

Supplementary Figures

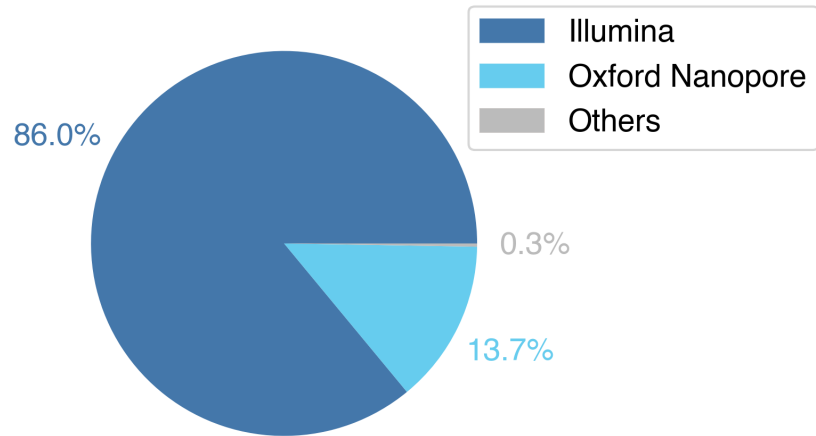


Figure S1 Sequenced genomes present on SRA database sequenced by Illumina, Nanopore, and others sequencing technologies that include Pacbio, IonTorrent and BGISEQ.

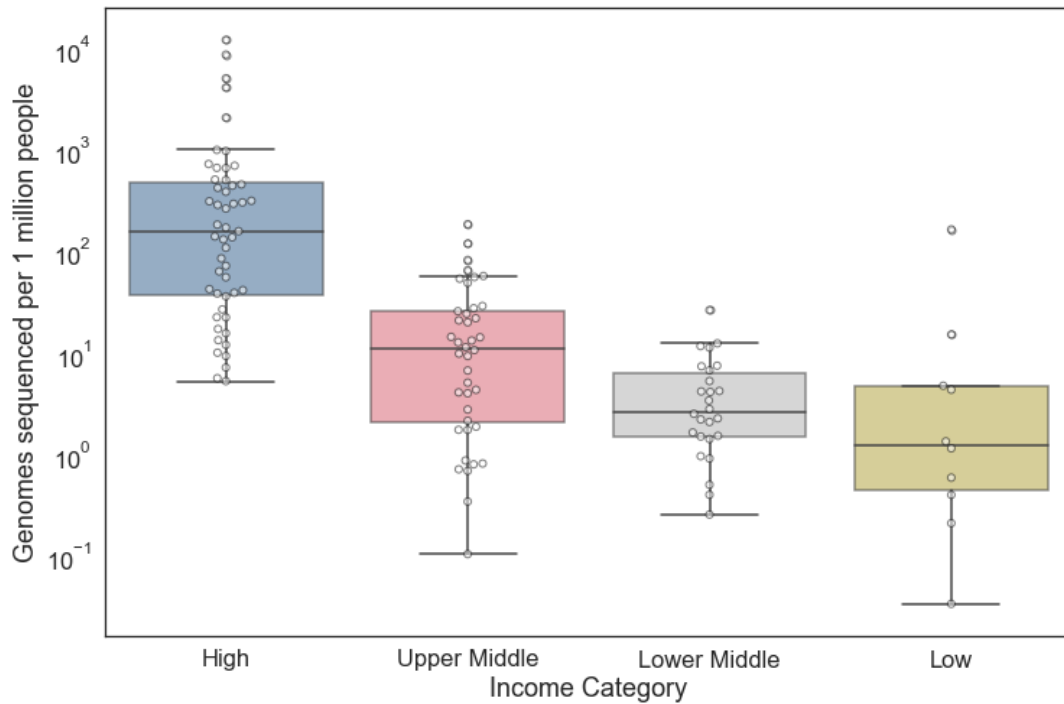


Figure S2 The number of genomes sequenced per million people per income category

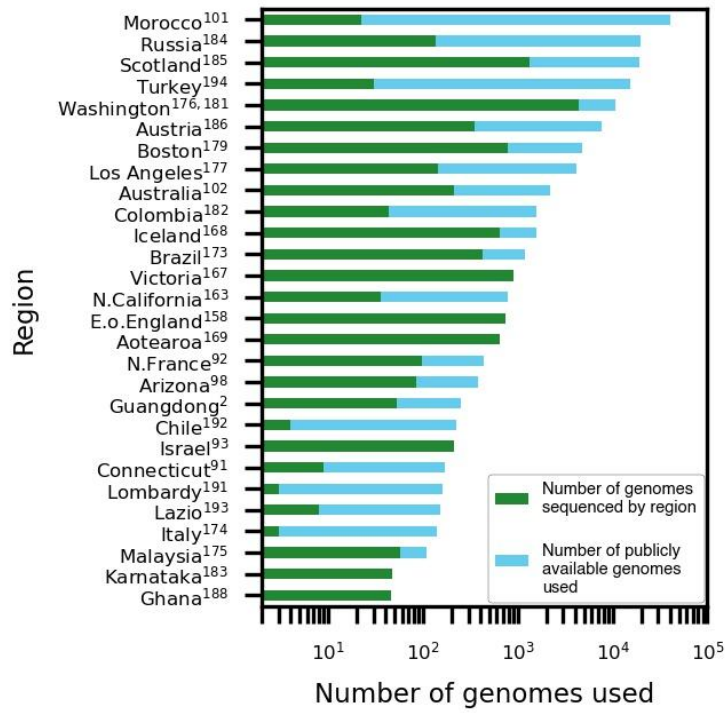


Figure S3 Number of genomes sequenced and publicly available across regions that were used for outbreak investigation studies in different regions.

Supplementary Table 1: Public repositories collecting SARS-CoV-2 genomes.

Database	Genomes*	Raw data*	Link
GISAID	1,013,054	-	https://www.gisaid.org/
COG-UK	413,687	-	https://www.cogconsortium.uk/
NCBI GenBank	151,209	333,257	https://www.ncbi.nlm.nih.gov/sars-cov-2/

* - Submitted by April 7, 2021

Supplementary Table 2: Summary of genomic analyses of SARS-CoV-2

The region covered (country)	Date of publication	Date of sample collection	Number of genomes sequenced by country/region	Total number of genomes used for analysis (combined with publicly available data)	Phylogeny Method	Citation
Guangdong Province (China)	April 30 2020	January 9-31, 2020; February 5, 2020, February 10-12, 2020	53	250	-Maximum likelihood phylogeny	²
Northern France	April 29, 2020	December 2019	97	438	-Maximum likelihood phylogeny	⁹²
Connecticut (United States)	May 28, 2020	March 14, 2020	9	168	-Maximum likelihood phylogeny -Clade-defining nucleotide substitution	⁹¹
Victoria, Australia	September 1, 2020	January 6, 2020 - April 14, 2020	903	903	-Maximum likelihood phylogeny -Bayesian maximum clade credibility tree -Bayesian	¹⁶⁸

					phylodynamic analysis	
Iceland	June 11, 2020	March 31, 2020 - April 1, 2020	643	1547	-Median-joining haplotype network	169
Israel	May 22, 2020	March 17, 2020 - April 22, 2020	212	212	-Maximum likelihood phylogeny -Phylogeographic analysis	93
Northern California (United States)	July 31, 2020	January 29, 2020 - March 20, 2020	36	789	-Maximum likelihood phylogeny	164
East of England (United Kingdom)	July 14, 2020	March 13, 2020 - April 25, 2020	747	747	-Maximum likelihood phylogeny	158
Aotearoa (New Zealand)	August 20, 2020	February 26, 2020 - May 22, 2020	649	649	-Maximum likelihood time-scaled -Bayesian maximum clade credibility tree	170
Multi-country Australia India United Kingdom China	August 10, 2020	May 27, 2020	1364 109 400	18,168	-Maximum likelihood phylogeny -Haplotype networks	171

United States			112 365			
Netherlands	July 16, 2020	February 25 -28, 2020; March 1 - 11, 2020	189	Combined with all available full-length SARS-CoV-2 genomes from GISAID as of March 22, 2020	-Maximum likelihood phylogeny -Bayesian phylogeny -Maximum clade credibility tree	167
Japan	June 5, 2020	March 5 - 15, 2020	10	Combined with all available full-length SARS-CoV-2 genomes from GISAID as of March 30, 2020	-Haplotype networks	172
Morocco	June 25, 2020	June 7, 2020	22	40,390	-Maximum likelihood phylogeny -Time resolved tree	101
Global	April 28, 2020	March 4, 2020	160	413	-Phylogenetic network analysis	173
Australia	July 9, 2020	January 21, 2020 - March 28, 2020	209	2194	-Genomic clustering -Maximum likelihood phylogeny -agent-based	102

					modeling	
Brazil	July 23, 2020	March 5, 2020 - April 30, 2020	427	1182	-Maximum likelihood and molecular clock phylogenies -Time-resolved maximum clade credibility phylogeny	174
Italy	March 27, 2020	N/A	3 (Sequences from GISAID)	141	-Maximum likelihood phylogeny	175
Malaysia	August 27, 2020	February - April 2020	58	108	-Maximum likelihood phylogeny	176
Washington (US)	October 30, 2020	February 20, 2020 - March 15, 2020	455	493	-Maximum likelihood phylogeny	177
Los Angeles, California (US)	September 18, 2020	February 28, 2020 - June 22, 2020	142	4095 (286 Sequences from LA county, of which 144 are obtained from	-Maximum parsimony phylogeny	178

				GISAID + 3809 genomes from across the world)		
Houston, Texas (US)	September 29, 2020	March 6, 2020 - July 7, 2020	5,085	Combined with genomes acquired through GISAID on 19 August 2020.	-Maximum likelihood phylogeny	179
Baltimore, Maryland (US)	August 23, 2020	March 11 - 31, 2020	114	Combined with datasets from GISAID having 886 and 2593 genomes	-Maximum likelihood phylogeny	94
Boston, Massachusetts (US)	August 25, 2020	January 29, 2020 - April 18 2020 March 4, 2020 - May 9, 2020	772	4783	Markov chain Monte Carlo Phylodynamics	180
Norfolk (United Kingdom)	September 30, 2020	March 2020 - August 2020	1035	The phylogenetic tree was estimated as part of the COG-UK phylogenetic pipeline (2020-09-07)	Maximum likelihood phylogeny	181

Washington (US)	September 30, 2020	February 2020 - July 2020	3940	10051	Markov chain Monte Carlo Phylodynamics	182
Arizona (US)	September 4, 2020	March 5, 2020 - April 2, 2020	84	376	-Maximum likelihood phylogeny -Bayesian maximum clade credibility phylogeny -Markov chain Monte Carlo Phylodynamics	98
Colombia (South America)	September 6, 2020	January 1, 2020 - April 19, 2020	43	1583	-Maximum likelihood phylogeny	183
Karnataka (India)	July 24, 2020	March 5 2020 - May 21, 2020	47	47	Maximum likelihood phylogeny	184
Russia	July 17, 2020	March 11 2020 - April 23 2020	135	19834	Maximum likelihood phylogeny	185
Scotland	December 21, 2020	February 2020 - March 2020	1,314	19370	Maximum likelihood phylogeny	186
Austria	December 9th, 2020	February 24 2020 - May 7	345	7666	Maximum likelihood phylogeny	187

		2020				
Global	August 19th, 2020	Up to March 30th, 2020	2492	2492 (all from GISAID database)	Amino-acid heterogeneity analysis	188
Ghana	December 16th, 2020	March 12th–April 1st 2020, May 25th - 27th 2020	46	46	Maximum likelihood phylogeny	189
Global	December 22nd, 2020	April 17th, 2003 - March 24th, 2020	120	120	Maximum likelihood phylogeny & MRP supertree method	190
Global	September 23rd, 2020	December 2019 - April 18th 2020	Greater than 6,000	Greater than 6,000	Maximum likelihood phylogeny	191
Lombardy (Italy)	March 29th, 2020	February 21st, 2020	3	160 (157 from GISAID)	Bayesian Markov Chain Monte Carlo method	192
Chile	March 29th, 2020	March 3rd - March 5th, 2020	4	222 (218 from GISAID)	Maximum likelihood phylogeny	193
Lazio (Italy)	August 26th, 2020	February 27th, 2020 - March 23rd, 2020	8	150	Maximum likelihood phylogeny	194

Turkey	June 21st, 2020	May 1st, 2020	30	15,277	Maximum likelihood phylogeny	195
Global	July 17th, 2020	December 24th, 2019-February 9th, 2020	112	112	Maximum likelihood phylogeny	196

Supplementary Table 3: Summary of wastewater analyses of SARS-CoV-2

The region covered (country)	Date of publication	Date of sample collection	Number of samples taken from WWTP (waste water treatment plants)	Total number of samples used for analysis (combined with publicly available data)	Method of NAAT Analysis	Citation
Federal State of North Rhine-Westphalia (Germany)	January 20, 2021	April 8, 2020	9	9	One-Step RT-qPCR	197
Rome and Milan (Italy)	September 20, 2020	February 3, 2020 - April 2, 2020	12	12	RT-PCR Analysis:ORFlab, Novel, nested RT-PCR targeting spike region	198

					RT-qPCR targeting RdRP gene	
Australia	August 1, 2020	February 24, March 20, March 26, March 28-30, April 4, 2020	3	3	RT-qPCR	126
Paris (France)	May 6, 2020	March 2020	3	3	RT-qPCR	122
The Netherlands	May 20, 2020	February, March 2020	4	4	RT-qPCR	125
New Haven (Connecticut, US)	September 18, 2020	March 19, 2020 - June 1, 2020	73	73	qRT-PCR	114
Massachusetts (US)	April 7, 2020	January 8, 2020 January 11, 2020 March 18-25, 2020	12	12	RT-qPCR	123
Istanbul (Turkey)	May 6, 2020	April 21-25, 2020	9 (7 WWTP and 2 manholes)	9	RT-qPCR	199

Israel	May 1, 2020	N/A	5	5	qPCR	200
Valencia (Spain)	April 29, 2020	February 12- April 14, 2020	15	15	RT-qPCR	121
Wuhan (China) *used wastewater in septic tanks in hospitals	May 14, 2020	February 26, 2020 March 1, 2020 March 10, 2020	3	3	RT-qPCR	201
Santa Catalina (Brazil)	June 29, 2020	October 30, 2019 - March 4, 2020	6	6	RT-qPCR	202
Japan	June 18, 2020	March 17, 2020 - May 7, 2020	13 (5 from WWTP and 3 from river)	13	qPCR Nested PCR	203
India	June 18, 2020	May 8, 2020 May 27, 2020	N/A	N/A	RT-PCR	204
Bozeman (Montana)	September 22, 2020	March 30, 2020 - June 12, 2020	17	17	RT-PCR	205

Czech Republic	July 30, 2020	April - June 2020	112 (from 33 different WWTP)	112	RT-qPCR	206
Alameda County Marin County (Northern California, US)	September 14, 2020	May 19, 2020 - July 15, 2020	N/A	N/A	qRT-PCR	120
Santa Clara County (California, US)	September 15, 2020	March 16 - May 31, June 2 - July 12	89	89	RT-qPCR	207
Frankfurt (Germany)	October 27th, 2020	April 2020 - August 2020	44	44	RT-qPCR	208
Ottawa, Ontario and Gatineau, Quebec, Canada	October 23rd, 2020	April 1st, 2020 - June 30th, 2020	23 (14 from Ottawa, and 9 from Gatineau)	23	RT-qPCR and RT-ddPCR	209
Istanbul (Turkey)	May 16th, 2020	May 7th, 2020	9	9	RT-qPCR	210
Quito River	November 15th, 2020	June 5th, 2020	3	3	qRT-PCR	211

(Ecuador)						
Cruiseship in Australia, Flights from Los Angeles–Brisbane, Hong Kong–Brisbane, New Delhi–Sydney	July 5th, 2020	April 23rd, 2020; July 6th, 2020; October 5th, 2020	21	21	RT-qPCR and RT-ddPCR	212
Murcia (Spain)	16 May 2020	2 March to 14 April, 2020	72	72	RT-qPCR	213
Milan Metropolitan Area, Italy	May 5, 2020.	April, 14th and April, 22th, 2020	16	16	RT-PCR	214
Milan/Lombardy, Turin/Piedmont and Bologna/Emilia Romagna (Italy)	June 26, 2020	12 September 2018 and 19 June 2019, 9 October 2019 and 28 February 2020	40	40	nested RT-PCR and RT-qPCR	215
Montpellier, France	July 09, 2020. July 09, 2020	May 7th, 18th, 26th, June 4th, 15th and 25th	N/A	N/A	RT-qPCR	216

Louisiana, USA	November 15th, 2020	January to April 2020	15	15	RT-qPCR	217
Syracuse, NY and Onondaga County, NY	May 23rd, 2020	May 6th and 13th, 2020	11	11	RT-qPCR	218