

Summary report for UK introductions

This report gives summaries of UK specific lineages for week 2020-05-01. There are time lags due to batching, curation and analysis, the most recently sampled sequence is 2020-04-25. The analysis (eg time since last sample) is therefore undertaken from this date. 10725 sequences in the UK have been included in this analysis. 478 lineages have been recorded, 387 of which only contain one sequence.

Note: 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.

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

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

Of the 44 that remain: 10 are pending extinction, ie last seen three weeks ago. 13 lineages have gone quiet, ie haven't been seen this week. 2 lineages have reactivated. 19 lineages have been continuously circulating.

The following table contains information about lineages and the number of sequences in each country in the UK for each lineage, in reverse size order. 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.

It is also written to "summary_files" as "introduction_summary.tsv" for further use.

Introduction	England	Scotland	Wales	Northern Ireland	Date range	Total sequences	Global lineage	Time since last sample (days)
UK833	3306 (86.09%)	157 (4.09%)	354 (9.22%)	23 (0.6%)	Feb-15, Apr-25	3840	B.1.1, B.1.24, B.1.10, B.1	0
UK74	1632 (69.57%)	396 (16.88%)	314 (13.38%)	4 (0.17%)	Feb-03, Apr-24	2346	B.1.11, B.1.20, B.1.8, B.1, B.1.7, B.1.22, B.1.5, B.1.5.3	1
UK70	975 (92.95%)	27 (2.57%)	45 (4.29%)	2 (0.19%)	Feb-09, Apr-22	1049	B.2, B.2.1	3
UK51	422 (73.91%)	123 (21.54%)	24 (4.2%)	2 (0.35%)	Feb-03, Apr-24	571	B, B.4, B.3, B.8	1
UK65	233 (73.73%)	59 (18.67%)	23 (7.28%)	1 (0.32%)	Feb-13, Apr-22	316	B.2	3
UK15	247 (79.68%)	24 (7.74%)	37 (11.94%)	2 (0.65%)	Mar-06, Apr-24	310	B.1.3, B.1, B.1.27	1

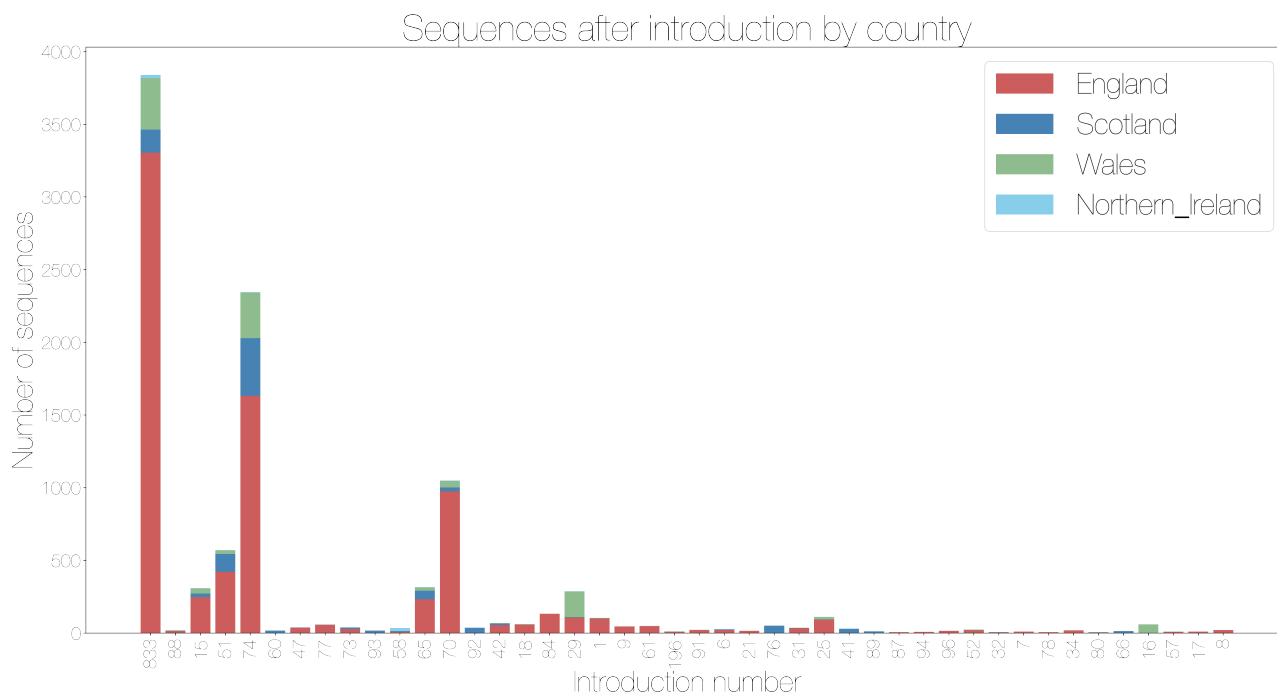
Introduction	England	Scotland	Wales	Northern Ireland	Date range	Total sequences	Global lineage	Time since last sample (days)
UK29	107 (37.28%)	3 (1.05%)	176 (61.32%)	1 (0.35%)	Mar-09, Apr-19	287	B.3	6
UK84	133 (100.0%)	0 (0%)	0 (0%)	0 (0%)	Mar-11, Apr-20	133	B.1.13	5
UK25	92 (82.88%)	2 (1.8%)	16 (14.41%)	1 (0.9%)	Feb-25, Apr-13	111	B.2.2, B.2	12
UK1	99 (95.19%)	2 (1.92%)	3 (2.88%)	0 (0%)	Feb-27, Apr-19	104	B.2.5	6
UK42	55 (80.88%)	11 (16.18%)	2 (2.94%)	0 (0%)	Feb-27, Apr-20	68	B.1	5
UK18	58 (95.08%)	0 (0%)	3 (4.92%)	0 (0%)	Mar-11, Apr-20	61	B.2.4, B.2	5
UK16	0 (0%)	0 (0%)	60 (100.0%)	0 (0%)	Mar-10, Mar-30	60	B.3	26
UK77	57 (96.61%)	2 (3.39%)	0 (0%)	0 (0%)	Mar-06, Apr-23	59	B.1	2
UK76	1 (1.96%)	50 (98.04%)	0 (0%)	0 (0%)	Mar-16, Apr-14	51	B.1	11
UK61	48 (97.96%)	0 (0%)	1 (2.04%)	0 (0%)	Mar-17, Apr-19	49	B.1.30	6
UK9	45 (100.0%)	0 (0%)	0 (0%)	0 (0%)	Feb-24, Apr-19	45	B.1	6

Introduction	England	Scotland	Wales	Northern Ireland	Date range	Total sequences	Global lineage	Time since last sample (days)
UK73	28 (73.68%)	10 (26.32%)	0 (0%)	0 (0%)	Mar- 20, Apr- 23	38	B.1	2
UK47	37 (97.37%)	1 (2.63%)	0 (0%)	0 (0%)	Mar- 12, Apr- 23	38	B.3	2
UK92	0 (0%)	36 (100.0%)	0 (0%)	0 (0%)	Mar- 22, Apr- 21	36	B.1	4
UK31	34 (94.44%)	1 (2.78%)	1 (2.78%)	0 (0%)	Mar- 06, Apr- 13	36	B, B.3	12
UK58	9 (25.71%)	5 (14.29%)	0 (0%)	21 (60.0%)	Mar- 13, Apr- 22	35	B	3
UK41	2 (6.67%)	27 (90.0%)	1 (3.33%)	0 (0%)	Mar- 12, Apr- 12	30	A.2	13
UK6	21 (77.78%)	6 (22.22%)	0 (0%)	0 (0%)	Mar- 01, Apr- 17	27	B.1	8
UK52	21 (84.0%)	0 (0%)	4 (16.0%)	0 (0%)	Mar- 11, Apr- 08	25	B	17
UK91	20 (90.91%)	0 (0%)	2 (9.09%)	0 (0%)	Mar- 12, Apr- 18	22	B.1	7
UK8	20 (95.24%)	0 (0%)	1 (4.76%)	0 (0%)	Mar- 12, Mar- 28	21	B.2.1	28
UK88	14 (73.68%)	1 (5.26%)	4 (21.05%)	0 (0%)	Mar- 15, Apr- 24	19	B.2	1

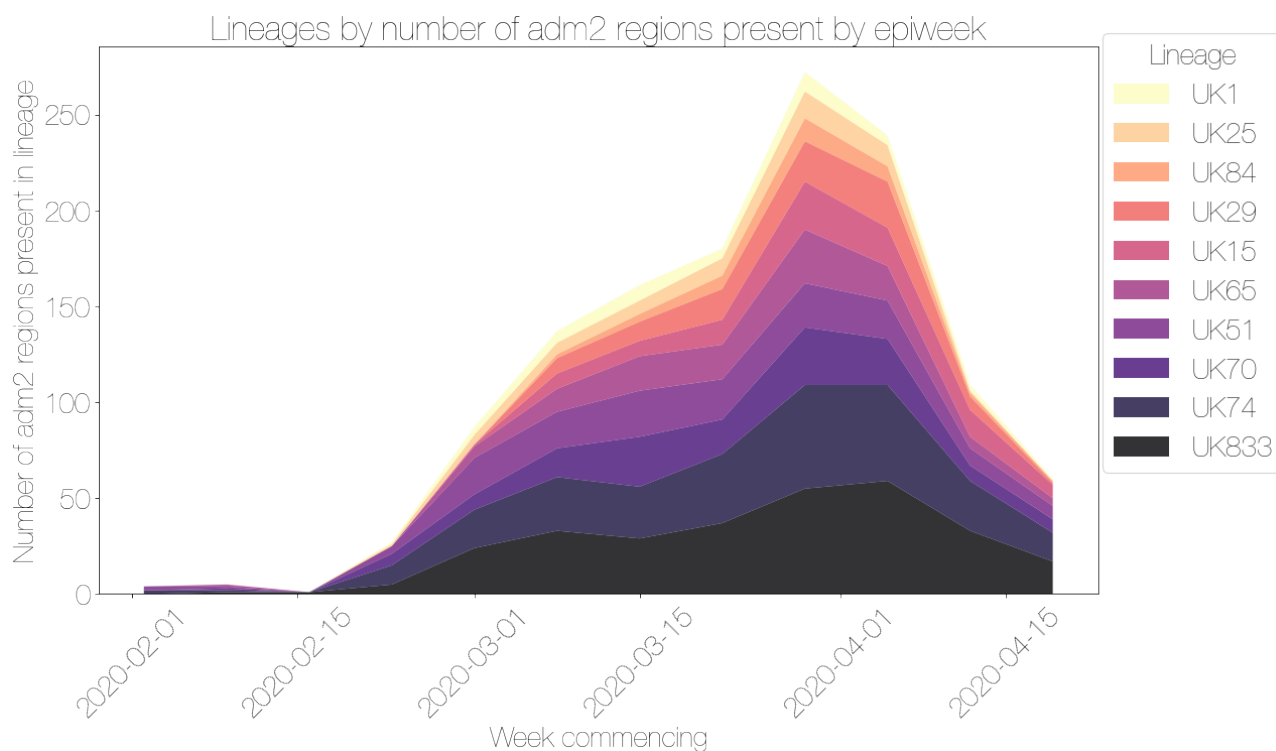
Introduction	England	Scotland	Wales	Northern Ireland	Date range	Total sequences	Global lineage	Time since last sample (days)
UK34	19 (100.0%)	0 (0%)	0 (0%)	0 (0%)	Mar-14, Apr-02	19	B.2.1	23
UK93	0 (0%)	18 (100.0%)	0 (0%)	0 (0%)	Mar-21, Apr-22	18	B.1	3
UK60	0 (0%)	17 (100.0%)	0 (0%)	0 (0%)	Mar-18, Apr-23	17	B.1	2
UK96	15 (100.0%)	0 (0%)	0 (0%)	0 (0%)	Mar-19, Apr-10	15	B.1	15
UK21	15 (100.0%)	0 (0%)	0 (0%)	0 (0%)	Mar-13, Apr-14	15	B.2	11
UK66	1 (7.14%)	13 (92.86%)	0 (0%)	0 (0%)	Mar-12, Apr-01	14	B	24
UK196	8 (66.67%)	1 (8.33%)	3 (25.0%)	0 (0%)	Mar-26, Apr-18	12	B.1	7
UK89	0 (0%)	12 (100.0%)	0 (0%)	0 (0%)	Mar-13, Apr-11	12	B.1	14
UK17	10 (90.91%)	0 (0%)	1 (9.09%)	0 (0%)	Mar-12, Mar-29	11	B.1.11	27
UK7	10 (100.0%)	0 (0%)	0 (0%)	0 (0%)	Mar-07, Apr-03	10	B.8	22
UK57	9 (90.0%)	0 (0%)	1 (10.0%)	0 (0%)	Mar-16, Mar-30	10	B.8	26

Introduction	England	Scotland	Wales	Northern Ireland	Date range	Total sequences	Global lineage	Time since last sample (days)
UK94	9 (100.0%)	0 (0%)	0 (0%)	0 (0%)	Mar-16, Apr-11	9	B.2	14
UK87	7 (100.0%)	0 (0%)	0 (0%)	0 (0%)	Mar-21, Apr-11	7	B.1	14
UK78	7 (100.0%)	0 (0%)	0 (0%)	0 (0%)	Mar-10, Apr-02	7	B.2.1	23
UK80	3 (50.0%)	2 (33.33%)	1 (16.67%)	0 (0%)	Mar-20, Apr-01	6	B.1	24
UK32	5 (83.33%)	1 (16.67%)	0 (0%)	0 (0%)	Mar-05, Apr-03	6	B.2.2	22

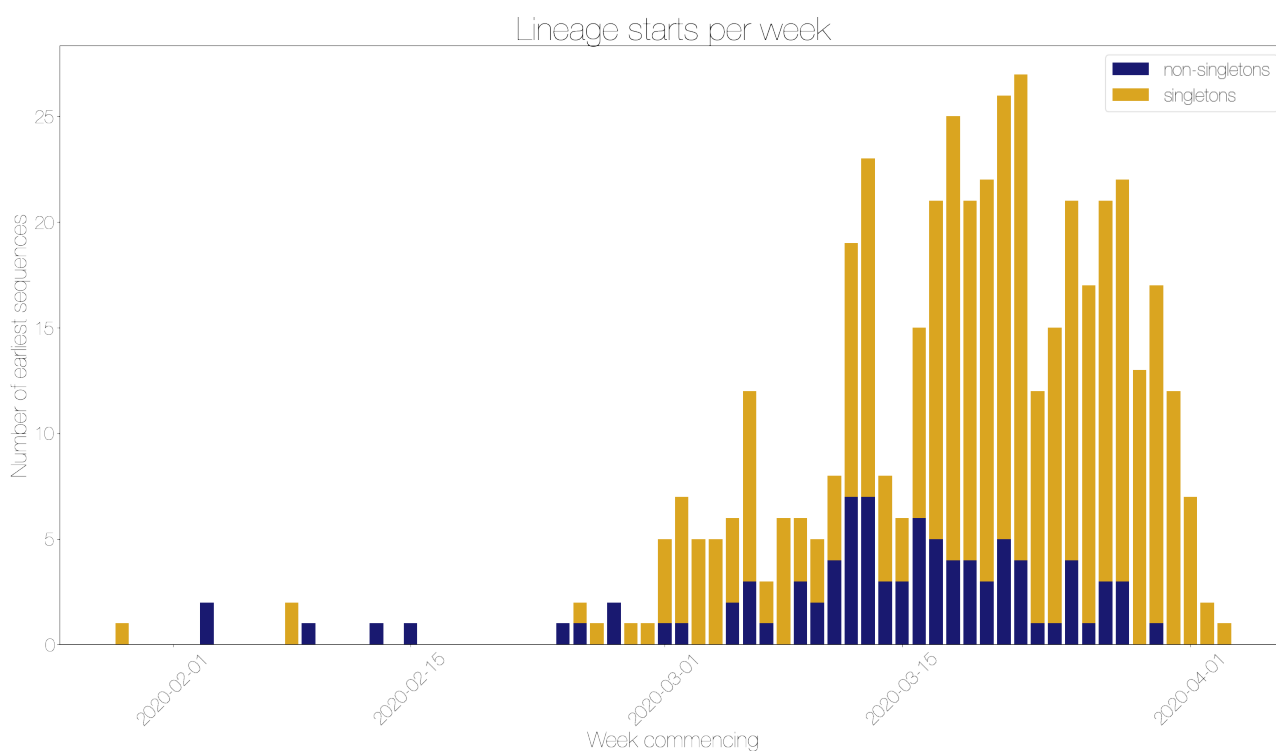
These data is represented in the stacked bar chart below. Note that the number of sequences is likely to be due more to differing sampling efforts in different regions, rather than genuine differences in numbers of cases.



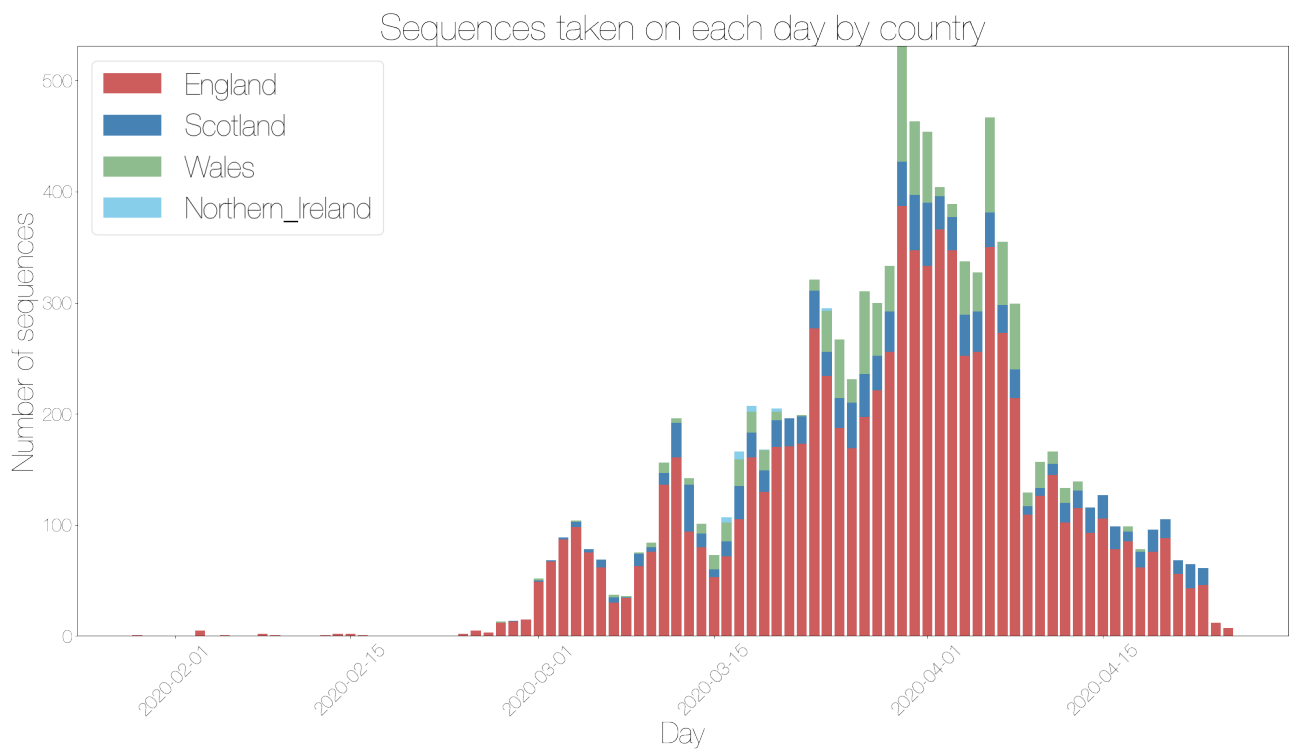
The relative growth and decline of the ten most sampled lineages in terms of number of counties they are present in is shown below.



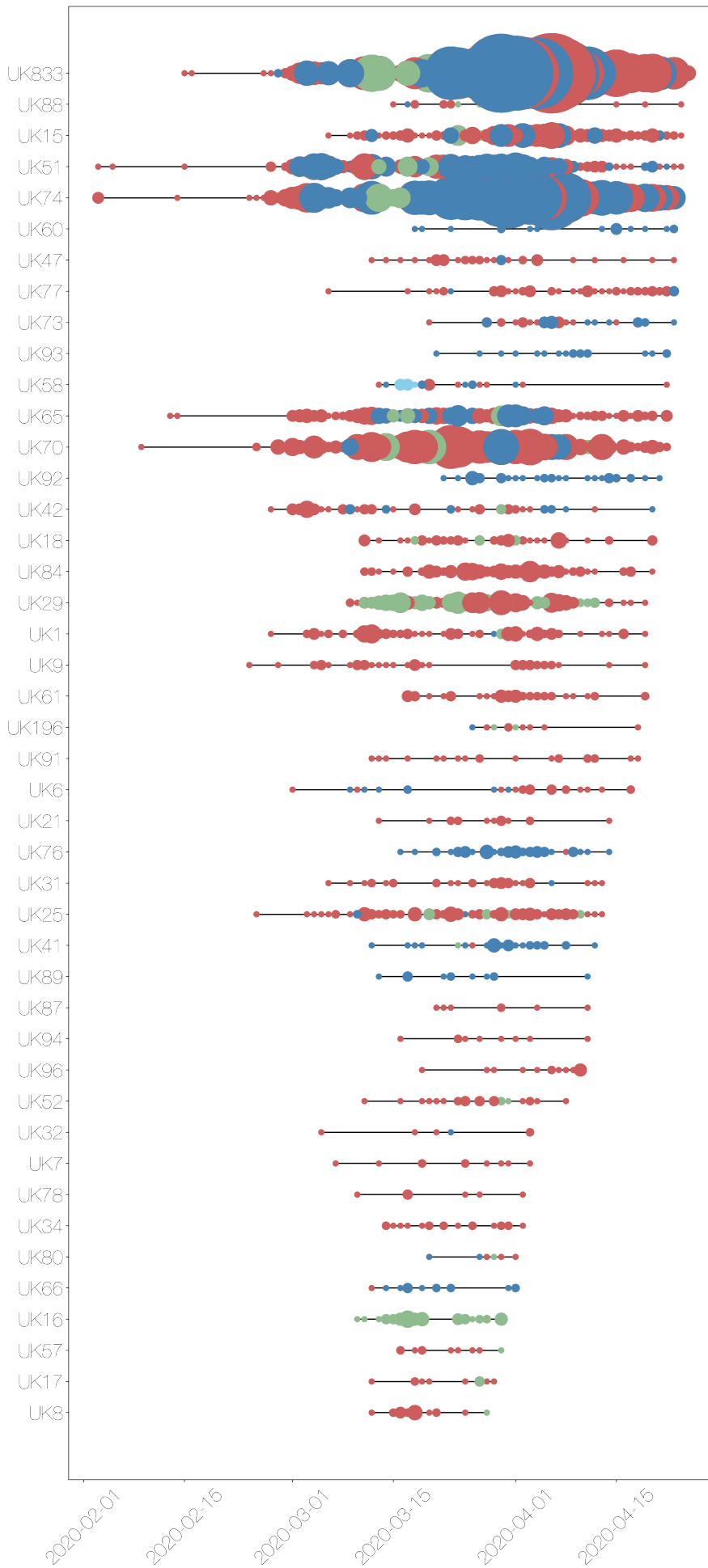
The date of first sequence in the cluster is shown below for every cluster with date information.



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.



These introductions are shown on the timeline below. 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.



Error in callback (for post_execute):

```
-----KeyboardInterrupt Traceback (most recent call last)
in 5 input_geojsons = [uk_json, channels, NI_json] 6 --> 7 new_uncleans = map.make_map(input_geojsons,
adm2_cleaning_file, metadata_file, output_directory, week) ~/anaconda3/envs/report/lib/python3.7/site-
packages/UK_full_report/utils/mapping.py in make_map(input_geojsons, adm2_cleaning_file, metadata_file,
overall_output_dir, week) 339 england, scotland, wales, n_i, channels, plot_dict = parse_countries(with_seq_counts)
340 -> 341 plot_map(england, scotland, wales, n_i, channels, plot_dict) 342 343 if missing_sequences:
~/anaconda3/envs/report/lib/python3.7/site-packages/UK_full_report/utils/mapping.py in plot_map(england,
scotland, wales, n_i, channels, plot_dict) 247 scotland = scotland.to_crs("EPSG:3395") 248 wales =
wales.to_crs("EPSG:3395") -> 249 n_i = n_i.to_crs("EPSG:3395") 250 channels = channels.to_crs("EPSG:3395")
251 ~/anaconda3/envs/report/lib/python3.7/site-packages/geopandas/geodataframe.py in to_crs(self, crs,
epsg, inplace) 562 else: 563 df = self.copy() -> 564 geom = df.geometry.to_crs(crs=crs, epsg=epsg) 565
df.geometry = geom 566 df.crs = geom.crs ~/anaconda3/envs/report/lib/python3.7/site-packages/geopandas/geoseries.py
in to_crs(self, crs, epsg) 435 436 transformer = Transformer.from_crs(self.crs, crs, always_xy=True) -> 437
result = self.apply(lambda geom: transform(transformer.transform, geom)) 438 result.class = GeoSeries 439
result.crs = crs ~/anaconda3/envs/report/lib/python3.7/site-packages/pandas/core/series.py in apply(self,
func, convert_dtype, args, **kwds) 3846 else: 3847 values = self.astype(object).values -> 3848 mapped
= lib.map_infer(values, f, convert=convert_dtype) 3849 3850 if len(mapped) and isinstance(mapped[0],
Series): pandas/_libs/lib.pyx in pandas._libs.lib.map_infer() ~/anaconda3/envs/report/lib/python3.7/site-
packages/geopandas/geoseries.py in (geom) 435 436 transformer = Transformer.from_crs(self.crs, crs,
always_xy=True) -> 437 result = self.apply(lambda geom: transform(transformer.transform, geom)) 438 re-
sult.class = GeoSeries 439 result.crs = crs ~/anaconda3/envs/report/lib/python3.7/site-packages/shapely/ops.py
in transform(func, geom) 235 elif geom.type == 'Polygon': 236 shell = type(geom.exterior)( -> 237
zip(func(izip(geom.exterior.coords)))) 238 holes = list(type(ring)(zip(func(izip(ring.coords)))) 239 for ring in
geom.interiors) KeyboardInterrupt:
```

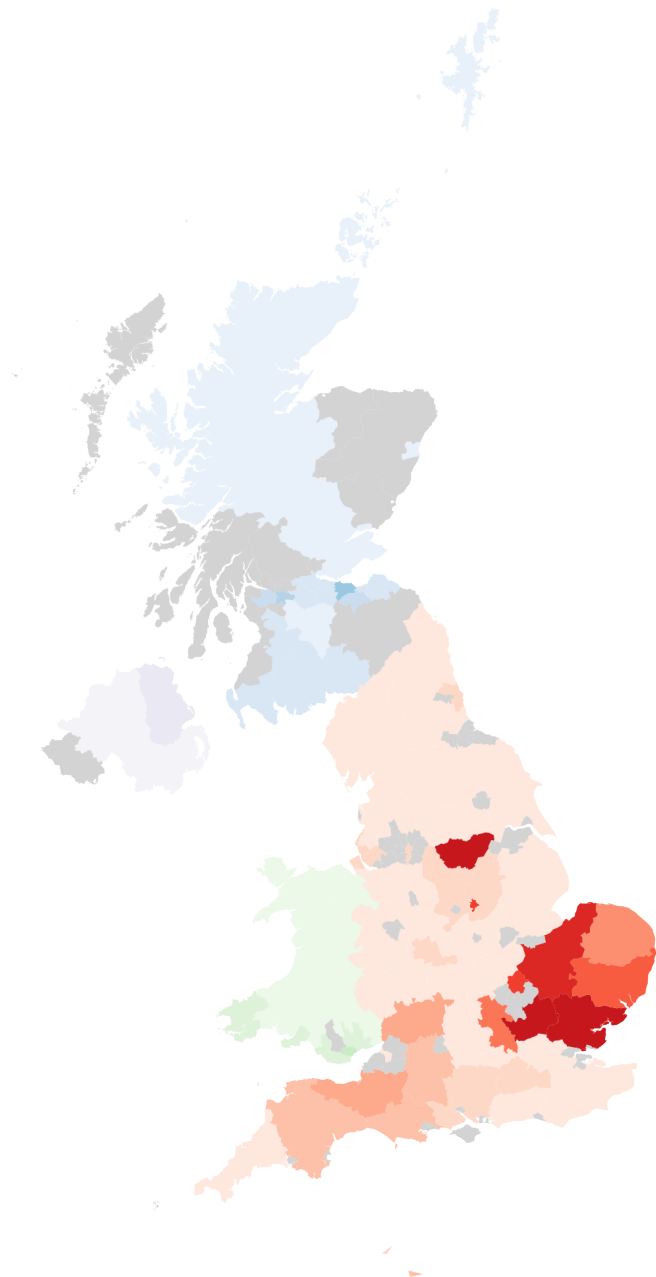
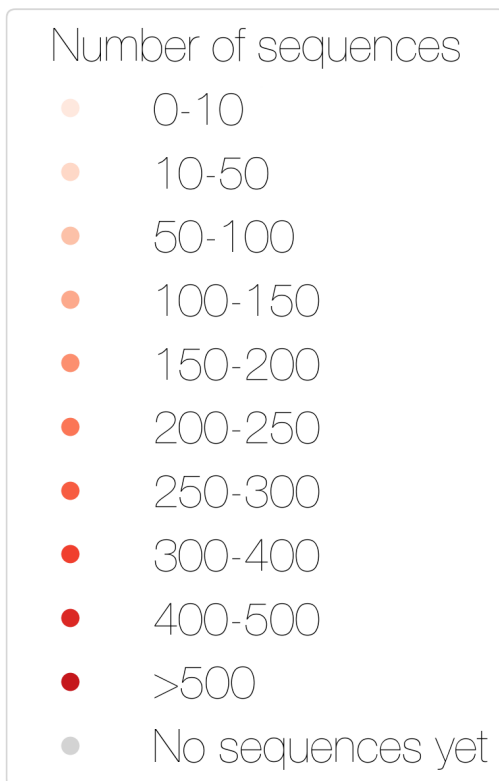
```
-----KeyboardInterrupt Traceback (most recent call
last)~/anaconda3/envs/report/lib/python3.7/site-packages/ipykernel/pylab/backend_inline.py in flush_figures()
119 # ignore the tracking, just draw and close all figures 120 try: -> 121 return show(True) 122 except Ex-
ception as e: 123 # safely show traceback if in IPython, else raise ~/anaconda3/envs/report/lib/python3.7/site-
packages/ipykernel/pylab/backend_inline.py in show(close, block) 41 display( 42 figure_manager.canvas.figure,
-> 43 metadata=_fetch_figure_metadata(figure_manager.canvas.figure) 44 ) 45 finally: ~/anaconda3/envs/report/lib/python
packages/IPython/core/display.py in display(include, exclude, metadata, transient, display_id, objs, kwargs)
311 publish_display_data(data=obj, metadata=metadata, kwargs) 312 else: -> 313 format_dict, md_dict
= format(obj, include=include, exclude=exclude) 314 if not format_dict: 315 # nothing to display (e.g.
ipython_display took over) ~/anaconda3/envs/report/lib/python3.7/site-packages/IPython/core/formatters.py
in format(self, obj, include, exclude) 178 md = None 179 try: -> 180 data = formatter(obj) 181 ex-
cept: 182 # FIXME: log the exception in call(self, obj) ~/anaconda3/envs/report/lib/python3.7/site-
packages/IPython/core/formatters.py in catch_format_error(method, self, args, **kwargs) 222 """"show
traceback on failed format call"""" 223 try: -> 224 r = method(self, *args, kwargs) 225 except NotImplementedError:
226 # don't warn on NotImplementedError ~/anaconda3/envs/report/lib/python3.7/site-
packages/IPython/core/formatters.py in call(self, obj) 339 pass 340 else: -> 341 return printer(obj)
342 # Finally look for special method names 343 method = get_real_method(obj, self.print_method)
~/anaconda3/envs/report/lib/python3.7/site-packages/IPython/core/pylabtools.py in (fig) 246 247 if
'png' in formats: -> 248 png_formatter.for_type(Figure, lambda fig: print_figure(fig, 'png', kwargs)) 249
if 'retina' in formats or 'png2x' in formats: 250 png_formatter.for_type(Figure, lambda fig: retina_figure(fig,
kwargs)) ~/anaconda3/envs/report/lib/python3.7/site-packages/IPython/core/pylabtools.py in
print_figure(fig, fmt, bbox_inches, kwargs) 130 FigureCanvasBase(fig) 131 -> 132 fig.canvas.print_figure(bytes_io,
kw) 133 data = bytes_io.getvalue() 134 if fmt == 'svg': ~/anaconda3/envs/report/lib/python3.7/site-
```

```

packages/matplotlib/backend_bases.py in print_figure(self, filename, dpi, facecolor, edgecolor, orien-
tation, format, bbox_inches, kwargs) 2103 orientation=orientation, 2104 bbox_inches_restore=_bbox_inches_restore,
-> 2105 **kwargs) 2106 finally: 2107 if bbox_inches and restore_bbox: ~/anaconda3/envs/report/lib/python3.7/site-
packages/matplotlib/backends/backend_agg.py in print_png(self, filename_or_obj, metadata, pil_kwargs,
*args, **kwargs) 535 with cbook.open_file_cm(filename_or_obj, "wb") as fh: 536 _png.write_png(renderer._renderer,
fh, self.figure.dpi, -> 537 metadata={default_metadata, metadata}) 538 539 def print_to_buffer(self): Key-
boardInterrupt:

```

COVID-19 sequences from each Admn2 region UK



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

Other results modules for UK introductions can be added in here if required.
```

Appendix

The plot below shows the number of sequences from each country that don't have specific enough location data to plot on the map.

