UK lineages summary report

This report gives summaries of UK specific lineages sequenced by SANG for week 2020-05-29. There are time lags due to batching, curation and analysis, the most recently sampled sequence is 2020-05-10. The analysis (eg time since last sample) is therefore undertaken from this date. 4791 sequences in the UK from the sequencing centre SANG have been included in this analysis.

A few notes: the size of a lineage may be due to a low amount of transmission of this lineage, but it is likely also that it just hasn't been sampled as frequently, especially for newer lineages. It's also important to realise that these lineages are *estimates* of how we think the virus is spreading in the UK after being introduced from abroad, as the low evolutionary rate of the virus makes it difficult to separate lineages with certainty.

The minimum number of introductions is 6130 and the maximum is 9084

Sequences which were replicates or too error-prone were removed from this analysis.

1759 are lineages which only contained five sequences or fewer, and so have been left out of visualisation in the interests of clarity

Furthermore, those sequences which haven't been sampled in the last month are not shown.

Of the 93 that remain: 40 are pending extinction, ie last seen three weeks ago. 36 lineages have gone quiet, ie haven't been seen this week. 7 lineages have reactivated. 10 lineages have been continuously circulating.

The following table contains information about the ten largest lineages lineages and the number of sequences the dataset. Information about other lienages is found in the appendix, along with the raw data for all of the other figures.

Each entry is the count of sequences from each lineage in each country, with the percentage of the total sequences from that lineage that this count represents.

"Activity score" is calculated by taking the average gap between sampling for each lineage, and dividing it by the number of days since the lineage was last sampled. Therefore the higher the number, the more active the lineage is. If the score is above 1, then it has been sampled *more* recently than expected given its average gap size. We might interpret this as an increase in activity. If the score is below 1, it has been sampled *less* recently than expect given its average gap size, so we might interpret this as a decrease in activity.

The global lineages are correct as of the data release on 2020-05-19

It is written to "summary_files" as "lineage_summary.tsv" for further use, and the full list of lineages is available in the same directory as "all_lineages.csv"

Lineage name	England	Scotland	Northern Ireland	Date range	Total sequences	Global lineage	Time since last sample (days)	Activity score
UK5	296 (96.73%)	9 (2.94%)	1 (0.33%)	Mar-07, May-09	306	B.1.1, B.1.1.1	1	0.2066
UK2464	92 (84.4%)	17 (15.6%)	0 (0%)	Mar-12, May-01	109	B.1.p11	9	0.0514
UK9	93 (100.0%)	0 (0%)	0 (0%)	Mar-19, May-03	93	B.1.13	7	0.0699
UK494	68 (100.0%)	0 (0%)	0 (0%)	Mar-21, May-01	68	B.1.p11	9	0.068

Lineage name	England	Scotland	Northern Ireland	Date range	Total sequences	Global lineage	Time since last sample (days)	Activity score
UK19	59 (100.0%)	0 (0%)	0 (0%)	Mar-14, May-10	59	B.1	0	active today
UK177	55 (100.0%)	0 (0%)	0 (0%)	Mar-27, May-01	55	B.1.1	9	0.072
UK115	53 (100.0%)	0 (0%)	0 (0%)	Mar-17, Apr-20	53	B.2.1	20	0.0327
UK66	52 (100.0%)	0 (0%)	0 (0%)	Mar-18, May-01	52	B.1.1.8	9	0.0959
UK31	51 (100.0%)	0 (0%)	0 (0%)	Mar-21, May-08	51	B.1	2	0.48
UK701	44 (97.78%)	0 (0%)	1 (2.22%)	Feb-27, May-10	45	B.1	0	active today

These data is represented in the figure one. Note that the number of sequences is likely to be due more to differing sampling efforts in different regions, rather than genuine differences in numbers of cases.

The raw data for this bar chart are in the table above.

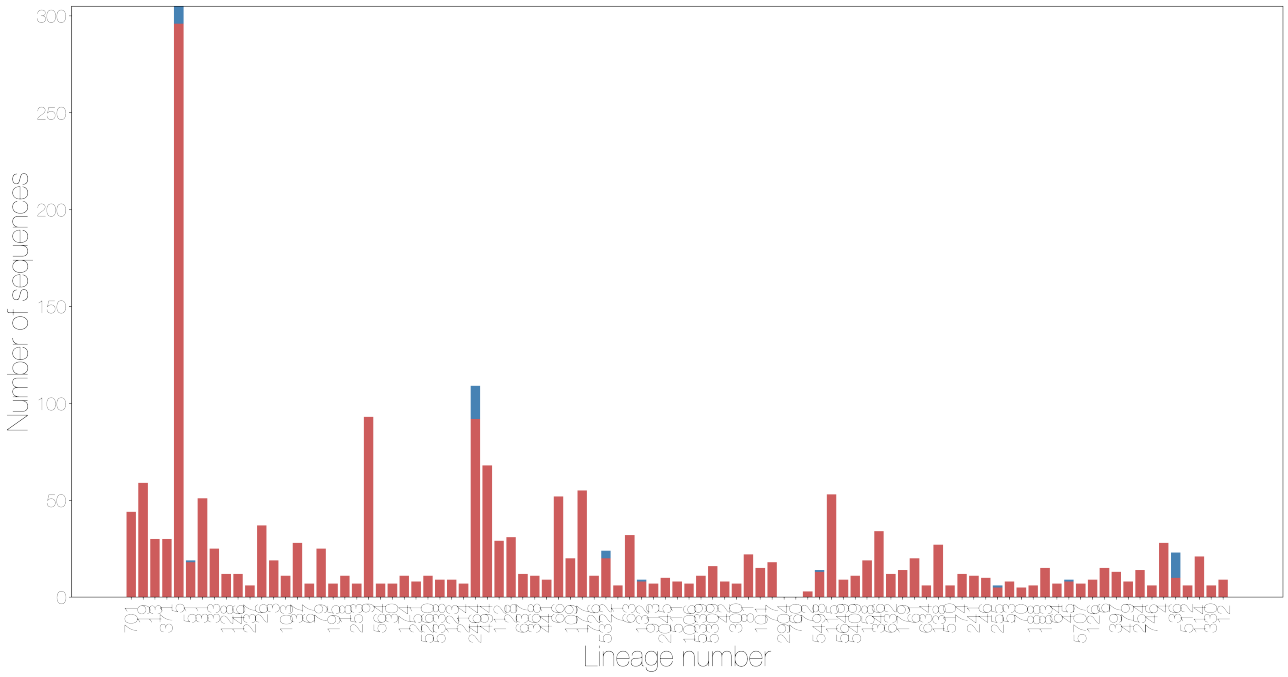


Figure 1: Number of sequences sampled in a lineage by country

Different sequencing centres have different delays in turn around from receipt of samples to submission of sequence data. This will affect all of the figures shown after this if lineages have geographical variation, as some regions have less up to date data.

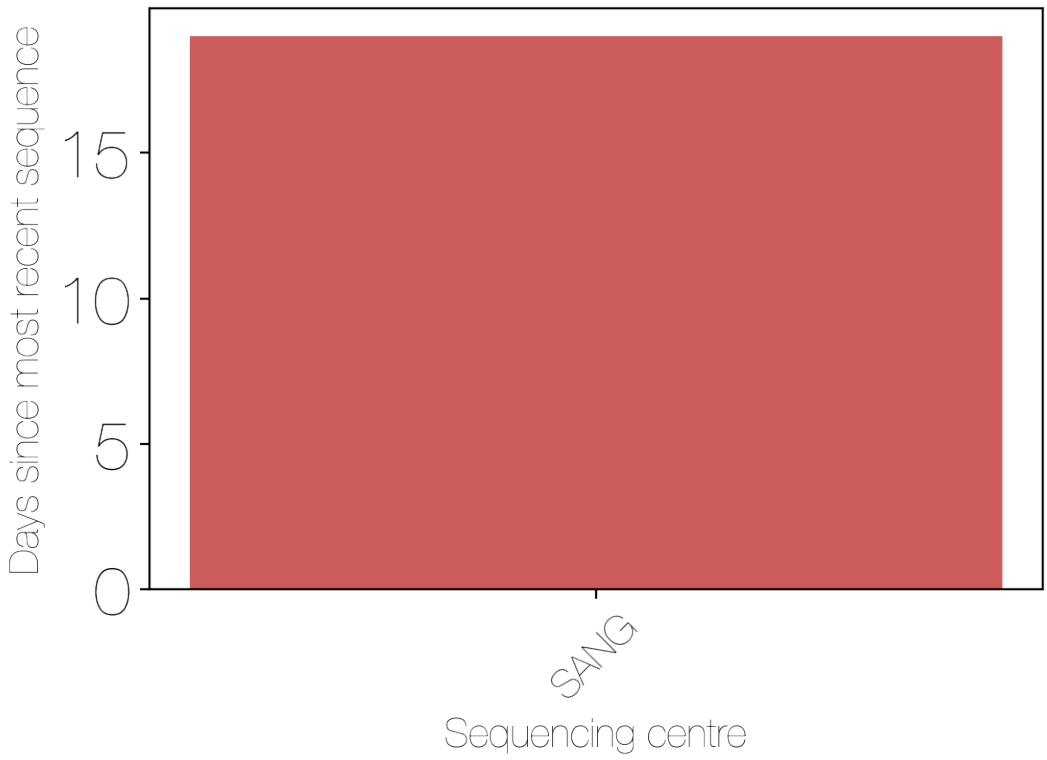


Figure 2: Lag since the most recent sequence from each sequencing centre to most current date

The relative growth and decline of the ten most sampled lineages in terms of number of counties they are present in is shown in figure three.

These lineages are shown on the timeline. Each line represents the length of the cluster, from oldest to most recent sampling date. The dots are sized by the number of sequences taken on that date, and again are colour coded by country. The raw data has been written to a summary file.

The date of first sequence in the cluster is shown in figure five for every cluster with date information.

For comparison, here is a plot of the day that every sequence was taken, coloured by country. Note that sequences without dates were not included.

The map shows the number of sequences sampled in each admin2 region in the UK. The colour scale is the same for all four countries, but with different underlying base colours.

There are 410 sequences without enough geographical information to map from this centre.

———————————————————————————FileNotFoundError Traceback (most recent call last) in 5 input_geojsons = [uk_json, channels, NI_json] 6 —-> 7 map_output = map.make_map(input_geojsons, adm2_cleaning_file, metadata_file, output_directory, week, sequencing_centre, country) 8 9 if type(map_output) != bool: ~/anaconda3/envs/report/lib/python3.7/site-packages/UK_full_report/utils/mapping.py in make_map(input_geojs adm2_cleaning_file, metadata_file, overall_output_dir, week, sequencing_centre, country) 540 sort_missing_sequences(mis missing_sequences, sequencing_centre, country) 541 –> 542 new_unclean_locs = find_new_locs_cleaning(metadata_file, mapping_dictionary, all_uk, output_dir, sequencing_centre) 543 544 return new_unclean_locs, cleaned ~/anaconda3/envs/report/lib/python3.7/site-packages/UK_full_report/utils/mapping.py in find_new_locs_cleaning(metada mapping_dictionary, all_uk, output_dir, sequencing_centre) 426 427 new_unclean_locs = False –> 428 fw = open(output_dir + “unclean_locations.csv”, ‘w’) 429 430 for i in all_uk[“NAME_2”]: FileNotFoundError: [Errno 2] No such file or directory: ‘UK_full_report/regional_reports/results/results_SANG/summary_files/unclean_locations.csv’

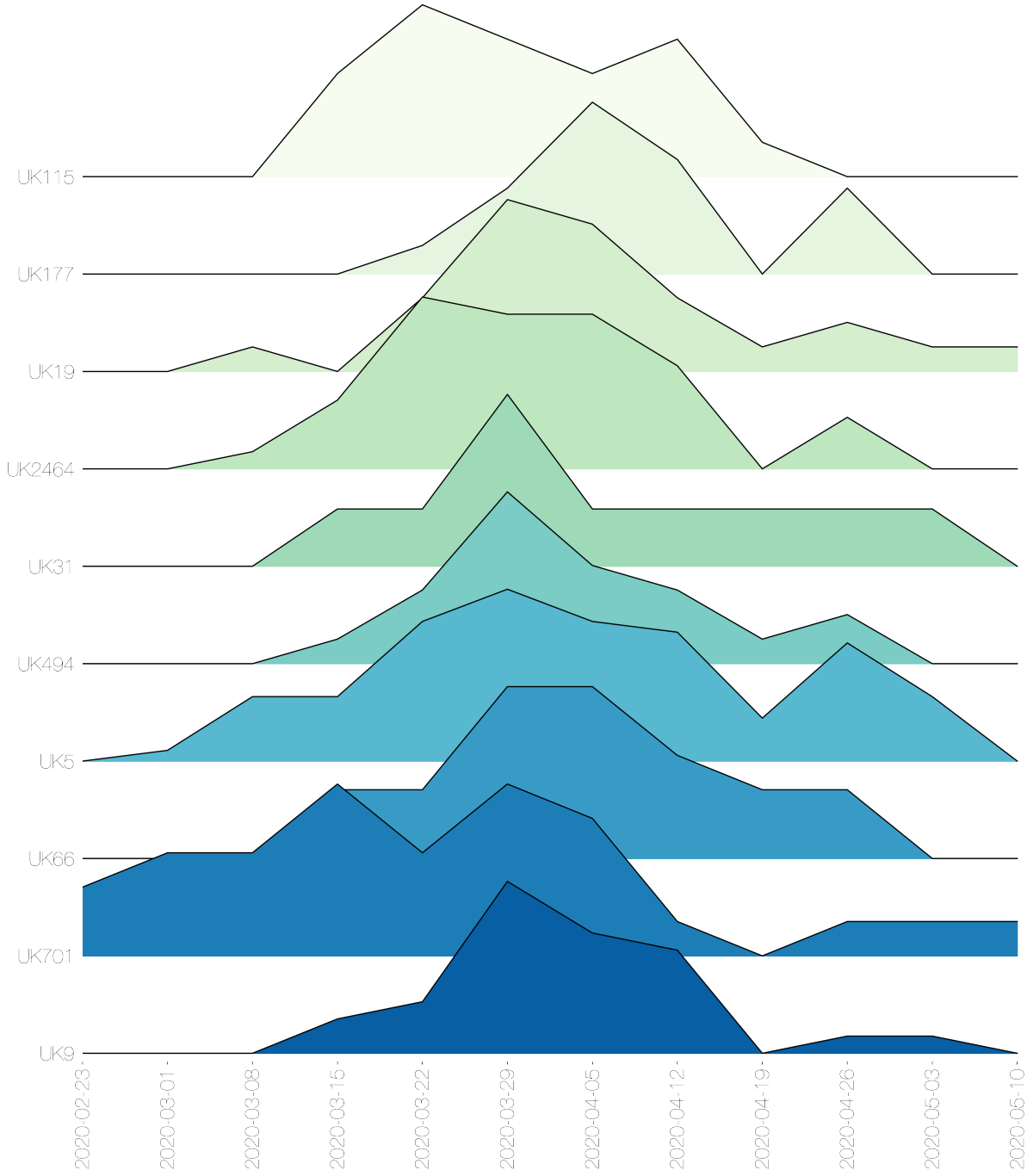


Figure 3: Lineages by number of adm2 regions present by epiweek

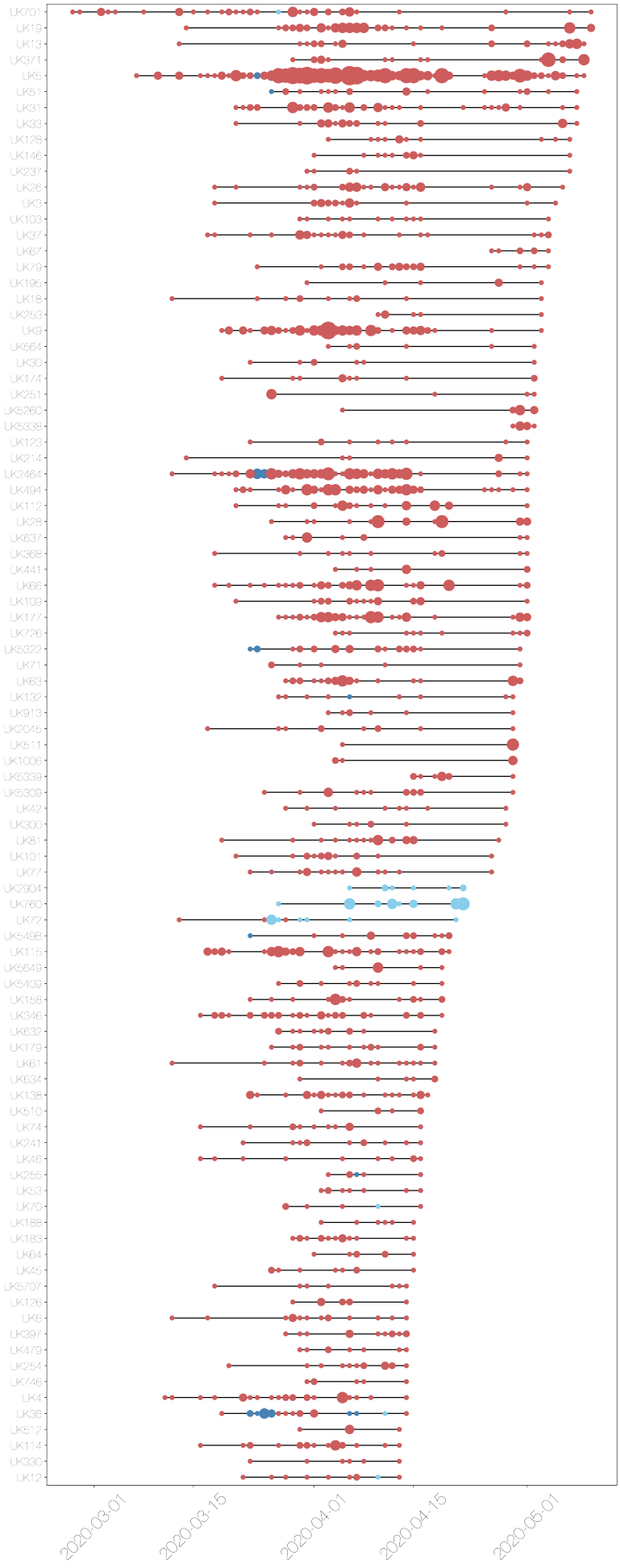


Figure 4: Timeline of lineages, sized by number of sequences from each country.

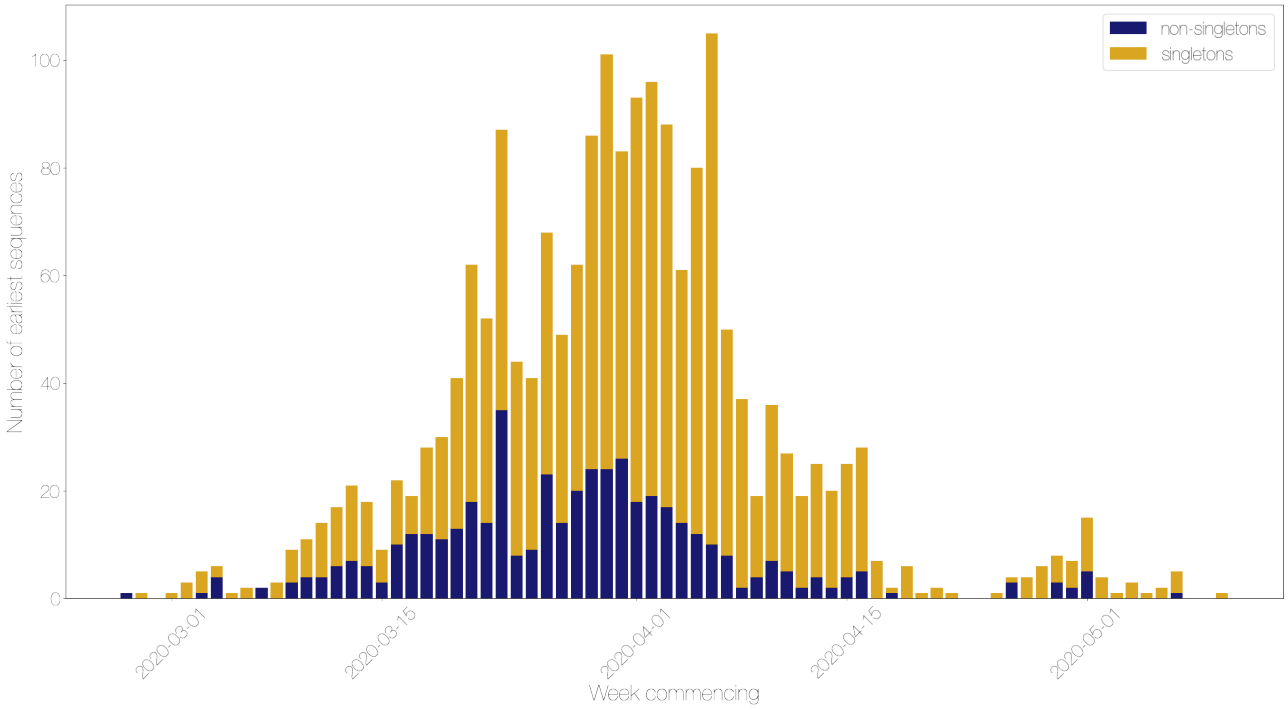


Figure 5: Lineage starts per week, split by singletons and non-singletons

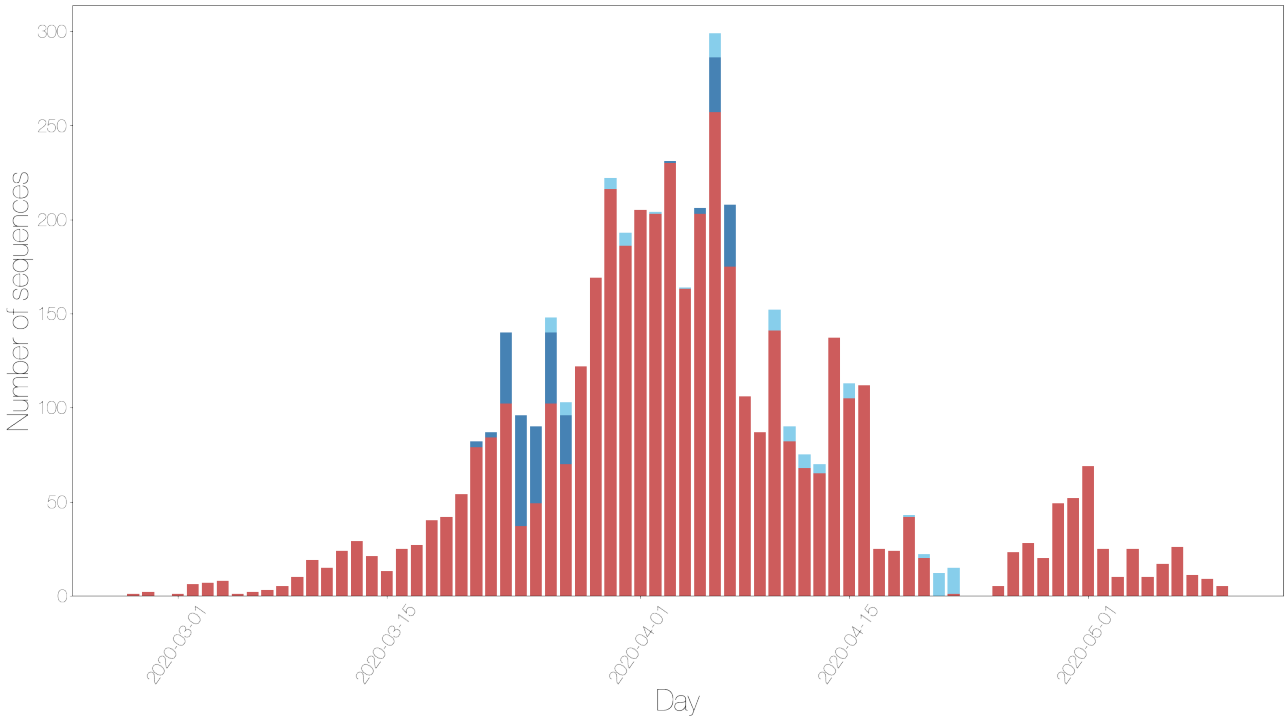
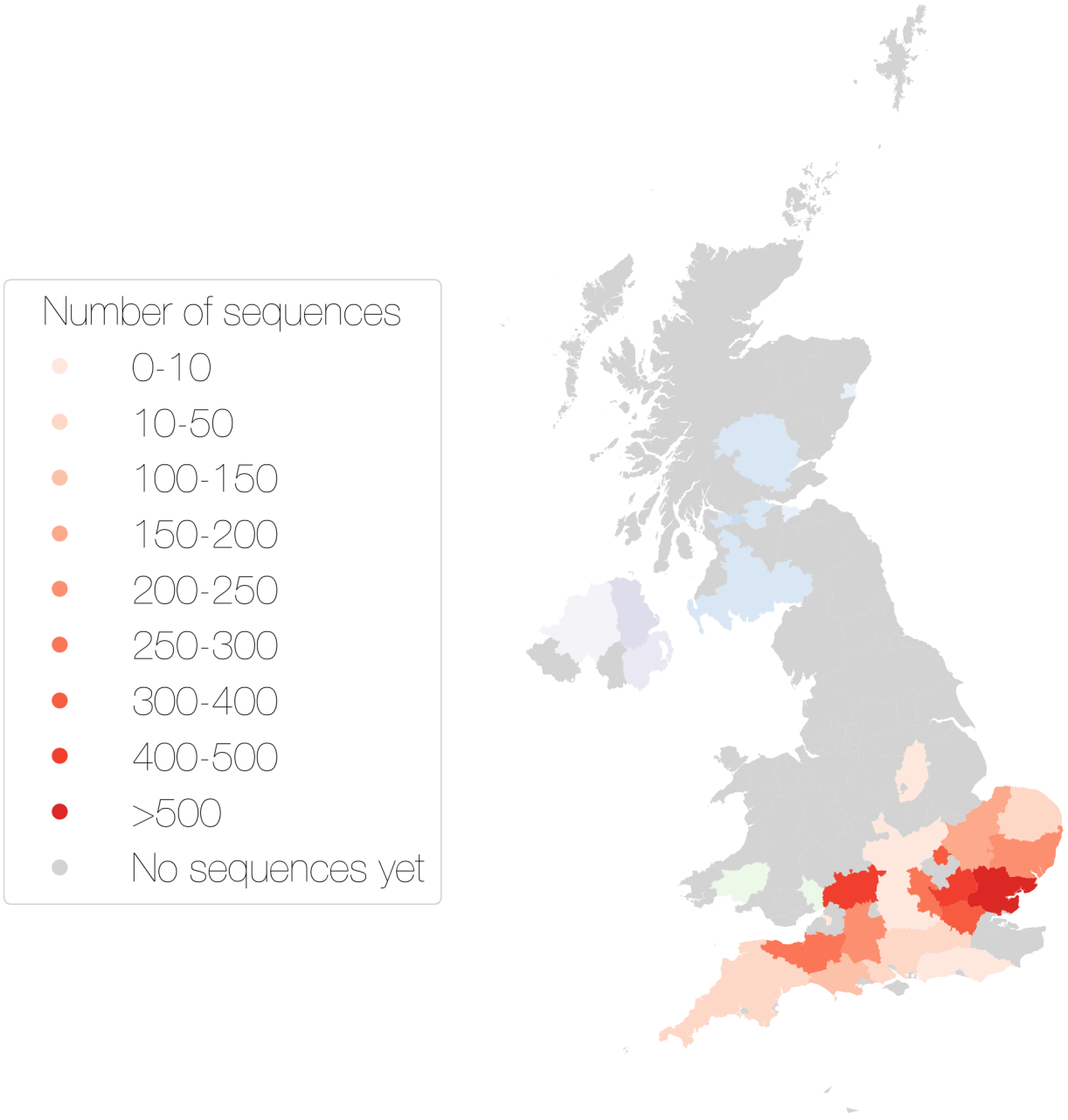


Figure 6: Sequences taken on each day by country



———————————————————————————NameError Traceback (most recent call last) in —-> 1 if not no_seqs: 2 if new_uncleans: 3 print("There are some sequences with locations that are not matched to real Admin2 regions, some manual curation required.") NameError: name 'no_seqs' is not defined

Other results modules for UK lineage analysis can be added in here if required.

Appendix

Below are the raw data tables for each of the figures in the report.

Table S1 Description of all lineages that have been circulating in the last month, and have more than 5 sequences.

Lineage name	England	Scotland	Northern Ireland	Date range	Total sequences	Global lineage	Time since last sample (days)	Activity score
UK5	296 (96.73%)	9 (2.94%)	1 (0.33%)	Mar-07, May-09	306	B.1.1, B.1.1.1	1	0.2066
UK2464	92 (84.4%)	17 (15.6%)	0 (0%)	Mar-12, May-01	109	B.1.p11	9	0.0514
UK9	93 (100.0%)	0 (0%)	0 (0%)	Mar-19, May-03	93	B.1.13	7	0.0699
UK494	68 (100.0%)	0 (0%)	0 (0%)	Mar-21, May-01	68	B.1.p11	9	0.068
UK19	59 (100.0%)	0 (0%)	0 (0%)	Mar-14, May-10	59	B.1	0	active today
UK177	55 (100.0%)	0 (0%)	0 (0%)	Mar-27, May-01	55	B.1.1	9	0.072
UK115	53 (100.0%)	0 (0%)	0 (0%)	Mar-17, Apr-20	53	B.2.1	20	0.0327
UK66	52 (100.0%)	0 (0%)	0 (0%)	Mar-18, May-01	52	B.1.1.8	9	0.0959
UK31	51 (100.0%)	0 (0%)	0 (0%)	Mar-21, May-08	51	B.1	2	0.48
UK701	44 (97.78%)	0 (0%)	1 (2.22%)	Feb-27, May-10	45	B.1	0	active today
UK26	37 (100.0%)	0 (0%)	0 (0%)	Mar-18, May-06	37	B.1.1.3	4	0.3403
UK346	34 (100.0%)	0 (0%)	0 (0%)	Mar-16, Apr-19	34	B.1.72, B.1	21	0.0491
UK760	0 (0%)	0 (0%)	32 (100.0%)	Mar-27, Apr-22	32	B.1.1	18	0.0466
UK63	32 (100.0%)	0 (0%)	0 (0%)	Mar-28, Apr-30	32	B.1.1	10	0.1065
UK28	31 (100.0%)	0 (0%)	0 (0%)	Mar-26, May-01	31	B.1.1.10	9	0.1333

Lineage name	England	Scotland	Northern Ireland	Date range	Total sequences	Global lineage	Time since last sample (days)	Activity score
UK40	0 (0%)	31 (100.0%)	0 (0%)	Mar-21, Apr-07	31	B.16	33	0.0172
UK371	30 (100.0%)	0 (0%)	0 (0%)	Mar-29, May-09	30	B.1.1	1	1.4138
UK13	30 (100.0%)	0 (0%)	0 (0%)	Mar-13, May-09	30	B.1.1	1	1.9655
UK112	29 (100.0%)	0 (0%)	0 (0%)	Mar-21, May-01	29	B.1.1	9	0.1627
UK37	28 (100.0%)	0 (0%)	0 (0%)	Mar-17, May-04	28	B.1.30	6	0.2963
UK4	28 (100.0%)	0 (0%)	0 (0%)	Mar-11, Apr-14	28	B	26	0.0484
UK138	27 (100.0%)	0 (0%)	0 (0%)	Mar-23, Apr-17	27	B.2.1	23	0.0418
UK79	25 (100.0%)	0 (0%)	0 (0%)	Mar-24, May-04	25	B.1	6	0.2847
UK33	25 (100.0%)	0 (0%)	0 (0%)	Mar-21, May-08	25	B.1.1	2	1.0
UK5322	20 (80.0%)	4 (16.0%)	1 (4.0%)	Mar-23, Apr-30	25	B.1.1	10	0.1583
UK36	10 (41.67%)	13 (54.17%)	1 (4.17%)	Mar-19, Apr-14	24	B.1	26	0.0435
UK8	24 (100.0%)	0 (0%)	0 (0%)	Mar-10, Apr-08	24	B	32	0.0394
UK81	22 (100.0%)	0 (0%)	0 (0%)	Mar-19, Apr-27	22	B.1.1	13	0.1429
UK114	21 (100.0%)	0 (0%)	0 (0%)	Mar-16, Apr-13	21	B.1.1	27	0.0519
UK61	20 (100.0%)	0 (0%)	0 (0%)	Mar-12, Apr-18	20	B.3	22	0.0885
UK109	20 (100.0%)	0 (0%)	0 (0%)	Mar-21, May-01	20	B.1.5	9	0.2398
UK158	19 (100.0%)	0 (0%)	0 (0%)	Mar-23, Apr-19	19	B.1.1	21	0.0714
UK52	1 (5.26%)	18 (94.74%)	0 (0%)	Mar-23, Apr-07	19	B.1, B.1.p73	33	0.0253
UK3	19 (100.0%)	0 (0%)	0 (0%)	Mar-18, May-05	19	B.1	5	0.5333

Lineage name	England	Scotland	Northern Ireland	Date range	Total sequences	Global lineage	Time since last sample (days)	Activity score
UK51	18 (94.74%)	1 (5.26%)	0 (0%)	Mar-26, May-08	19	B.1.36	2	1.1944
UK77	18 (100.0%)	0 (0%)	0 (0%)	Mar-23, Apr-26	18	B.2	14	0.1429
UK39	0 (0%)	18 (100.0%)	0 (0%)	Mar-24, Apr-07	18	A.2	33	0.025
UK5675	18 (100.0%)	0 (0%)	0 (0%)	Mar-22, Apr-01	18	B.2	39	0.0151
UK274	16 (100.0%)	0 (0%)	0 (0%)	Mar-15, Apr-08	16	B.3	32	0.05
UK225	0 (0%)	16 (100.0%)	0 (0%)	Mar-21, Apr-07	16	B.2	33	0.0343
UK5309	16 (100.0%)	0 (0%)	0 (0%)	Mar-25, Apr-29	16	B.1.1.10, B.1.1	11	0.2121
UK101	15 (100.0%)	0 (0%)	0 (0%)	Mar-21, Apr-26	15	B.1.5	14	0.1837
UK183	15 (100.0%)	0 (0%)	0 (0%)	Mar-29, Apr-15	15	B.1.1	25	0.0486
UK6	15 (100.0%)	0 (0%)	0 (0%)	Mar-12, Apr-14	15	B.1	26	0.0907
UK254	14 (100.0%)	0 (0%)	0 (0%)	Mar-20, Apr-14	14	B.1.1	26	0.074
UK179	14 (100.0%)	0 (0%)	0 (0%)	Mar-26, Apr-18	14	B.1.1.p11	22	0.0804
UK5498	13 (92.86%)	1 (7.14%)	0 (0%)	Mar-23, Apr-20	14	B.1.1, B.2	20	0.1077
UK173	13 (100.0%)	0 (0%)	0 (0%)	Mar-16, Apr-11	13	B	29	0.0747
UK23	12 (92.31%)	1 (7.69%)	0 (0%)	Mar-18, Apr-03	13	B.9, B	37	0.036
UK72	3 (23.08%)	0 (0%)	10 (76.92%)	Mar-13, Apr-21	13	B.10	19	0.1711
UK397	13 (100.0%)	0 (0%)	0 (0%)	Mar-28, Apr-14	13	B.1.1.13	26	0.0545
UK632	12 (100.0%)	0 (0%)	0 (0%)	Mar-27, Apr-18	12	B.1.1	22	0.0909
UK62	12 (100.0%)	0 (0%)	0 (0%)	Mar-21, Apr-09	12	B.3	31	0.0557
UK637	12 (100.0%)	0 (0%)	0 (0%)	Mar-28, May-01	12	B.1.1	9	0.3434
UK146	12 (100.0%)	0 (0%)	0 (0%)	Apr-01, May-07	12	B.1.1	3	1.0909
UK128	12 (100.0%)	0 (0%)	0 (0%)	Apr-03, May-07	12	B.1.1	3	1.0303

Lineage name	England	Scotland	Northern Ireland	Date range	Total sequences	Global lineage	Time since last sample (days)	Activity score
UK74	12 (100.0%)	0 (0%)	0 (0%)	Mar-16, Apr-16	12	B.1	24	0.1174
UK5339	11 (100.0%)	0 (0%)	0 (0%)	Apr-15, Apr-29	11	B.1.1	11	0.1273
UK5260	11 (100.0%)	0 (0%)	0 (0%)	Apr-05, May-02	11	B.1.1	8	0.3375
UK103	11 (100.0%)	0 (0%)	0 (0%)	Mar-30, May-04	11	B.1.1	6	0.5833
UK18	11 (100.0%)	0 (0%)	0 (0%)	Mar-12, May-03	11	B.1.1.7	7	0.7429
UK174	11 (100.0%)	0 (0%)	0 (0%)	Mar-19, May-02	11	B.1.5	8	0.55
UK726	11 (100.0%)	0 (0%)	0 (0%)	Apr-04, May-01	11	B.1	9	0.3
UK368	11 (100.0%)	0 (0%)	0 (0%)	Mar-18, May-01	11	B.1	9	0.4889
UK241	11 (100.0%)	0 (0%)	0 (0%)	Mar-22, Apr-16	11	B.1.5.3	24	0.1042
UK5409	11 (100.0%)	0 (0%)	0 (0%)	Mar-27, Apr-19	11	B.1.1	21	0.1095
UK2045	10 (100.0%)	0 (0%)	0 (0%)	Mar-17, Apr-29	10	B.1, B	11	0.4343
UK513	10 (100.0%)	0 (0%)	0 (0%)	Apr-03, Apr-11	10	B.1.p11	29	0.0307
UK46	10 (100.0%)	0 (0%)	0 (0%)	Mar-16, Apr-16	10	B.2.1	24	0.1435
UK12	9 (90.0%)	0 (0%)	1 (10.0%)	Mar-22, Apr-13	10	B.1.p11	27	0.0905
UK354	10 (100.0%)	0 (0%)	0 (0%)	Mar-28, Apr-07	10	B.1.1	33	0.0337
UK88	0 (0%)	10 (100.0%)	0 (0%)	Mar-23, Apr-07	10	B.1	33	0.0505
UK134	10 (100.0%)	0 (0%)	0 (0%)	Mar-09, Apr-07	10	B.1	33	0.0976
UK444	9 (100.0%)	0 (0%)	0 (0%)	Mar-31, Apr-09	9	B.1.1	31	0.0363
UK569	9 (100.0%)	0 (0%)	0 (0%)	Mar-23, Apr-10	9	B.1.1	30	0.075
UK5649	9 (100.0%)	0 (0%)	0 (0%)	Apr-04, Apr-19	9	B.2.6	21	0.0893

Lineage name	England	Scotland	Northern Ireland	Date range	Total sequences	Global lineage	Time since last sample (days)	Activity score
UK5338	9 (100.0%)	0 (0%)	0 (0%)	Apr-29, May-02	9	B.1.1	8	0.0469
UK441	9 (100.0%)	0 (0%)	0 (0%)	Apr-04, May-01	9	B.1.1	9	0.375
UK45	8 (88.89%)	1 (11.11%)	0 (0%)	Mar-26, Apr-15	9	B.1.1	25	0.1
UK291	9 (100.0%)	0 (0%)	0 (0%)	Mar-13, Apr-05	9	B.2.1	35	0.0821
UK95	9 (100.0%)	0 (0%)	0 (0%)	Mar-17, Apr-05	9	B.2.1	35	0.0679
UK2200	3 (33.33%)	6 (66.67%)	0 (0%)	Mar-20, Apr-07	9	B.1.5.6, B.1.5	33	0.0682
UK132	8 (88.89%)	1 (11.11%)	0 (0%)	Mar-27, Apr-29	9	B.1	11	0.375
UK126	9 (100.0%)	0 (0%)	0 (0%)	Mar-29, Apr-14	9	B.1.1	26	0.0769
UK58	6 (66.67%)	3 (33.33%)	0 (0%)	Mar-17, Apr-09	9	B.1	31	0.0927
UK123	9 (100.0%)	0 (0%)	0 (0%)	Mar-23, May-01	9	B.1	9	0.5417
UK163	8 (100.0%)	0 (0%)	0 (0%)	Mar-27, Apr-07	8	B.1.1	33	0.0476
UK53	8 (100.0%)	0 (0%)	0 (0%)	Apr-02, Apr-16	8	B.1.1.4	24	0.0833
UK42	8 (100.0%)	0 (0%)	0 (0%)	Mar-28, Apr-28	8	B.1, B.1.35	12	0.369
UK479	8 (100.0%)	0 (0%)	0 (0%)	Mar-30, Apr-14	8	B.1.1	26	0.0824
UK341	8 (100.0%)	0 (0%)	0 (0%)	Mar-23, Apr-12	8	B.1	28	0.102
UK86	8 (100.0%)	0 (0%)	0 (0%)	Mar-23, Mar-30	8	B.1	41	0.0244
UK2904	0 (0%)	0 (0%)	8 (100.0%)	Apr-06, Apr-22	8	B.1.p11	18	0.127
UK645	8 (100.0%)	0 (0%)	0 (0%)	Mar-29, Apr-08	8	B.2.1	32	0.0446
UK318	8 (100.0%)	0 (0%)	0 (0%)	Mar-20, Apr-10	8	B	30	0.1
UK511	8 (100.0%)	0 (0%)	0 (0%)	Apr-05, Apr-29	8	B.1.1	11	0.3117
UK251	8 (100.0%)	0 (0%)	0 (0%)	Mar-26, May-02	8	B.1.1	8	0.6607
UK759	8 (100.0%)	0 (0%)	0 (0%)	Mar-28, Apr-04	8	B.1.1	36	0.0278

Lineage name	England	Scotland	Northern Ireland	Date range	Total sequences	Global lineage	Time since last sample (days)	Activity score
UK214	7 (100.0%)	0 (0%)	0 (0%)	Mar-14, May-01	7	B.1.1	9	0.8889
UK253	7 (100.0%)	0 (0%)	0 (0%)	Apr-10, May-03	7	B.1.1	7	0.5476
UK276	7 (100.0%)	0 (0%)	0 (0%)	Mar-18, Apr-06	7	B.1.1	34	0.0931
UK195	7 (100.0%)	0 (0%)	0 (0%)	Mar-31, May-03	7	B.1.1	7	0.7857
UK119	7 (100.0%)	0 (0%)	0 (0%)	Mar-23, Apr-07	7	B.2.5	33	0.0758
UK1006	7 (100.0%)	0 (0%)	0 (0%)	Apr-04, Apr-29	7	B.1.1	11	0.3788
UK5707	7 (100.0%)	0 (0%)	0 (0%)	Mar-18, Apr-14	7	B.2	26	0.1731
UK913	7 (100.0%)	0 (0%)	0 (0%)	Apr-03, Apr-29	7	B.1	11	0.3939
UK2240	7 (100.0%)	0 (0%)	0 (0%)	Mar-18, Apr-06	7	B.1	34	0.0931
UK5174	6 (85.71%)	0 (0%)	1 (14.29%)	Mar-26, Apr-07	7	B.1.1.7	33	0.0606
UK64	7 (100.0%)	0 (0%)	0 (0%)	Apr-01, Apr-15	7	B.1	25	0.0933
UK30	7 (100.0%)	0 (0%)	0 (0%)	Mar-23, May-02	7	B.1.1	8	0.8333
UK14	4 (57.14%)	3 (42.86%)	0 (0%)	Mar-04, Apr-01	7	B	39	0.1197
UK564	7 (100.0%)	0 (0%)	0 (0%)	Apr-03, May-02	7	B.1.1	8	0.6042
UK300	7 (100.0%)	0 (0%)	0 (0%)	Apr-01, Apr-28	7	B.1.1	12	0.375
UK67	7 (100.0%)	0 (0%)	0 (0%)	Apr-26, May-04	7	B.1.1	6	0.2222
UK5672	6 (100.0%)	0 (0%)	0 (0%)	Mar-21, Apr-04	6	B.2	36	0.0778
UK541	6 (100.0%)	0 (0%)	0 (0%)	Apr-01, Apr-12	6	B.1.1	28	0.0786
UK287	6 (100.0%)	0 (0%)	0 (0%)	Mar-31, Apr-06	6	B.1	34	0.0353
UK71	6 (100.0%)	0 (0%)	0 (0%)	Mar-26, Apr-30	6	B	10	0.7
UK512	6 (100.0%)	0 (0%)	0 (0%)	Mar-30, Apr-13	6	B.1.1	27	0.1037

Lineage name	England	Scotland	Northern Ireland	Date range	Total sequences	Global lineage	Time since last sample (days)	Activity score
UK188	6 (100.0%)	0 (0%)	0 (0%)	Apr-02, Apr-15	6	B.1	25	0.104
UK510	6 (100.0%)	0 (0%)	0 (0%)	Apr-02, Apr-16	6	B.1.1	24	0.1167
UK70	5 (83.33%)	0 (0%)	1 (16.67%)	Mar-28, Apr-16	6	B.2	24	0.1583
UK255	5 (83.33%)	1 (16.67%)	0 (0%)	Apr-03, Apr-16	6	B.1.1	24	0.1083
UK659	6 (100.0%)	0 (0%)	0 (0%)	Mar-21, Mar-30	6	B	41	0.0439
UK330	6 (100.0%)	0 (0%)	0 (0%)	Mar-23, Apr-13	6	B.1.1	27	0.1556
UK746	6 (100.0%)	0 (0%)	0 (0%)	Mar-31, Apr-14	6	B.1.5	26	0.1077
UK634	6 (100.0%)	0 (0%)	0 (0%)	Mar-30, Apr-18	6	B.1.1	22	0.1727
UK716	6 (100.0%)	0 (0%)	0 (0%)	Mar-31, Apr-08	6	B.1.1	32	0.05
UK517	6 (100.0%)	0 (0%)	0 (0%)	Mar-29, Apr-12	6	B.1.1	28	0.1
UK237	6 (100.0%)	0 (0%)	0 (0%)	Mar-31, May-07	6	B.1.1	3	2.4667

Table S2 Raw data for figure three showing the number of admin2 regions a lineage is present in over time

Week commencing	UK5	UK2464	UK9	UK494	UK19	UK177	UK115	UK66	UK31	UK701
2020-02-23	0	0	0	0	0	0	0	0	0	2
2020-03-01	1	0	0	0	0	0	0	0	0	3
2020-03-08	6	1	0	0	1	0	0	0	0	3
2020-03-15	6	4	2	1	0	0	3	2	1	5
2020-03-22	13	10	3	3	3	1	5	2	1	3
2020-03-29	16	9	10	7	7	3	4	5	3	5
2020-04-05	13	9	7	4	6	6	3	5	1	4
2020-04-12	12	6	6	3	3	4	4	3	1	1
2020-04-19	4	0	0	1	1	0	1	2	1	0
2020-04-26	11	3	1	2	2	3	0	2	1	1
2020-05-03	6	0	1	0	1	0	0	0	1	1
2020-05-10	0	0	0	0	1	0	0	0	0	1

Table S3 is not appropriate for this report and so has been omitted.

Table S4 Raw data for figure six showing when lineages started per day, divided by singletons and non-singletons

Day	Number of singleton starts	Number of non-singleton starts	Total
2020-02-27	0	1	1
2020-02-28	1	0	1
2020-03-01	1	0	1
2020-03-02	3	0	3
2020-03-03	4	1	5
2020-03-04	2	4	6
2020-03-05	1	0	1
2020-03-06	2	0	2
2020-03-07	0	2	2
2020-03-08	3	0	3
2020-03-09	6	3	9
2020-03-10	7	4	11
2020-03-11	10	4	14
2020-03-12	11	6	17
2020-03-13	14	7	21
2020-03-14	12	6	18
2020-03-15	6	3	9
2020-03-16	12	10	22
2020-03-17	7	12	19
2020-03-18	16	12	28
2020-03-19	19	11	30
2020-03-20	28	13	41
2020-03-21	44	18	62
2020-03-22	38	14	52
2020-03-23	52	35	87
2020-03-24	36	8	44
2020-03-25	32	9	41
2020-03-26	45	23	68
2020-03-27	35	14	49
2020-03-28	42	20	62
2020-03-29	62	24	86
2020-03-30	77	24	101
2020-03-31	57	26	83
2020-04-01	75	18	93
2020-04-02	77	19	96
2020-04-03	71	17	88
2020-04-04	47	14	61
2020-04-05	68	12	80
2020-04-06	95	10	105
2020-04-07	42	8	50
2020-04-08	35	2	37
2020-04-09	15	4	19
2020-04-10	29	7	36
2020-04-11	22	5	27
2020-04-12	17	2	19
2020-04-13	21	4	25

Day	Number of singleton starts	Number of non-singleton starts	Total
2020-04-14	18	2	20
2020-04-15	21	4	25
2020-04-16	23	5	28
2020-04-17	7	0	7
2020-04-18	1	1	2
2020-04-19	6	0	6
2020-04-20	1	0	1
2020-04-21	2	0	2
2020-04-22	1	0	1
2020-04-25	1	0	1
2020-04-26	1	3	4
2020-04-27	4	0	4
2020-04-28	6	0	6
2020-04-29	5	3	8
2020-04-30	5	2	7
2020-05-01	10	5	15
2020-05-02	4	0	4
2020-05-03	1	0	1
2020-05-04	3	0	3
2020-05-05	1	0	1
2020-05-06	2	0	2
2020-05-07	4	1	5
2020-05-10	1	0	1

Table S5 Raw data for figure seven showing the number of sequences taken over time.

Day	England	Scotland	Northern Ireland
2020-02-27	1	0	0
2020-02-28	2	0	0
2020-03-01	1	0	0
2020-03-02	6	0	0
2020-03-03	7	0	0
2020-03-04	8	0	0
2020-03-05	1	0	0
2020-03-06	2	0	0
2020-03-07	3	0	0
2020-03-08	5	0	0
2020-03-09	10	0	0
2020-03-10	19	0	0
2020-03-11	15	0	0
2020-03-12	24	0	0
2020-03-13	29	0	0
2020-03-14	21	0	0
2020-03-15	13	0	0
2020-03-16	25	0	0
2020-03-17	27	0	0
2020-03-18	40	0	0
2020-03-19	42	0	0
2020-03-20	54	0	0
2020-03-21	79	3	0
2020-03-22	84	3	0
2020-03-23	102	38	0
2020-03-24	37	59	0
2020-03-25	49	41	0
2020-03-26	102	38	8
2020-03-27	70	26	7
2020-03-28	122	0	0
2020-03-29	169	0	0
2020-03-30	216	0	6
2020-03-31	186	0	7
2020-04-01	205	0	0
2020-04-02	203	0	1
2020-04-03	230	1	0
2020-04-04	163	0	1
2020-04-05	203	3	0
2020-04-06	257	29	13
2020-04-07	175	33	0
2020-04-08	106	0	0
2020-04-09	87	0	0
2020-04-10	141	0	11
2020-04-11	82	0	8
2020-04-12	68	0	7
2020-04-13	65	0	5
2020-04-14	137	0	0

Day	England	Scotland	Northern Ireland
2020-04-15	105	0	8
2020-04-16	112	0	0
2020-04-17	25	0	0
2020-04-18	24	0	0
2020-04-19	42	0	1
2020-04-20	20	0	2
2020-04-21	0	0	12
2020-04-22	1	0	14
2020-04-25	5	0	0
2020-04-26	23	0	0
2020-04-27	28	0	0
2020-04-28	20	0	0
2020-04-29	49	0	0
2020-04-30	52	0	0
2020-05-01	69	0	0
2020-05-02	25	0	0
2020-05-03	10	0	0
2020-05-04	25	0	0
2020-05-05	10	0	0
2020-05-06	17	0	0
2020-05-07	26	0	0
2020-05-08	11	0	0
2020-05-09	9	0	0
2020-05-10	5	0	0

Table S6 Raw data for the map with the number of sequences assigned to each admin2 region.

———————————————————————————NameError Traceback (most recent call last) in —-> 1 if not no_seqs: 2 print(mapping_data.to_markdown()) NameError: name ‘no_seqs’ is not defined

```
---------------------------------------------------------------------------FileNotFoundError
Traceback (most recent call last)<ipython-input-1-c2b516fe2325> in
<module>
----> 1 writing.write_summary_files(summary_output, dataframe,
omitted, week, intro_alls, timeline_df)
~/anaconda3/envs/report/lib/python3.7/site-
packages/UK_full_report/utils/writing_summary_files.py in
write_summary_files(output_dir, dataframe, omitted, week, intro_alls,
timeline_data)
     55 def write_summary_files(output_dir, dataframe, omitted, week,
intro_alls, timeline_data):
     56
---> 57     write_summary_table(dataframe, output_dir)
     58     write_omitteds(omitted, output_dir)
     59     write_singletons(intro_alls, output_dir)
~/anaconda3/envs/report/lib/python3.7/site-
packages/UK_full_report/utils/writing_summary_files.py in
write_summary_table(dataframe, output_dir)
      4 def write_summary_table(dataframe, output_dir):
      5
----> 6     dataframe.to_csv(output_dir + "/lineage_summary.tsv",
sep="\t")
      7
      8 def write_all_lins(intro_alls, output_dir):
~/anaconda3/envs/report/lib/python3.7/site-
packages/pandas/core/generic.py in to_csv(self, path_or_buf, sep,
na_rep, float_format, columns, header, index, index_label, mode,
encoding, compression, quoting, quotechar, line_terminator, chunksize,
date_format, doublequote, escapechar, decimal)
   3202             decimal=decimal,
   3203         )
-> 3204         formatter.save()
   3205
   3206         if path_or_buf is None:
~/anaconda3/envs/report/lib/python3.7/site-
packages/pandas/io/formats/csvs.py in save(self)
    186                 self.mode,
    187                 encoding=self.encoding,
--> 188                 compression=dict(self.compression_args,
method=self.compression),
    189             )
    190             close = True
~/anaconda3/envs/report/lib/python3.7/site-
packages/pandas/io/common.py in get_handle(path_or_buf, mode,
encoding, compression, memory_map, is_text)
    426         if encoding:
    427             # Encoding
--> 428             f = open(path_or_buf, mode, encoding=encoding,
newline="")
    429         elif is_text:
```

```
    430             # No explicit encoding
FileNotFoundError: [Errno 2] No such file or directory:
'UK_full_report/regional_reports/results/results_SANG/summary_files/lineage_summary.tsv'
```