279	11,64	29,11	165,60	8,72	11,01	29,67	90
Калинівський	70653	8157	11,69	29,23	120,90	8,65	11,24	30,21	121
Козятинський	79926	9483	11,39	28,47	109,75	8,87	11,45	29,68	386
Крижопільський	44216	5366	13,50	33,75	167,29	7,52	10,83	29,83	200
Липовецький	55568	6517	12,04	30,09	149,35	8,43	11,03	30,01	86
Літинський	45416	5369	14,55	36,37	151,13	6,97	10,73	30,67	103
Могилів-Подільський	77431	9591	9,61	24,03	134,25	10,54	11,60	28,56	91
Мурованокуриловецький	40300	4477	18,44	46,10	134,33	5,49	10,63	30,52	162
Немирівський	63236	8177	10,49	26,23	150,57	9,67	11,28	29,30	268
Оратівський	28320	3898	13,22	33,05	266,79	7,76	10,62	28,20	624
Піщанський	28718	3696	14,10	35,25	246,60	7,26	10,49	29,14	485
Погребищенський	44668	5475	13,56	33,89	164,91	7,49	10,84	29,83	99
Теплицький	36710	4322	13,22	33,05	205,79	7,71	10,62	29,97	123
Тиврівський	52831	7655	8,96	22,41	210,98	11,39	11,52	26,59	35
Томашпільський	43012	5908	12,23	30,59	189,80	8,32	11,03	28,34	201
Тростянецький	63627	9204	9,80	24,51	160,18	10,36	11,55	27,62	19
Тульчинський	77243	9592	10,36	25,91	124,78	9,76	11,52	29,06	29
Хмільницький	77656	8145	12,31	30,79	104,44	8,20	11,16	31,30	22
Чернівецький	34205	3507	16,26	40,65	179,51	6,25	10,27	31,80	496
Чечельницький	30243	3746	16,21	40,54	203,61	6,29	10,42	30,02	371
Шаргородський	73625	7459	11,89	29,72	114,10	8,50	11,08	31,45	280
Ямпільський	49014	5392	13,10	32,75	155,53	7,74	10,80	30,66	590

Примітки. N — кількість носіїв прізвищ, S — кількість прізвищ, I^{jj} — показник ізонімії, F — коефіцієнт інбридингу, v — індекс міграції, a — показник різноманітності прізвищ, H — показник ентропії розподілу прізвищ, R — показник надмірності розподілу прізвищ, I_{20} — індекс місця.

Популяційні показники Волинської області [123]

<i>Популяції</i>	<i>N</i>	<i>S</i>	$I^{jj} \times 10^{-4}$	$F \times 10^{-5}$	$\nu \times 10^{-4}$	$a \times 10^2$	<i>H</i>	<i>R</i>	<i>I</i> ₂₀
Волинська область	1050021	59545	7,14	17,85	13,33	14,02	12,80	36,00	111
місто Луцьк	201938	23348	6,33	15,84	78,13	15,90	12,68	28,04	12
Райони									
Володимир-Волинський	59852	8660	9,43	23,58	177,00	10,78	11,55	27,24	27
Горохівський	62153	6350	13,06	32,65	123,02	7,74	11,02	30,78	31
Іваничівський	83071	10578	8,93	22,33	134,65	11,34	11,76	28,06	21
Камінь-Каширський	63540	3021	26,77	66,92	58,64	3,75	9,52	40,35	80
Ківерцівський	73831	7434	16,20	40,50	83,48	6,22	10,88	32,69	100
Ковельський	110627	10345	10,72	26,81	84,20	9,39	11,38	32,08	27
Локачинський	24580	3138	17,93	44,83	226,49	5,70	10,29	29,48	65
Луцький	50680	7362	11,27	28,17	174,89	9,02	11,33	27,52	47
Любешівський	42656	2305	33,54	83,86	69,65	2,99	9,07	41,03	242
Любомльський	44309	3886	20,90	52,25	107,76	4,83	10,06	34,80	175
Маневицький	63087	3845	22,52	56,30	70,23	4,46	9,93	37,70	49
Ратнівський	50867	2505	35,64	89,11	54,96	2,81	9,10	41,77	111
Рожищенський	41473	5105	18,36	45,90	131,08	5,51	10,55	31,19	53
Старовижівський	31760	2557	23,95	59,87	131,16	4,22	9,65	35,47	118
Турійський	27563	4097	15,55	38,88	232,92	6,57	10,60	28,15	19
Шацький	18034	1811	40,53	101,32	136,28	2,49	8,96	36,60	493

Примітки. *N* — кількість носіїв прізвищ, *S* — кількість прізвищ, I^{jj} — показник ізонімії, *F* — коефіцієнт інбридингу, ν — індекс міграції, *a* — показник різноманітності прізвищ, *H* — показник ентропії розподілу прізвищ, *R* — показник надмірності розподілу прізвищ, *I*₂₀ — індекс місця.

Популяційні показники Дніпропетровської області

Популяції	N	S	$I^{jj} \times 10^{-4}$	$F \times 10^{-5}$	$\nu \times 10^{-4}$	$a \times 10^2$	H	R	I_{20}
Дніпропетровська область	3134360	188651	2,81	7,03	10,94	35,60	14,54	32,79	48
місто Дніпро	920623	97032	2,91	7,29	37,26	34,43	14,14	28,64	12
Райони									
Апостолівський	78012	16316	4,18	10,45	306,45	24,66	12,63	22,26	145
Васильківський	40521	7294	7,94	19,85	310,64	12,99	11,49	24,95	174
Верхньодніпровський	89011	14593	5,69	14,22	197,35	17,92	12,31	25,13	81
Дніпропетровський	89302	19212	5,18	12,95	216,14	19,73	12,67	22,95	232
Криворізький	668959	76542	2,96	7,41	50,43	33,91	14,00	27,63	16
Криничанський	289302	38933	3,73	9,32	92,70	27,07	13,36	26,38	23
Магдалинівський	40691	7774	7,20	18,01	340,85	14,36	11,61	24,21	341
Межівський	31723	5371	10,56	26,39	298,31	9,75	11,01	26,39	582
Нікопольський	179209	30067	3,51	8,78	158,84	28,92	13,21	24,29	27
Новомосковський	118632	21091	5,40	13,49	156,15	18,82	12,64	25,02	148
Павлоградський	113071	21244	4,43	11,06	199,74	23,05	12,74	24,10	106
Петриківський	25766	4965	12,05	30,12	321,74	8,57	10,84	26,02	487
Петропавлівський	53535	10625	6,80	16,99	274,71	15,12	11,91	24,20	302
Покровський	49065	7046	10,31	25,78	197,42	9,88	11,21	28,03	258
П'ятихатський	108021	18475	4,66	11,66	198,48	21,87	12,64	24,39	173
Синельниківський	55646	12003	5,58	13,96	321,71	18,50	12,22	22,46	86
Солонянський	42392	8640	5,82	14,54	405,26	17,91	11,93	22,39	259
Софіївський	27648	6536	8,10	20,25	446,16	12,91	11,50	22,08	272
Томаківський	30604	6447	8,35	20,89	390,81	12,45	11,43	23,29	515
Царичанський	33111	5279	11,27	28,17	267,70	9,11	10,88	27,52	743
Широківський	32002	7763	6,95	17,37	449,32	15,06	11,75	21,47	443
Юр'ївський	17514	4155	9,92	24,79	575,27	10,69	10,98	22,09	1578

Примітки. N — кількість носіїв прізвищ, S — кількість прізвищ, I^{jj} — показник ізонімії, F — коефіцієнт інбридингу, ν — індекс міграції, a — показник різноманітності прізвищ, H — показник ентропії розподілу прізвищ, R — показник надмірності розподілу прізвищ, I_{20} — індекс місця.

Генетичні характеристики популяцій АР Крим області

Популяції	N	S	$I^{jj} \times 10^{-4}$	$F \times 10^{-5}$	$\nu \times 10^{-4}$	$a \times 10^2$	H	R	I_{20}
АР Крим	2189533	186020	1,61	4,03	38,23	62,31	15,07	26,94	46
місто Севастополь	239153	50198	2,05	5,12	204,05	49,81	2,25	87,43	18
місто Сімферополь	373580	72980	1,61	4,02	166,31	63,18	1,69	90,88	16
Райони									
Бахчисарайський	97542	25539	2,90	7,25	353,28	35,72	4,47	73,00	84
Білогірський	71642	18995	3,39	8,47	412,09	30,79	5,87	63,63	132
Джанкойський	136801	31618	2,08	3,35	350,55	49,70	3,75	78,03	33
Кіровський	60612	17453	2,85	7,14	577,72	37,16	6,43	59,53	78
Красногвардійський	97367	25687	2,18	5,45	471,15	48,14	4,49	72,93	43
Красноперекіпський	90329	23725	2,06	5,15	537,62	51,32	4,84	70,58	54
Ленінський	229012	49811	1,84	4,60	237,05	55,61	2,35	86,78	28
Нижньогірський	58998	16743	2,34	5,84	725,52	46,15	6,53	58,77	50
Первомайський	46052	13296	3,01	7,53	721,16	35,79	8,12	47,59	211
Роздольненський	33284	11391	2,65	6,62	1134,50	42,59	9,32	37,96	75
Сакський	219914	50488	1,74	4,35	261,51	59,05	2,45	86,19	24
Сімферопольський	136104	38121	3,30	8,25	222,65	30,99	3,72	78,20	58
Советський	37056	12021	3,20	7,99	844,38	34,18	8,87	41,53	106
Чорноморський	36939	12051	3,14	7,85	862,34	34,86	8,90	41,36	294
Яли Бою (ПБК)	225148	52335	1,93	4,83	229,92	52,98	1,96	88,99	26

Примітки. N — кількість носіїв прізвищ, S — кількість прізвищ, I^{jj} — показник ізонімії, F — коефіцієнт інбридингу, ν — індекс міграції, a — показник різноманітності прізвищ, H — показник ентропії розподілу прізвищ, R — показник надмірності розподілу прізвищ, I_{20} — індекс місця.

Онлайн сервіс створення та перевірки кваліфікованого та удосконаленого електронного підпису

ПРОТОКОЛ
створення та перевірки кваліфікованого та удосконаленого електронного підпису

Дата та час: 12:59:52 15.06.2026

Назва файлу з підписом: Shuba_diss.pdf.asice

Розмір файлу з підписом: 6.4 МБ

Перевірені файли:

Назва файлу без підпису: Shuba_diss.pdf

Розмір файлу без підпису: 7.7 МБ

Результат перевірки підпису: Підпис створено та перевірено успішно. Цілісність даних підтверджено

Підписувач: ШУБА ВЛАДИСЛАВ ВОЛОДИМИРОВИЧ

П.І.Б.: ШУБА ВЛАДИСЛАВ ВОЛОДИМИРОВИЧ

Країна: Україна

РНОКПП: 3627403235

Організація (установа): ФІЗИЧНА ОСОБА

Час підпису (підтверджено кваліфікованою позначкою часу для підпису від Надавача): 12:59:50
15.06.2026

Сертифікат виданий: КНЕДП АЦСК АТ КБ "ПРИВАТБАНК"

Серійний номер: 5E984D526F82F38F0400000087CD19015033FE06

Алгоритм підпису: ДСТУ 4145

Тип підпису: Удосконалений

Тип контейнера: Підпис та дані в архіві (розширений) (ASiC-E)

Формат підпису: З повними даними для перевірки (XAdES-B-LT)

Сертифікат: Кваліфікований

Версія від: 2026.04.06 13:00