

 Open access • Posted Content • DOI:10.1101/2021.03.05.21253014

Investigating the first wave of the COVID-19 pandemic in Ukraine using epidemiological and genomic sequencing data — [Source link](#)

[Yuriy Gankin](#), [Alina Nemira](#), [Vladimir Koniukhovskii](#), [Gerardo Chowell](#) ...+3 more authors

Institutions: [Georgia State University](#), [EPAM Systems](#), [University of South Florida](#)

Published on: 08 Mar 2021 - [medRxiv](#) (Cold Spring Harbor Laboratory Press)

Topics: [Pandemic](#) and [Viral phylodynamics](#)

Related papers:

- [Estimating the effect of social inequalities on the mitigation of COVID-19 across communities in Santiago de Chile.](#)
- [SARS-CoV-2 transmission dynamics in Belarus revealed by genomic and incidence data analysis](#)
- [The impact of early public health interventions on SARS-CoV-2 transmission and evolution](#)
- [Impact of reduction of susceptibility to SARS-CoV-2 on epidemic dynamics in four early-seeded metropolitan regions.](#)
- [Prediction of the COVID-19 epidemic trends based on SEIR and AI models.](#)

Share this paper:    

View more about this paper here: <https://typeset.io/papers/investigating-the-first-wave-of-the-covid-19-pandemic-in-4sp1mb87xk>

Investigating the first stage of the COVID-19 pandemic in Ukraine using epidemiological and genomic data

Yuriy Gankin^{a,*}, Alina Nemira^{b,*}, Vladimir Koniukhovskii^{c,*}, Gerardo Chowell^d, Thomas A. Weppelmann^e, Pavel Skums^{b,**}, Alexander Kirpich^{d,**}

^a*Quantori, Cambridge, Massachusetts, United States of America*

^b*Department of Computer Science, Georgia State University, Atlanta, Georgia, United States of America*

^c*EPAM Systems, Saint Petersburg, Russian Federation*

^d*Department of Population Health Sciences, School of Public Health, Georgia State University, Atlanta, Georgia, United States of America*

^e*Department of Internal Medicine, University of South Florida, Tampa, Florida, United States*

Abstract

The novel coronavirus SARS-CoV-2 was first detected in China in December 2019 and has rapidly spread around the globe. The World Health Organization declared COVID-19 a pandemic in March 2020 just three months after the introduction of the virus. Individual nations have implemented and enforced a variety of social distancing interventions to slow the virus spread, that had different degrees of success. Understanding the role of non-pharmaceutical interventions (NPIs) on COVID-19 transmission in different settings is highly important. While most such studies have focused on China, neighboring Asian countries, Western Europe, and North America, there is a scarcity of studies for Eastern Europe. The aim of this study is to contribute to filling this gap by analyzing the characteristics of the *first months of the epidemic* in Ukraine using agent-based modelling and phylodynamics. Specifically, first we studied the dynamics of COVID-19 incidence and mortality and explored the impact of epidemic NPIs. Our stochastic model suggests, that even a small delay of weeks could have increased the number

*Equal Contribution

**Equal Contribution

Corresponding Author: Alexander Kirpich

Email address: akirpich@gsu.edu (Alexander Kirpich)

Preprint submitted to a research journal

August 5, 2021

of cases by up to 50%, with the potential to overwhelm hospital systems. Second, the genomic data analysis suggests that there have been multiple introductions of SARS-CoV-2 into Ukraine during the early stages of the epidemic. Our findings support the conclusion that the implemented travel restrictions may have had limited impact on the epidemic spread. Third, the basic reproduction number for the epidemic that has been estimated independently from case counts data and from genomic data suggest sustained intra-country transmissions.

Keywords: SARS-CoV-2, agent-based, COVID-19, Ukraine, phylogenetics, phylodynamics

1 Introduction

SARS-CoV-2 virus causing COVID-19 was first detected in December 2019 in the Chinese city of Wuhan [1][2][3][4][5], and has rapidly spread around the globe, prompting the World Health Organization (WHO) to declare a pandemic in March, 2020 [6], just three month after the first reported case. Despite having much lower case-fatality rate than other recent coronavirus pandemics such as the severe acute respiratory syndrome (SARS) and Middle East respiratory syndrome (MERS), the novel coronavirus has claimed more lives just within a few months of introduction than both of those epidemics combined [7]. As of June 29, 2021 there were more than 182 million infections with over 3.9 million deaths [8]. In the absence of vaccines during the early pandemic period, non-pharmaceutical interventions (specifically, non-pharmaceutical epidemic mitigation interventions) were the only tools at the disposal of public health authorities to prevent and to mitigate the virus spread [9][10][11]. The strategies implemented and enforced by governments around the world were highly variable and included frequent sanitation of public spaces, enforced social distancing, wearing of masks, closure of schools, churches, and ban of mass gatherings [12][13][14].

Even well after a year since the epidemic started, fundamental questions regarding the effects of non-pharmaceutical interventions (NPIs) [15][16] and the genomic evolution of SARS-CoV-2 [17][18][19][20] during the introductory period remain. Additionally, recent modeling efforts aimed at shedding light on those questions have mostly focused on China [1][2][3][4], the rest of Asia [21][22][23], Western and Central Europe [24][22][25], and North America [26][27][28][29][5], largely neglecting Africa, the Middle East, and Eastern

26 Europe.

27 In Eastern Europe, post-socialist economics and healthcare systems are
28 inherently different from Western Europe. The available SARS-CoV-2 trans-
29 mission models for Eastern Europe are based on relatively simple *SIR* or
30 similar compartmental models [30][31][32] where individuals are assigned to
31 groups and all individuals within a given group are expected to have the
32 same characteristics. To the best of our knowledge no agent-based modeling
33 studies have been conducted for Ukraine to evaluate the impact of spatial
34 heterogeneities in key transmission drivers such as density of infected indi-
35 viduals and their geographic locations. Furthermore, the number of genomic
36 epidemiology studies on the COVID-19 pandemic in Eastern Europe has been
37 limited. In this paper, we sought to fill the knowledge gap for the Ukrainian
38 epidemic [33], which provides a unique setting for studying the COVID-19
39 spread under the ex-USSR healthcare system, and with the epidemic miti-
40 gation policies similar to the rest of Europe.

41 The first confirmed case in Ukraine was reported on March 3, 2020 and
42 was an individual who has recently traveled from Italy. The first death was
43 reported on March 13, 2020 [34][33]. The Ukrainian government started to
44 implement quarantine measure on March 12, 2020 [33][32] while the cases
45 continued to rise possibly because of the delayed detections of existing infec-
46 tions and returns of infected Ukrainians from abroad [32] (Figure 3B). As a
47 result, more strict measures have been implemented on April 6, 2020 which
48 included the closure of schools, universities, shopping malls, and mandatory
49 mask regiment in public places [32]. Those measures were slightly softened
50 on April 24, 2020 and many services resumed even though some restrictions
51 lasted till the end of June 2020 [33]. As a summary, Ukrainian officials took
52 the epidemic very seriously from the beginning and started to implement the
53 mitigation efforts and corresponding regulations almost immediately after
54 multiple cases in the country have been detected. At the same time imple-
55 mentation of the proposed mitigation efforts did vary from region to region,
56 and so did the compliance with those regulations [35][36].

57 **2. Methods**

58 *2.1. Agent-based Stochastic Model*

59 To investigate the COVID-19 epidemic in Ukraine and to assess its dy-
60 namics under different mitigation scenarios, we utilized our general stochas-
61 tic agent-based modeling framework [37]. The model was adjusted to the

62 Ukrainian settings and fit into the observed Ukrainian data. The summary
63 of the framework together with the adaptation details are outlined below.

64 In brief, the model simulates the epidemics evolving over the discrete
65 time interval $(1, \dots, T)$ with time points $1 \leq t \leq T$ corresponding to calendar
66 days and over the certain geographical area projected on a plain. Infected
67 individuals are represented as agents with multiple characteristics that in-
68 clude geographic coordinates; age; infection time, severity and current sta-
69 tus; disease stage; infectivity rate and infectivity radius which determines
70 how frequently and where it produces secondary infections. The summaries
71 of empirical reproduction numbers of individual agents which are generated
72 by model simulations are used for the estimation of the population basic
73 reproduction number \mathcal{R}_0 [38]. The geographical part of the model includes
74 circular local epidemic spread areas $\mathbf{E} = \{E_1, E_2, \dots, E_I\}$ characterized by
75 their centers and radii. The centers of these areas represent hotspots of the
76 infection introduction into the local population (e.g. transport hubs or ad-
77 ministrative centers). The model incorporates NPI measures via a reduced
78 infection transmission parameters which are effective starting from a certain
79 calendar date customizable within the model.

80 In this study, we used epidemic spread areas and the corresponding in-
81 cidence and mortality data reported by the National Security and Defence
82 Council of Ukraine [39]. It includes daily reports for individual administra-
83 tive regions ("oblast") under the control of the Ukrainian government start-
84 ing from March, 2020. The reported data was separated into three parts.
85 The initially reported cases from March 12, 2020 to April 12, 2020 were ret-
86 rospectively incorporated into the model as the initial conditions [37]. The
87 reported and model-produced data from April 22, 2020 to July 12, 2020 were
88 used for model calibration, and from July 13, 2020 to August 1, 2020 – for
89 model validation. The data before April 22, 2020 were used solely for the
90 initial conditions to increase the model fit robustness, since the initial num-
91 ber of cases was relatively small in comparison to subsequent periods. The
92 August 1, 2020 has been selected as the end date of our simulations to agree
93 with the dates of genomic analysis based on available analyzed SARS-CoV-2
94 sequences collection times [40].

95 Optimization of model parameters has been performed by minimizing
96 the sum of squared differences between the model-produced outputs (across
97 multiple runs) and the calibration data using the Nelder–Mead numerical
98 minimization method [41]. The population basic reproduction number \mathcal{R}_0
99 [38] has been estimated from the model-produced distribution quantiles (5%,

100 median, 95%) of the reproduction numbers of individuals and summarized
101 across multiple stochastic runs [37]. The estimates for \mathcal{R}_0 were produced
102 from the model fit to real data with the assumption that interventions have
103 started almost immediately after the virus introduction.

104 In addition to the simulations based on the model fit to the actual case
105 count, mortality and NPI data, two alternative simulation scenarios were
106 considered under the hypothetical assumptions that NPIs that caused re-
107 duced transmissibility were implemented one (on April 19, 2020) and two
108 (on April 26, 2020) weeks after the simulation start time. The results of
109 simulations under these three scenarios were compared to assess the effect of
110 timely NPI implementations.

111 The additional details about the model can be found in our earlier study
112 [37], and the model implementation tailored to Ukrainian data is available
113 at <https://github.com/quantori/COVID19-Ukraine-Transmission>.

114 *2.2. Genomic Epidemiology Analysis*

115 Sixty high-quality SARS-CoV-2 genomes from Ukraine sampled between
116 April 24, 2020 and August 7, 2020 were extracted from GISAID [40]. These
117 genomes were utilized to construct a maximum likelihood phylogeny using
118 Nextstrain build for SARS-CoV-2 with the default country-specific subsam-
119 pling settings [42]. The obtained timed phylogeny contained Ukrainian se-
120 quences together with a representative subsample of 6479 sequences from
121 other geographic regions, and included inferred ancestral geographic traits of
122 internal nodes. Using these traits, intra-country transmission clusters were
123 identified as clades with the most recent common ancestors (MRCA) esti-
124 mated as originating from Ukraine. For each cluster, confidence intervals for
125 emergence times for MRCA and its parent were also obtained.

126 Next, a phylodynamic analysis of the three largest clusters and the en-
127 tire Ukrainian SARS-CoV-2 population was performed using BEAST v1.10.4
128 [43]. We used a strict molecular clock, HKY+ Γ nucleotide substitution
129 model, a tree prior with exponential growth coalescent. Priors for the pa-
130 rameters were defined in BEAUti v 1.10.4 and were the following: a) nor-
131 mal $\mathcal{N}(mean = 8.0e-4, st.dev = 2.0e-5)$ for the clock rate, b) log-normal
132 $\mathcal{LN}(mean = 1.0, st.dev = 1.25)$ for the population size, c) double ex-
133 ponential (Laplace) distribution $\mathcal{DEXP}(\mu = 0, b = 100)$ for the growth
134 rate, d) normal $\mathcal{N}(mean = 0, st.dev = 1)$ for the freqParameter, e) ex-
135 ponential $\mathcal{EXP}(mean = 0.5, offset = 0)$ for the gammaShape parameter,

136 and f) log-normal $\mathcal{LN}(mean = 1.0, st.dev = 1.25)$ for the kappa param-
137 eter. The detailed parameters file is available in XML format at [https:](https://github.com/alanira/COVID19-Ukraine-phyldynamics)
138 [//github.com/alanira/COVID19-Ukraine-phyldynamics](https://github.com/alanira/COVID19-Ukraine-phyldynamics). The param-
139 eters were estimated after 30,000,000 iterations of Markov Chain Monte Carlo
140 (MCMC) sampling, with the initial 10% values discarded as burn-in. The
141 results were accepted if the effective sample sizes were above 200 for all pa-
142 rameters. The estimated exponential growth rates were used to calculate
143 the basic reproduction numbers \mathcal{R}_0 under the assumption that SARS-CoV-2
144 generation intervals (i.e. times between infection onset and onward trans-
145 mission) were gamma-distributed [44]. We used the formula

$$\hat{\mathcal{R}}_0 = \left[1 + \frac{\hat{f}\hat{\sigma}^2}{\hat{\mu}} \right]^{\frac{\hat{\mu}^2}{\hat{\sigma}^2}}, \quad (1)$$

146 where $\hat{\mu}$ and $\hat{\sigma}$ are the mean and standard deviation of the aforementioned
147 gamma distribution [45][46][47][48]. For these values, we used the estimates
148 $\hat{\mu} = 5.20$ and $\hat{\sigma} = 1.72$ from [49] and $\hat{\mu} = 3.95$ and $\hat{\sigma} = 1.51$ from [49]. The
149 formula (1) defines a strictly monotone transformation of \hat{f} , and, therefore,
150 it also straightforwardly transforms the 95% highest posterior density (HPD)
151 intervals for f into those for \mathcal{R}_0 .

152 3. Results

153 3.1. Agent-based Stochastic Model

154 The visual results of the first scenario (model fit) and the corresponding
155 outputs are summarized in Figure 1. Blue curves in Figure 1 correspond
156 to the reported data. They are captured by the model fits which is also
157 indicated by the corresponding median and 90% pointwise model prediction
158 bands across five hundred runs. The calibration interval is highlighted by
159 cyan background.

160 For each of the three considered scenarios the median value across five
161 hundred simulations were computed at each time point and presented to-
162 gether with the corresponding 5-th and 95-th percentiles across five hundred
163 stochastic realizations to form the 90% prediction intervals (PI-s). The cor-
164 responding results are summarized in Tables 1 and 2 for the model-predicted
165 cases and deaths, respectively. The three scenario summaries from Table 1
166 can be directly compared. For comparison the actual number of reported

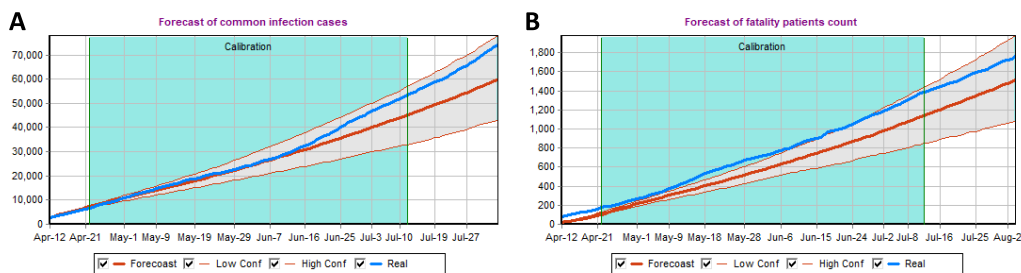


Figure 1: The model predictions together with the reported cases (Panel A) and reported death (Panel B) are presented. The model calibration time interval is highlighted in green. Red lines correspond to the median of the five hundred model-produced runs together with the corresponding 90% prediction bands to quantify the uncertainty. The actual observed case and death counts are displayed in blue for visual comparison.

167 cases by August 1, 2020 was 71,056 [39] which validates the model fit since
168 August 1, 2020 was outside of the calibration interval. The hypothetical
169 April 19th and April 26th intervention start dates produce larger number of
170 cases in comparison to the original fitted scenario. The median estimates
171 can be compared directly. The hypothetical April 19th scenario results in
172 16% predicted increase in cumulative number of cases on June 1, 2020 and
173 in 20% predicted increase in cumulative number of cases on August 1, 2020
174 in comparison to the fitted scenario. The hypothetical April 26th scenario
175 results in 36% predicted increase in cumulative number of cases on June 1,
176 2020 and in 46% predicted increase in cumulative number of cases on August
177 1, 2020 when compared to the fitted scenario.

178 The hypothetical April 19th scenario results in 14% increase in cumulative
179 number of deaths predicted on June 1, 2020 and in 20% increase in cumulative
180 number of deaths predicted on August 1, 2020. The hypothetical April 26th
181 scenario results in 32% increase in cumulative number of deaths predicted on
182 June 1, 2020 and in 46% increase in cumulative number of deaths predicted
183 on August 1, 2020.

184 Interestingly, the median results for the hypothetical April 19th scenario
185 displayed better alignment with the actual data. This suggests the delayed
186 impact of NPIs in transmission mitigation caused by the time needed to
187 put the prescribed measures into effect. Furthermore, the obtained results
188 demonstrate the importance of the early epidemic mitigation measures which
189 cause the reduction in transmission probability parameters and, therefore, a

Scenario	June 1, 2020	July 1, 2020	August 1, 2020
Real Data	24,012	44,998	71,056
April 12, 2020	23,724 (19,093; 28,250)	38,932 (29,181; 48,782)	57,810 (41,864; 75,029)
April 19, 2020	27,511 (22,681; 32,872)	46,193 (36,238; 57,100)	69,358 (50,642; 88,407)
April 26, 2020	32,220 (26,016; 38,645)	55,141 (41,974; 69,992)	84,227 (63,992; 109,069)

Table 1: The model outputs are presented together with the reported data. The predicted number of cumulative cases produced by the model over time for three different epidemic mitigation scenarios for three initiation dates together with the corresponding 90% prediction intervals.

Scenario	June 1, 2020	July 1, 2020	August 1, 2020
Real Data	718	1,173	1,709
April 12, 2020	568 (464; 667)	967 (738; 1,203)	1,461 (1,050; 1,904)
April 19, 2020	653 (541; 783)	1,143 (889; 1,412)	1,762 (1,308; 2,241)
April 26, 2020	752 (612; 906)	1,364 (1,049; 1,705)	2,141 (1,604; 2,719)

Table 2: The model outputs are presented together with the reported data. The predicted number of cumulative death produced by the model over time for three different epidemic mitigation scenarios for three initiation dates together with the corresponding 90% prediction intervals.

190 reduction in the number of cases, and (more importantly) deaths. At the
 191 same time the results for later mitigation efforts implementation dates should
 192 only be interpreted as sensitivity analysis, since the Ukrainian government
 193 has implemented quarantine measures from the beginning of the epidemic and
 194 there were no data to properly estimate the corresponding non-intervention
 195 transmission probability parameters [50]. Therefore, the corresponding non-
 196 quarantine probability parameters have been adopted from the previous anal-
 197 ysis [37].

198 The population basic reproduction number \mathcal{R}_0 estimate during the in-
 199 tervention was estimated to be 1.10 (median) with the corresponding 90%
 200 confidence interval from quantiles equal to (0.24; 1.88).

201 3.2. Genomic epidemiology of SARS-CoV-2

202 Despite a sparse sampling, the observed genomic diversity of SARS-CoV-
 203 2 in Ukraine is substantial, indicating both multiple introductions of the virus
 204 and sustained intra-country evolution (Figure 2). This agrees well with the
 205 patterns observed in other countries [51], and emphasizes the contribution of
 206 global movement of people to the rapid spread of SARS-CoV-2. Specifically,

207 Ukrainian sequences are distributed among eight lineages by the classification
208 of [52] as follows: *B.1* - 50.0 % of genomes, *B.1.1* - 28.3 %, *B.1.1.243* - 8.3 %,
209 *B.1.1.527* - 5.0 %, *B.1.1.325* - 3.3 %, and 1.7 % for each *B.1.131*, *B.1.1.194*,
210 *B*. Similarly, by Nextstrain classification the distribution of lineages is: 19A
211 - 1.7 %, 20A - 51.7 % and 20B - 46.7 % (Figure 2).

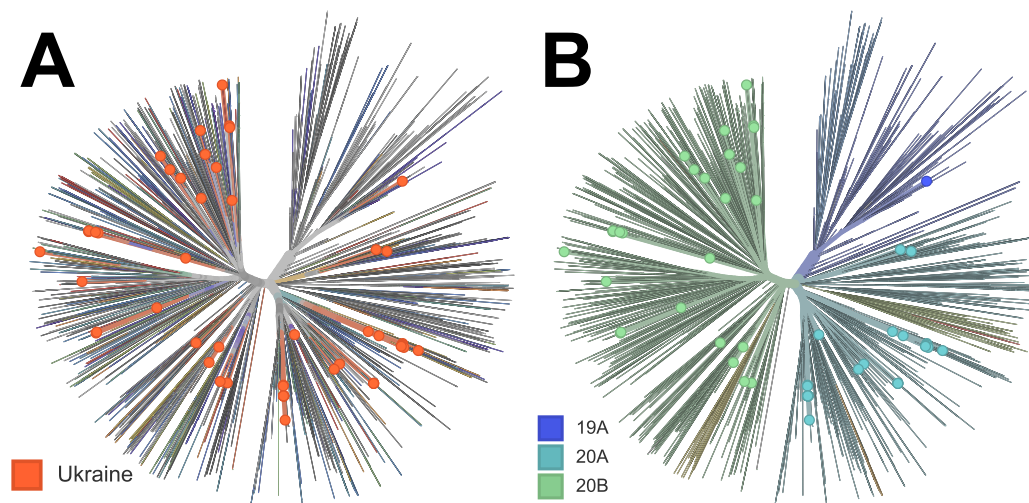


Figure 2: The global phylogenetic tree of SARS-CoV-2 genomes: A) distribution of Ukrainian SARS-CoV-2 genomes inside global SARS-CoV-2 population, B) the same tree with global SARS-CoV-2 lineages highlighted.

212 Seven Ukrainian clusters contain multiple sequences and jointly constitute
213 73.3% of all sampled genomes. Presence of these clusters and the correspond-
214 ing intra-country lineages indicate sustained internal transmissions (Figure
215 3A and Figure A1-A7 in Appendix). For each such lineage, a confidence in-
216 terval of its introduction time can be assessed by the union of the confidence
217 intervals for inferred dates of its Ukrainian MRCA v and the non-Ukrainian
218 parent of v (Table A2 in Appendix).

219 We analyzed these introduction times relatively to the implementation
220 time of the travel ban, that was established on March 16, 2020 [53] for foreign
221 citizens and on March 17, 2020 [54] for all travelers with the exception of
222 Ukrainian citizens returning from abroad. It turned out that three out of
223 seven lineages were most likely introduced into the country after the travel
224 ban date, as indicated by their introduction confidence intervals (Figure 4).
225 Similarly, a single lineage was likely imported before that date; for three
226 remaining lineages the travel ban date falls into their confidence intervals,

227 preventing us from the decisive conclusion, even though the date lies closer to
228 the left ends of all intervals. Thus, the analysis support the hypothesis that
229 the travel restrictions had limited effect on the virus importation control.

230 The estimates of the basic reproduction number \mathcal{R}_0 for three largest lin-
231 eages are summarized in Table 3. All estimates are significantly above one, in-
232 dicated sustained local transmission of SARS-CoV-2 during the first months
233 of the epidemic in Ukraine.

Cluster	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mathcal{R}}_0$ & 95% CI	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mathcal{R}}_0$ & 95% CI
First	5.20	1.72	1.31 (1.12; 1.52)	3.95	1.51	1.23 (1.09; 1.37)
Second	5.20	1.72	1.47 (1.1; 1.98)	3.95	1.51	1.34 (1.07; 1.68)
Third	5.20	1.72	1.48 (1.16; 1.94)	3.95	1.51	1.35 (1.12; 1.65)

Table 3: The estimates of the basic reproduction number \mathcal{R}_0 for three largest clusters together with the corresponding 95% confidence intervals (CI-s). The results are reported for two pairs of generation interval distribution parameters $\hat{\mu}$ and $\hat{\sigma}$ reported by two studies

234 4. Discussion

235 In this study, we have detailed the epidemic characteristics of the first
236 months of the COVID-19 pandemic in Ukraine and studied the effects of
237 NPIs. We considered two complementary approaches based on the stochas-
238 tic modeling applied to incidence data and genomic epidemiology methods
239 applied to sequencing data. Different types of data reflect various aspects
240 of the epidemics, and are prone to different biases. Therefore, such syn-
241 thetic approach facilitates ubiquitous understanding of the early stages of
242 the epidemic in Ukraine.

243 COVID-19 pandemic is characterized by a richness of available data, that
244 allow to utilize agent-based modelling and genomic analysis at the finest pos-
245 sible resolution. In Ukraine, we have an access to public health data on the
246 level of individual regions, which makes agent-based model predictions more
247 comprehensive. Similarly, the advances and cost reduction of next-generation
248 sequencing (NGS) methods allowed rapid genomic data acquisition at early
249 stages of the epidemic [40]. These data processed by advanced phyloge-
250 netic and phylodynamic models allow to assess the virus importation and
251 intra-country transmission dynamics from a “different angle” [55]. Further-
252 more, in cases when two methods produced independent estimations of the
253 basic reproduction number \mathcal{R}_0 , the obtained results are comparable, thus

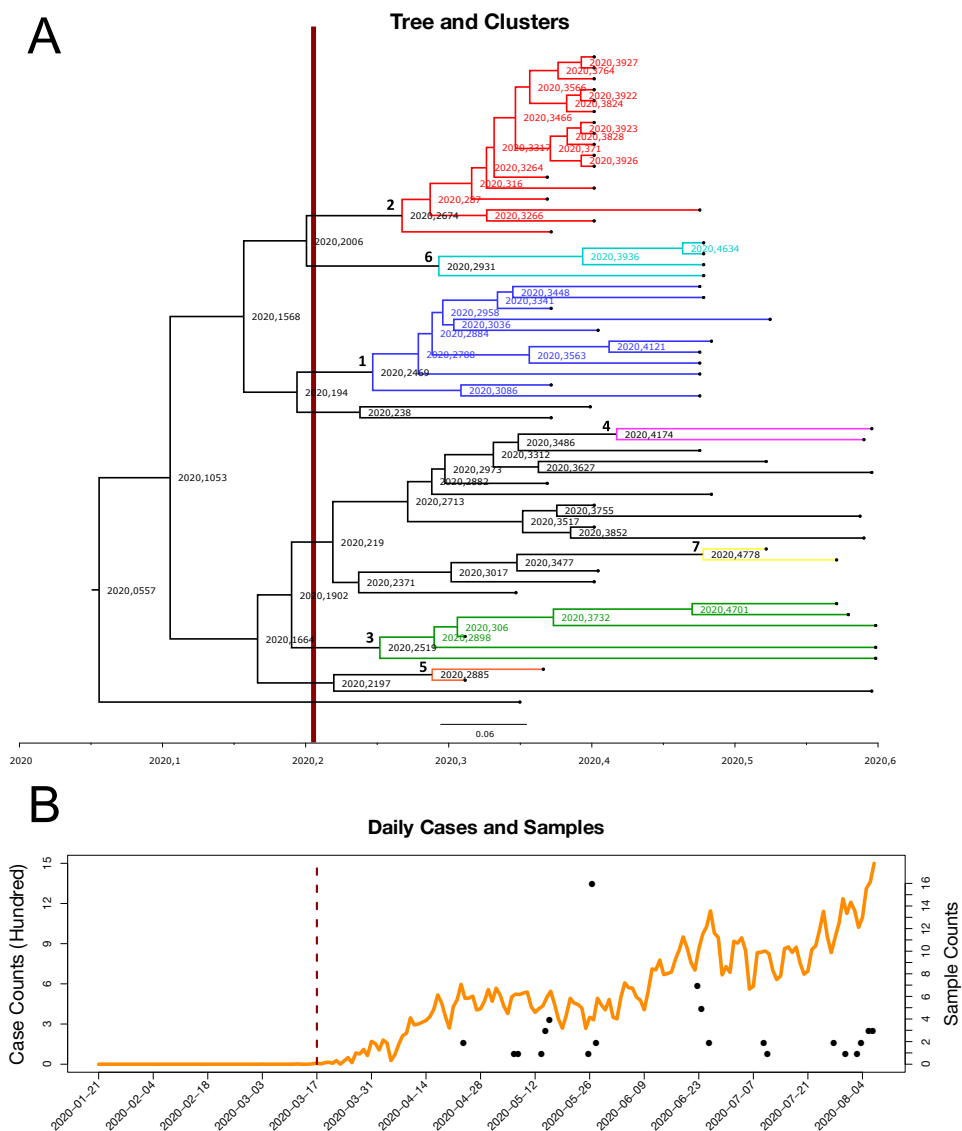


Figure 3: Panel A: The SARS-CoV-2 clusters are presented in the Ukrainian phylogenetic tree. Clusters colored by blue, red, green, pink, orange, azure, yellow, and numbered from one to seven, respectively. Panel B: Daily incidence of reported cases for Ukraine (orange) together with the sample counts and collection dates for sequenced samples (black). The travel restriction has happened on March 17, 2020, which is indicated by a vertical dark red bar time separator in both panels.

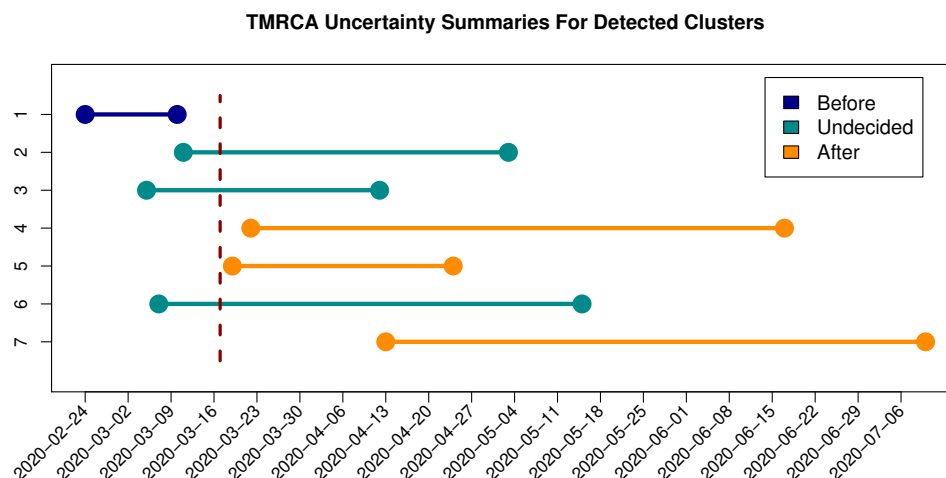


Figure 4: Introduction times for seven largest Ukrainian transmission lineages visualized from Table A2 from the supplement. The estimated intervals for introduction times are depicted as horizontal lines, the border closure date (March 17, 2020) is indicated by a vertical line.

254 highlighting their consistency. The uncertainty estimates for the stochastic
 255 estimates are wider, which may be due to the fact that stochastic model has
 256 more parameters and higher variability in the outputs while phylodynamic
 257 models has pretty strong priors.

258 The study has limitations since the available surveillance incidence and
 259 genomic molecular data are limited. Ukraine is one of the poorest countries
 260 in Europe (based on GDP per capita), and, therefore, the health care infras-
 261 tructure in Ukraine lacks in some parts the resources of its close and distant
 262 European neighbors [56]. As a result, both availability of screening tests and
 263 the reporting of incidence data during the initial epidemic likely (substancially)
 264 underestimated the burden of disease in terms of incidence counts.
 265 Likewise, as the pandemic evolved, the scarcity of genetic sequencing limited
 266 the number of sequence comparison in the phylogenetic analysis. As such, the
 267 actual number of viral clusters of local transmission remains unknown and
 268 should be interpreted as *at least seven* clusters which only form the “tip of the
 269 iceberg” of all transmission clusters. Moreover, the local population compli-
 270 ance with the NPI regulations implemented by officials is always a question,
 271 which might have reduced the effectiveness of such measures [35][36].

272 In summary, this study was among the first to explore the characteris-

273 tics of the initial pandemic as it spread to Ukraine *and* provided additional
274 *genomic analysis* not previously published.

275 **5. Acknowledgements**

276 The authors would like to acknowledge the authors, originating and sub-
277 mitting laboratories of the sequences from GISAID. The list of laboratories
278 is provided in Appendix. The authors would also like to acknowledge Sergey
279 Brutsky for his help with Ukraine incidence data collection and acquisition.

280 **6. Funding**

281 AN was supported by the GSU Brains and Behavior fellowship. PS was
282 supported by the National Institutes of Health grant 1R01EB025022 and by
283 the National Science Foundation grant 2047828.

284 **References**

- 285 [1] N. Zhu, D. Zhang, W. Wang, X. Li, B. Yang, J. Song, X. Zhao, B. Huang,
286 W. Shi, R. Lu, et al., A novel coronavirus from patients with pneumonia
287 in china, 2019, *New England Journal of Medicine* (2020).
- 288 [2] H. Lu, C. W. Stratton, Y.-W. Tang, Outbreak of pneumonia of un-
289 known etiology in wuhan china: the mystery and the miracle, *Journal*
290 *of Medical Virology* (2020).
- 291 [3] C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan,
292 J. Xu, X. Gu, et al., Clinical features of patients infected with 2019
293 novel coronavirus in wuhan, china, *The lancet* 395 (2020) 497–506.
- 294 [4] N. Chen, M. Zhou, X. Dong, J. Qu, F. Gong, Y. Han, Y. Qiu, J. Wang,
295 Y. Liu, Y. Wei, et al., Epidemiological and clinical characteristics of 99
296 cases of 2019 novel coronavirus pneumonia in wuhan, china: a descrip-
297 tive study, *The Lancet* 395 (2020) 507–513.
- 298 [5] M. Ciotti, M. Ciccozzi, A. Terrinoni, W.-C. Jiang, C.-B. Wang,
299 S. Bernardini, The covid-19 pandemic, *Critical Reviews in Clinical*
300 *Laboratory Sciences* 57 (2020) 365–388.
- 301 [6] WHO, [https://www.who.int/dg/speeches/detail/who-director-](https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-COVID-19---11-march-2020)
302 [general-s-opening-remarks-at-the-media-briefing-on-COVID-](https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-COVID-19---11-march-2020)
303 [19---11-march-2020](https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-COVID-19---11-march-2020), 2020. Accessed: 2021-02-20.
- 304 [7] E. Mahase, Coronavirus: covid-19 has killed more people than sars and
305 mers combined, despite lower case fatality rate, *BMJ* (2020).
- 306 [8] Johns Hopkins University - Coronavirus Resource Center, [https://](https://coronavirus.jhu.edu/map.html)
307 coronavirus.jhu.edu/map.html, 2021. Accessed: 2021-03-01.
- 308 [9] W.-H. Chen, U. Strych, P. J. Hotez, M. E. Bottazzi, The SARS-CoV-2
309 vaccine pipeline: an overview, *Current tropical medicine reports* (2020)
310 1–4.
- 311 [10] S. Caddy, Russian SARS-CoV-2 vaccine, *BMJ: British Medical Journal*
312 (Online) 370 (2020).

- 313 [11] A. Mullard, How COVID vaccines are being divvied up around the
314 world, *Nature* (2020). URL: [https://doi.org/10.1038/d41586-020-](https://doi.org/10.1038/d41586-020-03370-6)
315 [03370-6](https://doi.org/10.1038/d41586-020-03370-6). doi:10.1038/d41586-020-03370-6.
- 316 [12] N. Islam, S. J. Sharp, G. Chowell, S. Shabnam, I. Kawachi, B. Lacey,
317 J. M. Massaro, R. B. D’Agostino, M. White, Physical distancing inter-
318 ventions and incidence of coronavirus disease 2019: natural experiment
319 in 149 countries, *bmj* 370 (2020).
- 320 [13] WHO, [https://www.who.int/emergencies/diseases/novel-](https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public)
321 [coronavirus-2019/advice-for-public](https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public), 2020. Accessed: 2021-02-20.
- 322 [14] CDC, [https://www.cdc.gov/coronavirus/2019-ncov/prevent-](https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/social-distancing.html)
323 [getting-sick/social-distancing.html](https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/social-distancing.html), 2020. Accessed: 2021-02-20.
- 324 [15] F. Di Gennaro, D. Pizzol, C. Marotta, M. Antunes, V. Racalbutto,
325 N. Veronese, L. Smith, Coronavirus diseases (covid-19) current sta-
326 tus and future perspectives: a narrative review, *International journal of*
327 *environmental research and public health* 17 (2020) 2690.
- 328 [16] K. Iyengar, A. Mabrouk, V. K. Jain, A. Venkatesan, R. Vaishya, Learn-
329 ing opportunities from covid-19 and future effects on health care system,
330 *Diabetes & Metabolic Syndrome: Clinical Research & Reviews* 14 (2020)
331 943–946.
- 332 [17] K. Tizaoui, I. Zidi, K. H. Lee, R. Abou Ghayda, S. H. Hong, H. Li,
333 L. Smith, A. Koyanagi, L. Jacob, A. Kronbichler, et al., Update of
334 the current knowledge on genetics, evolution, immunopathogenesis, and
335 transmission for coronavirus disease 19 (covid-19), *International Journal*
336 *of Biological Sciences* 16 (2020) 2906.
- 337 [18] H. Yi, 2019 novel coronavirus is undergoing active recombination, *Clin-*
338 *ical Infectious Diseases* (2020).
- 339 [19] C. Li, Y. Yang, L. Ren, Genetic evolution analysis of 2019 novel coro-
340 navirus and coronavirus from other species, *Infection, Genetics and*
341 *Evolution* (2020) 104285.
- 342 [20] D. Benvenuto, M. Giovanetti, A. Ciccozzi, S. Spoto, S. Angeletti, M. Ci-
343 ccozzi, The 2019-new coronavirus epidemic: evidence for virus evolution,
344 *Journal of medical virology* 92 (2020) 455–459.

- 345 [21] W. E. Wei, Z. Li, C. J. Chiew, S. E. Yong, M. P. Toh, V. J. Lee,
346 Presymptomatic transmission of SARS-CoV-2—singapore, january 23–
347 march 16, 2020, *Morbidity and Mortality Weekly Report* 69 (2020) 411.
- 348 [22] N. Yamamoto, G. Bauer, Apparent difference in fatalities between cen-
349 tral europe and east asia due to SARS-CoV-2 and covid-19: Four hy-
350 potheses for possible explanation, *Medical Hypotheses* (2020) 110160.
- 351 [23] D. Roy, S. Tripathy, S. K. Kar, N. Sharma, S. K. Verma, V. Kaushal,
352 Study of knowledge, attitude, anxiety & perceived mental healthcare
353 need in indian population during covid-19 pandemic, *Asian Journal of*
354 *Psychiatry* (2020) 102083.
- 355 [24] H. Salje, C. T. Kiem, N. Lefrancq, N. Courtejoie, P. Bosetti, J. Paireau,
356 A. Andronico, N. Hozé, J. Richet, C.-L. Dubost, et al., Estimating the
357 burden of SARS-CoV-2 in france, *Science* (2020).
- 358 [25] A. Fokas, N. Dikaios, G. Kastis, Mathematical models and deep learning
359 for predicting the number of individuals reported to be infected with
360 SARS-CoV-2, *Journal of the Royal Society Interface* 17 (2020) 20200494.
- 361 [26] I. Ghinai, T. D. McPherson, J. C. Hunter, H. L. Kirking, D. Chris-
362 tiansen, K. Joshi, R. Rubin, S. Morales-Estrada, S. R. Black, M. Pacilli,
363 et al., First known person-to-person transmission of severe acute respi-
364 ratory syndrome coronavirus 2 (SARS-CoV-2) in the usa, *The Lancet*
365 (2020).
- 366 [27] S. Ellington, P. Strid, V. T. Tong, K. Woodworth, R. R. Galang, L. D.
367 Zambrano, J. Nahabedian, K. Anderson, S. M. Gilboa, Characteristics
368 of women of reproductive age with laboratory-confirmed SARS-CoV-2
369 infection by pregnancy status—united states, january 22–june 7, 2020,
370 *Morbidity and Mortality Weekly Report* 69 (2020) 769.
- 371 [28] J. D. Silverman, N. Hupert, A. D. Washburne, Using influenza surveil-
372 lance networks to estimate state-specific prevalence of SARS-CoV-2 in
373 the united states, *Science translational medicine* 12 (2020).
- 374 [29] J. R. Fauver, M. E. Petrone, E. B. Hodcroft, K. Shioda, H. Y. Ehrlich,
375 A. G. Watts, C. B. Vogels, A. F. Brito, T. Alpert, A. Muyombwe, et al.,
376 Coast-to-coast spread of SARS-CoV-2 during the early epidemic in the
377 united states, *Cell* (2020).

- 378 [30] A. Issanov, Y. Amanbek, A. Abbay, S. Adambekov, M. Aljofan,
379 A. Kashkynbayev, A. Gaipov, Covid-19 outbreak in post-soviet states:
380 modeling the best and worst possible scenarios (2020).
- 381 [31] I. Nesteruk, Long-term predictions for covid-19 pandemic dynamics in
382 ukraine, austria and italy, Europe PMC (2020).
- 383 [32] Y. N. Kyrychko, K. B. Blyuss, I. Brovchenko, Mathematical modelling of
384 the dynamics and containment of covid-19 in ukraine, Scientific reports
385 10 (2020) 1–11.
- 386 [33] A. Åslund, Responses to the covid-19 crisis in russia, ukraine, and
387 belarus, Eurasian Geography and Economics (2020) 1–14.
- 388 [34] Kyiv Post, [https://www.kyivpost.com/ukraine-politics/
389 developing-first-coronavirus-case-identified-in-
390 ukraine.html](https://www.kyivpost.com/ukraine-politics/developing-first-coronavirus-case-identified-in-ukraine.html), 2020. Accessed: 2021-02-20.
- 391 [35] UN report: Impact of COVID-19 on human rights in Ukraine
392 , [https://www.ohchr.org/Documents/Countries/UA/Ukraine_COVID-
393 19_HR_impact_EN.pdf](https://www.ohchr.org/Documents/Countries/UA/Ukraine_COVID-19_HR_impact_EN.pdf), 2020. Accessed: 2021-06-29.
- 394 [36] Atlantic Council: Ukraine’s local authorities and the Covid-19
395 pandemic, [https://www.atlanticcouncil.org/blogs/ukrainealert/
396 ukraines-local-authorities-and-the-covid-19-pandemic/](https://www.atlanticcouncil.org/blogs/ukrainealert/ukraines-local-authorities-and-the-covid-19-pandemic/), 2021.
397 Accessed: 2021-06-29.
- 398 [37] A. Kirpich, V. Koniukhovskii, V. Shvartc, P. Skums, T. A. Weppelmann,
399 E. Imyanitov, S. Semyonov, K. Barsukov, Y. Gankin, Development of
400 an interactive, agent-based local stochastic model of COVID-19 trans-
401 mission and evaluation of mitigation strategies illustrated for the state
402 of Massachusetts, USA, PLoS ONE 16 (2021) e0247182.
- 403 [38] G. N. Milligan, A. D. Barrett, An Essential Guide, Wiley Online Library,
404 2015.
- 405 [39] National Security And Defence Council of Ukraine, [bhttps://
406 covid19.rnbo.gov.ua/](https://covid19.rnbo.gov.ua/), 2021. Accessed: 2021-02-20.
- 407 [40] GISAID, A global initiative on sharing avian flu data., [https://
408 www.gisaid.org/](https://www.gisaid.org/), 2020. Accessed: 2021-02-20.

- 409 [41] J. A. Nelder, R. Mead, A simplex method for function minimization,
410 The computer journal 7 (1965) 308–313.
- 411 [42] github.com, <https://github.com/nextstrain/ncov>, 2021. Accessed:
412 2020-11-01.
- 413 [43] M. A. Suchard, P. Lemey, G. Baele, D. L. Ayres, A. J. Drummond,
414 A. Rambaut, Bayesian phylogenetic and phylodynamic data integration
415 using beast 1.10, Virus evolution 4 (2018) vey016.
- 416 [44] Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, K. S.
417 Leung, E. H. Lau, J. Y. Wong, et al., Early transmission dynamics in
418 wuhan, china, of novel coronavirus–infected pneumonia, New England
419 Journal of Medicine (2020).
- 420 [45] BEAST Software, https://beast.community/estimating_R0.html,
421 2020. Accessed: 2021-02-20.
- 422 [46] BEAST Software, [https://beast.community/
423 phylogenetics_of_epidemic_influenza](https://beast.community/phylogenetics_of_epidemic_influenza), 2020. Accessed: 2021-02-20.
- 424 [47] J. Wallinga, M. Lipsitch, How generation intervals shape the relationship
425 between growth rates and reproductive numbers, Proceedings of the
426 Royal Society B: Biological Sciences 274 (2007) 599–604.
- 427 [48] N. C. Grassly, C. Fraser, Mathematical models of infectious disease
428 transmission, Nature Reviews Microbiology 6 (2008) 477–487.
- 429 [49] T. Ganyani, C. Kremer, D. Chen, A. Torneri, C. Faes, J. Wallinga,
430 N. Hens, Estimating the generation interval for coronavirus disease
431 (covid-19) based on symptom onset data, march 2020, Eurosurveillance
432 25 (2020) 2000257.
- 433 [50] github.com, [https://github.com/quantori/COVID19-Ukraine-
434 Transmission](https://github.com/quantori/COVID19-Ukraine-Transmission), 2021. Accessed: 2020-11-01.
- 435 [51] J. L. Geoghegan, X. Ren, M. Storey, J. Hadfield, L. Jelley, S. Jefferies,
436 J. Sherwood, S. Paine, S. Huang, J. Douglas, et al., Genomic epidemi-
437 ology reveals transmission patterns and dynamics of SARS-CoV-2 in
438 aotearoa new zealand, Nature communications 11 (2020) 1–7.

- 439 [52] A. Rambaut, E. C. Holmes, Á. O’Toole, V. Hill, J. T. McCrone, C. Ruis,
440 L. du Plessis, O. G. Pybus, A dynamic nomenclature proposal for SARS-
441 CoV-2 lineages to assist genomic epidemiology, *Nature microbiology* 5
442 (2020) 1403–1407.
- 443 [53] baltsi.mfa.gov.ua - No Entry for Foreigners - May 26, 2020 ,
444 [https://baltsi.mfa.gov.ua/ru/news/timchasova-zaborona-do-
445 24-kvitnya-2020-r-na-peretinannya-derzhavnogo-kordonu-na-
446 vyzd-v-ukrayinu-dlya-inozemciv-ta-osib-bez-gromadyanstva](https://baltsi.mfa.gov.ua/ru/news/timchasova-zaborona-do-24-kvitnya-2020-r-na-peretinannya-derzhavnogo-kordonu-navyizd-v-ukrayinu-dlya-inozemciv-ta-osib-bez-gromadyanstva),
447 2021. Accessed: 2021-02-20.
- 448 [54] Closed borders, [https://www.kmu.gov.ua/news/uryad-
449 timchasovo-zaboroniv-vyizd-v-ukrayinu-inozemciv-ta-zakriv-
450 mizhnarodne-pasazhirske-aviaspoluchennya](https://www.kmu.gov.ua/news/uryad-timchasovo-zaboroniv-vyizd-v-ukrayinu-inozemciv-ta-zakriv-mizhnarodne-pasazhirske-aviaspoluchennya), 2021. Accessed:
451 2021-02-20.
- 452 [55] G. L. Armstrong, D. R. MacCannell, J. Taylor, H. A. Carleton, E. B.
453 Neuhaus, R. S. Bradbury, J. E. Posey, M. Gwinn, Pathogen genomics in
454 public health, *New England Journal of Medicine* 381 (2019) 2569–2580.
- 455 [56] [emerging-europe.com](https://emerging-europe.com/news/imf-ukraine-is-emerging-europes-poorest-country/), [https://emerging-europe.com/news/imf-
456 ukraine-is-emerging-europes-poorest-country/](https://emerging-europe.com/news/imf-ukraine-is-emerging-europes-poorest-country/), 2021. Accessed:
457 2021-03-01.