UK lineages summary report

This report gives summaries of lineages sampled in Scotland for week 2020-05-22. There are time lags due to batching, curation and analysis, the most recently sampled sequence is 2020-05-17. The analysis (eg time since last sample) is therefore undertaken from this date. 1992 sequences from Scotland have been included in this analysis. 571 lineages have been recorded, 439 of which only contain one sequence.

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 574 and the maximum is 940

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

523 are lineages which were sampled less than five times in Scotland, 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 48 that remain: 22 are pending extinction, ie last seen three weeks ago. 6 have not been seen for more than one month, and so are viewed as extinct, but will continue to be monitored. 9 lineages have gone quiet, ie haven't been seen this week. 3 lineages have reactivated. 8 lineages have been continuously circulating.

The following table contains information about lineages and the number of sequences the dataset, 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.

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"

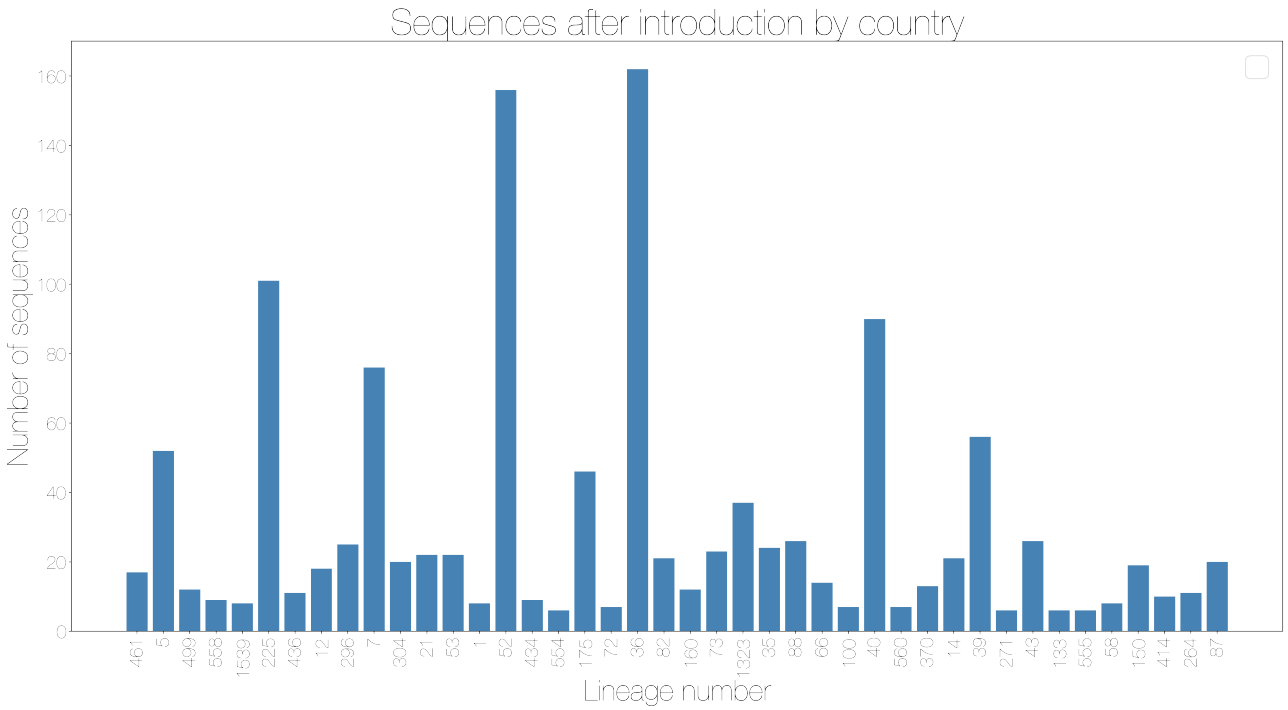
Lineage name	Date range	Number of sequences	Global lineage	Time since last sample (days)
UK36	Mar-20, May-04	162	B.1	13
UK52	Mar-01, May-08	156	B.1, B.1.p73	9
UK225	Mar-10, Apr-27	101	B.2.2, B.2, B.2.6	20
UK40	Mar-13, Apr-28	90	B, B.16	19
UK7	Mar-19, May-12	76	B.1.p11	5
UK39	Mar-12, Apr-27	56	A.2	20
UK5	Mar-13, May-06	52	B.1.1.1	11
UK175	Mar-22, May-04	46	B.1	13

Lineage name	Date range	Number of sequences	Global lineage	Time since last sample (days)
UK1323	Mar-17, May-01	37	B	16
UK43	Mar-23, Apr-26	26	A.5	21
UK88	Mar-22, Apr-29	26	B.1	18
UK296	Apr-08, May-13	25	B.1.5	4
UK35	Mar-25, May-01	24	B.1.5.6, B.1.5	16
UK73	Mar-19, May-02	23	B.1.p11	15
UK21	Mar-18, May-08	22	B.1.40	9
UK53	Apr-16, May-08	22	B.1.1.4	9
UK82	Mar-25, May-03	21	B.1.1, B.1.1.p11	14
UK14	Mar-14, Apr-27	21	B	20
UK304	Apr-16, May-12	20	B.1.1.14	5
UK87	Mar-13, Apr-20	20	B.1.70	27
UK150	Mar-21, Apr-22	19	B.1.1.p12	25
UK12	Apr-12, May-13	18	B.1.p11	4
UK461	Apr-18, May-17	17	B.1.5	0
UK502	Mar-06, Mar-20	16	B.1.69	58
UK156	Mar-18, Apr-18	14	B.1.71	29
UK66	Mar-28, Apr-28	14	B.1.1.8	19
UK370	Apr-08, Apr-27	13	B.1.1.10	20
UK499	Apr-24, May-15	12	B.1.5	2
UK160	Apr-01, May-02	12	B.1.1	15
UK261	Mar-15, Apr-08	12	A.3	39
UK436	Apr-13, May-14	11	B.1.5	3
UK264	Mar-29, Apr-22	11	B.1.p11	25

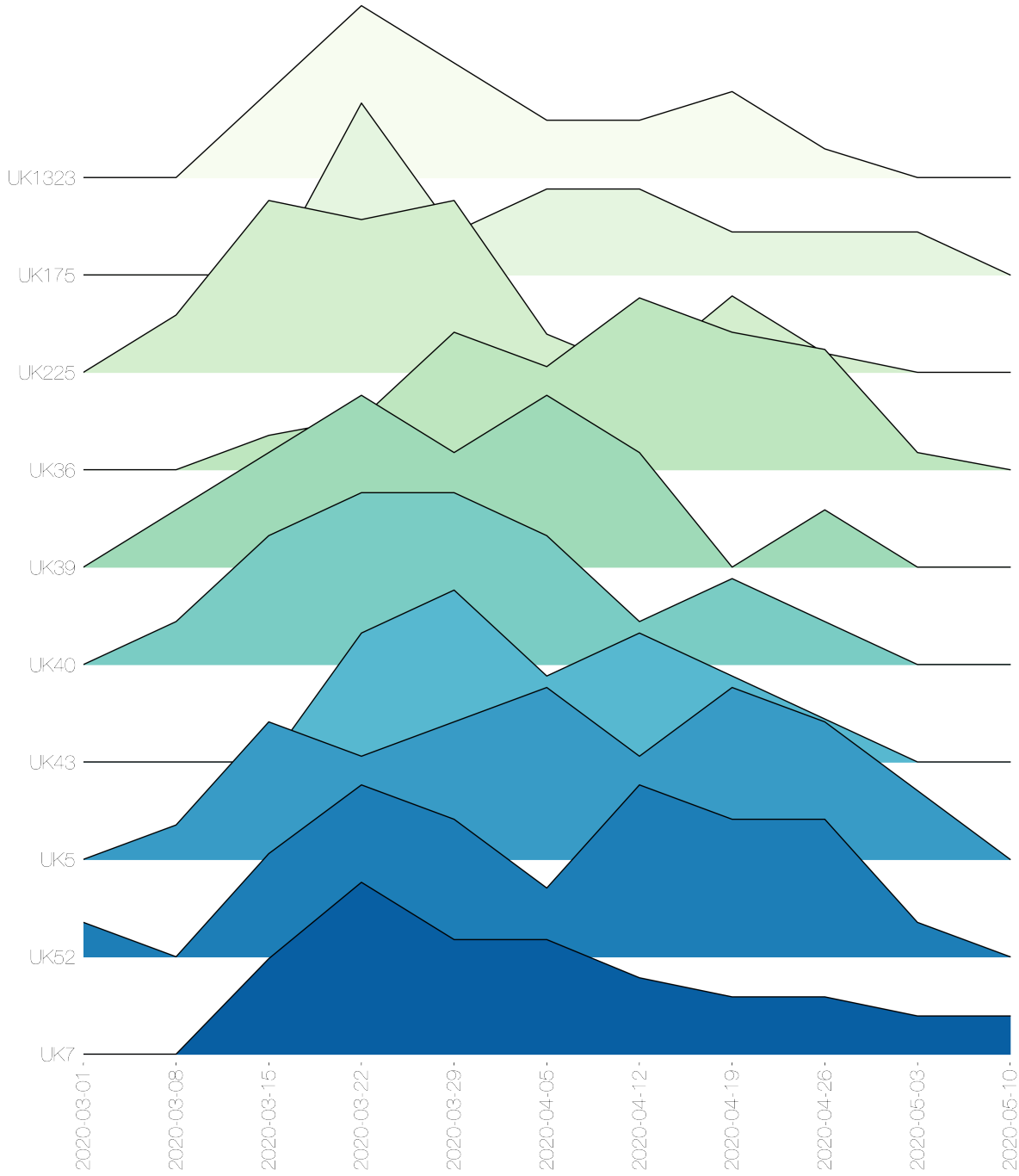
Lineage name	Date range	Number of sequences	Global lineage	Time since last sample (days)
UK414	Apr-05, Apr-22	10	B.1.5	25
UK434	Apr-20, May-06	9	B.1.5	11
UK558	Apr-24, May-15	9	B.1.5	2
UK137	Mar-10, Mar-31	8	B.1.1	47
UK1	Mar-03, Apr-08	8	B.1	39
UK1539	May-09, May-15	8	B.1.5	2
UK58	Mar-12, Apr-24	8	B.1	23
UK72	Mar-14, Apr-01	7	B.10	46
UK560	Apr-15, Apr-27	7	B.1.5	20
UK100	Apr-11, Apr-28	7	B.1.5	19
UK554	Apr-23, May-05	6	B.1.5	12
UK198	Mar-18, Apr-15	6	A, B.1.5	32
UK271	Apr-15, Apr-26	6	B.1	21
UK555	Apr-13, Apr-25	6	B.1.5	22
UK133	Mar-22, Apr-25	6	B.1	22
UK931	Mar-30, Apr-04	6	B.1.1	43

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



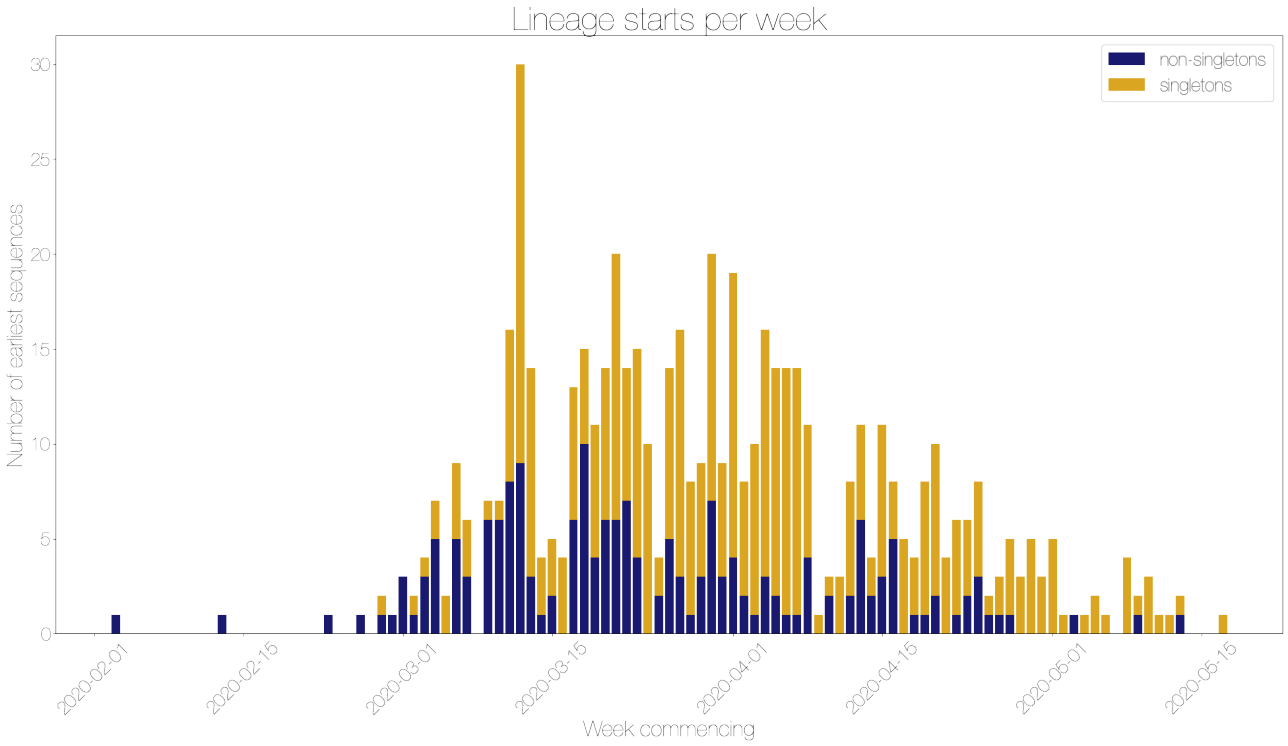
The relative growth and decline of the ten most sampled lineages in terms of number of counties they are present in is shown below. The raw data for the plot is shown below it, with each column representing a lineage, and the number of admin2 regions it is present in in each week.



Week commencing	UK36	UK52	UK225	UK40	UK7	UK39	UK5	UK175	UK1323	UK43
2020-03-01	0	1	0	0	0	0	0	0	0	0
2020-03-08	0	0	3	1	0	1	1	0	0	0
2020-03-15	2	3	9	3	5	2	4	0	3	0
2020-03-22	3	5	8	4	9	3	3	4	6	3
2020-03-29	8	4	9	4	6	2	4	1	4	4
2020-04-05	6	2	2	3	6	3	5	2	2	2
2020-04-12	10	5	0	1	4	2	3	2	2	3
2020-04-19	8	4	4	2	3	0	5	1	3	2
2020-04-26	7	4	1	1	3	1	4	1	1	1
2020-05-03	1	1	0	0	2	0	2	1	0	0
2020-05-10	0	0	0	0	2	0	0	0	0	0

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

NB the lineage may have started anywhere in the UK, but has been recorded at least once in Scotland

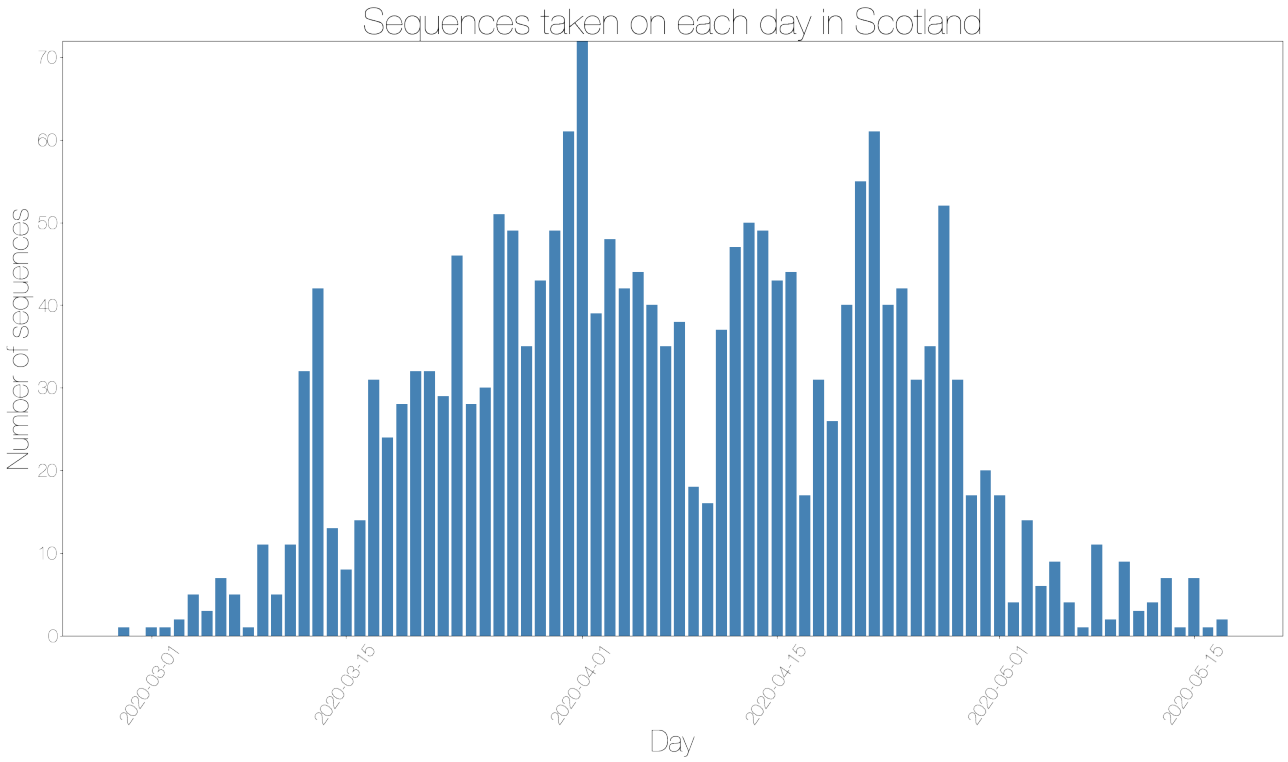


Day	Number of singleton starts	Number of non-singleton starts	Total
2020-02-03	0	1	1
2020-02-13	0	1	1
2020-02-23	0	1	1
2020-02-26	0	1	1
2020-02-28	1	1	2
2020-02-29	0	1	1
2020-03-01	0	3	3
2020-03-02	1	1	2
2020-03-03	1	3	4
2020-03-04	2	5	7
2020-03-05	2	0	2
2020-03-06	4	5	9
2020-03-07	3	3	6
2020-03-09	1	6	7
2020-03-10	1	6	7
2020-03-11	8	8	16
2020-03-12	21	9	30
2020-03-13	11	3	14
2020-03-14	3	1	4
2020-03-15	3	2	5
2020-03-16	4	0	4
2020-03-17	7	6	13
2020-03-18	5	10	15
2020-03-19	7	4	11
2020-03-20	8	6	14

Day	Number of singleton starts	Number of non-singleton starts	Total
2020-03-21	14	6	20
2020-03-22	7	7	14
2020-03-23	11	4	15
2020-03-24	10	0	10
2020-03-25	2	2	4
2020-03-26	9	5	14
2020-03-27	13	3	16
2020-03-28	7	1	8
2020-03-29	6	3	9
2020-03-30	13	7	20
2020-03-31	6	3	9
2020-04-01	15	4	19
2020-04-02	6	2	8
2020-04-03	9	1	10
2020-04-04	13	3	16
2020-04-05	12	2	14
2020-04-06	13	1	14
2020-04-07	13	1	14
2020-04-08	7	4	11
2020-04-09	1	0	1
2020-04-10	1	2	3
2020-04-11	3	0	3
2020-04-12	6	2	8
2020-04-13	5	6	11
2020-04-14	2	2	4
2020-04-15	8	3	11
2020-04-16	3	5	8
2020-04-17	5	0	5
2020-04-18	3	1	4
2020-04-19	7	1	8
2020-04-20	8	2	10
2020-04-21	4	0	4
2020-04-22	5	1	6
2020-04-23	4	2	6
2020-04-24	5	3	8
2020-04-25	1	1	2
2020-04-26	2	1	3
2020-04-27	4	1	5
2020-04-28	3	0	3
2020-04-29	5	0	5
2020-04-30	3	0	3
2020-05-01	5	0	5
2020-05-02	1	0	1
2020-05-03	0	1	1
2020-05-04	1	0	1
2020-05-05	2	0	2
2020-05-06	1	0	1
2020-05-08	4	0	4
2020-05-09	1	1	2

Day	Number of singleton starts	Number of non-singleton starts	Total
2020-05-10	3	0	3
2020-05-11	1	0	1
2020-05-12	1	0	1
2020-05-13	1	1	2
2020-05-17	1	0	1

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.

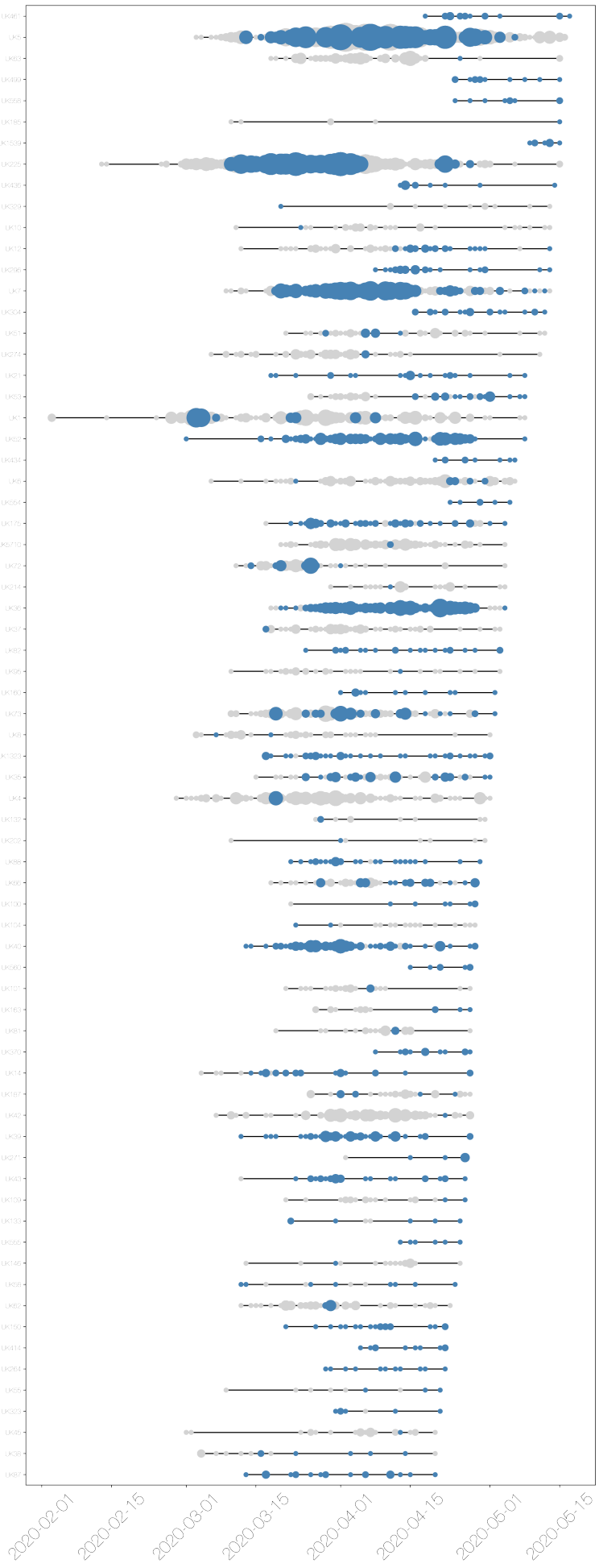


Day	Scotland
2020-02-28	1
2020-03-01	1
2020-03-02	1
2020-03-03	2
2020-03-04	5
2020-03-05	3
2020-03-06	7
2020-03-07	5
2020-03-08	1
2020-03-09	11
2020-03-10	5
2020-03-11	11
2020-03-12	32
2020-03-13	42
2020-03-14	13
2020-03-15	8

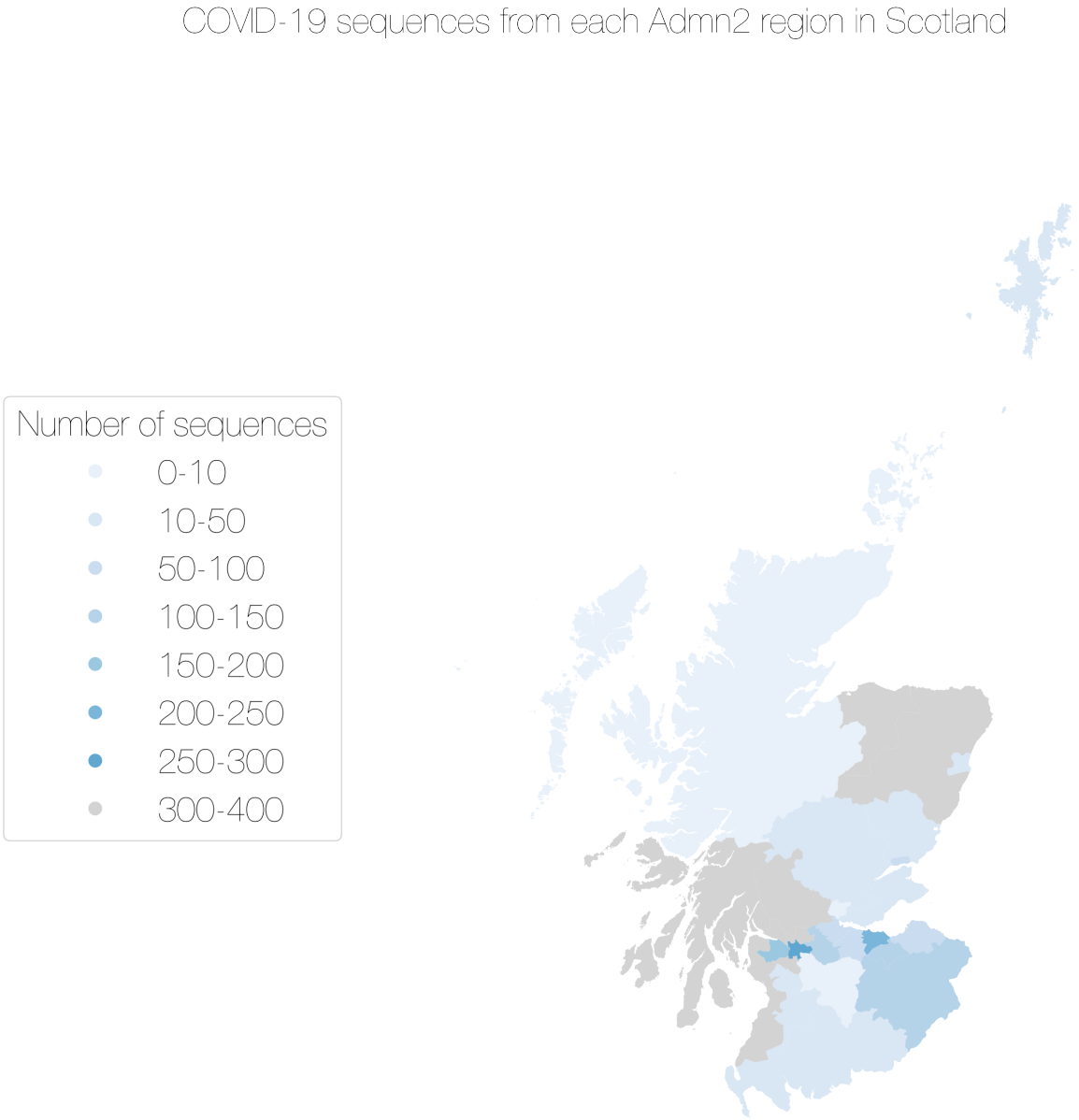
Day	Scotland
2020-03-16	14
2020-03-17	31
2020-03-18	24
2020-03-19	28
2020-03-20	32
2020-03-21	32
2020-03-22	29
2020-03-23	46
2020-03-24	28
2020-03-25	30
2020-03-26	51
2020-03-27	49
2020-03-28	35
2020-03-29	43
2020-03-30	49
2020-03-31	61
2020-04-01	72
2020-04-02	39
2020-04-03	48
2020-04-04	42
2020-04-05	44
2020-04-06	40
2020-04-07	35
2020-04-08	38
2020-04-09	18
2020-04-10	16
2020-04-11	37
2020-04-12	47
2020-04-13	50
2020-04-14	49
2020-04-15	43
2020-04-16	44
2020-04-17	17
2020-04-18	31
2020-04-19	26
2020-04-20	40
2020-04-21	55
2020-04-22	61
2020-04-23	40
2020-04-24	42
2020-04-25	31
2020-04-26	35
2020-04-27	52
2020-04-28	31
2020-04-29	17
2020-04-30	20
2020-05-01	17
2020-05-02	4
2020-05-03	14

Day	Scotland
2020-05-04	6
2020-05-05	9
2020-05-06	4
2020-05-07	1
2020-05-08	11
2020-05-09	2
2020-05-10	9
2020-05-11	3
2020-05-12	4
2020-05-13	7
2020-05-14	1
2020-05-15	7
2020-05-16	1
2020-05-17	2

These lineages 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. The raw data has been written to a summary file.


2020-02-01
2020-02-15
2020-03-01
2020-03-15
2020-04-01
2020-04-15
2020-05-01
2020-05-15

The map below 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.



Admin2	Country	Number of sequences	Sequence group
ABERDEEN	Scotland	21	10-50
ABERDEENSHIRE	Scotland	0	0
ANGLESEY	Wales	18	10-50
ANGUS	Scotland	10	10-50
ANTRIM	Northern Ireland	109	100-150
ARGYLL AND BUTE	Scotland	0	0
ARMAGH	Northern Ireland	12	10-50
BATH AND NORTH EAST SOMERSET	England	0	0
BEDFORDSHIRE	England	415	400-500
BERKSHIRE	England	7	1-10

Admin2	Country	Number of sequences	Sequence group
BLACKBURN WITH DARWEN	England	0	0
BLACKPOOL	England	0	0
BLAENAU GWENT	Wales	42	10-50
BOLTON	England	0	0
BOURNEMOUTH	England	0	0
BRIDGEND	Wales	83	50-100
BRIGHTON AND HOVE	England	0	0
BRISTOL	England	18	10-50
BUCKINGHAMSHIRE	England	326	300-400
BURY	England	0	0
CAERPHILLY	Wales	97	50-100
CAMBRIDGESHIRE	England	601	>500
CARDIFF	Wales	310	300-400
CARMARTHENSHIRE	Wales	75	50-100
CENTRAL BEDFORDSHIRE	England	0	0
CEREDIGION	Wales	16	10-50
CHESHIRE	England	8	1-10
CLACKMANNANSHIRE	Scotland	2	1-10
CONWY	Wales	37	10-50
CORNWALL	England	13	10-50
CUMBRIA	England	8	1-10
DARLINGTON	England	0	0
DENBIGHSHIRE	Wales	64	50-100
DERBY	England	0	0
DERBYSHIRE	England	25	10-50
DEVON	England	231	200-250
DORSET	England	140	100-150
DOWN	Northern Ireland	52	50-100
DUMFRIES AND GALLOWAY	Scotland	38	10-50
DUNDEE	Scotland	70	50-100
DURHAM	England	1	1-10
EAST AYRSHIRE	Scotland	36	10-50
EAST DUNBARTONSHIRE	Scotland	0	0
EAST LOTHIAN	Scotland	51	50-100
EAST RENFREWSHIRE	Scotland	0	0
EAST RIDING OF YORKSHIRE	England	20	10-50
EDINBURGH	Scotland	397	300-400
EILEAN SIAR	Scotland	2	1-10
ESSEX	England	1084	>500
FALKIRK	Scotland	62	50-100
FERMANAGH	Northern Ireland	3	1-10
FIFE	Scotland	41	10-50
FLINTSHIRE	Wales	46	10-50
GATESHEAD	England	0	0
GLASGOW	Scotland	606	>500
GLOUCESTERSHIRE	England	246	200-250
GREATER LONDON	England	2162	>500
GUERNSEY	Channel_islands	41	10-50
GWYNEDD	Wales	39	10-50

Admin2	Country	Number of sequences	Sequence group
HALTON	England	0	0
HAMPSHIRE	England	88	50-100
HARTLEPOOL	England	0	0
HEREFORDSHIRE	England	1	1-10
HERTFORDSHIRE	England	838	>500
HIGHLAND	Scotland	9	1-10
INVERCLYDE	Scotland	0	0
ISLE OF WIGHT	England	0	0
ISLES OF SCILLY	England	0	0
JERSEY	Channel_islands	77	50-100
KENT	England	27	10-50
KINGSTON UPON HULL	England	0	0
LANCASHIRE	England	6	1-10
LEICESTER	England	0	0
LEICESTERSHIRE	England	5	1-10
LINCOLNSHIRE	England	14	10-50
LONDONDERRY	Northern Ireland	10	10-50
LUTON	England	0	0
MANCHESTER	England	29	10-50
MEDWAY	England	0	0
MERSEYSIDE	England	59	50-100
MERTHYR TYDFIL	Wales	41	10-50
MIDDLESBROUGH	England	0	0
MIDLOTHIAN	Scotland	119	100-150
MILTON KEYNES	England	0	0
MONMOUTHSHIRE	Wales	46	10-50
MORAY	Scotland	0	0
NEATH PORT TALBOT	Wales	85	50-100
NEWPORT	Wales	112	100-150
NORFOLK	England	324	300-400
NORTH AYRSHIRE	Scotland	0	0
NORTH LANARKSHIRE	Scotland	103	100-150
NORTH LINCOLNSHIRE	England	0	0
NORTH SOMERSET	England	0	0
NORTH YORKSHIRE	England	4	1-10
NORTHAMPTONSHIRE	England	22	10-50
NORTHUMBERLAND	England	2	1-10
NOTTINGHAM	England	552	>500
NOTTINGHAMSHIRE	England	58	50-100
OLDHAM	England	0	0
ORKNEY ISLANDS	Scotland	1	1-10
OXFORDSHIRE	England	91	50-100
PEMBROKESHIRE	Wales	56	50-100
PERTHSHIRE AND KINROSS	Scotland	14	10-50
PETERBOROUGH	England	0	0
PLYMOUTH	England	0	0
POOLE	England	0	0
PORTSMOUTH	England	0	0
POWYS	Wales	35	10-50

Admin2	Country	Number of sequences	Sequence group
REDCAR AND CLEVELAND	England	0	0
RENFREWSHIRE	Scotland	157	150-200
RHONDDA, CYNON, TAFF	Wales	0	0
ROCHDALE	England	0	0
RUTLAND	England	0	0
SALFORD	England	0	0
SCOTTISH BORDERS	Scotland	102	100-150
SHETLAND ISLANDS	Scotland	14	10-50
SHROPSHIRE	England	1	1-10
SOMERSET	England	231	200-250
SOUTH AYRSHIRE	Scotland	0	0
SOUTH GLOUCESTERSHIRE	England	0	0
SOUTH LANARKSHIRE	Scotland	3	1-10
SOUTH YORKSHIRE	England	1058	>500
SOUTHAMPTON	England	0	0
SOUTHEND-ON-SEA	England	0	0
STAFFORDSHIRE	England	24	10-50
STIRLING	Scotland	0	0
STOCKPORT	England	0	0
STOCKTON-ON-TEES	England	0	0
STOKE-ON-TRENT	England	0	0
SUFFOLK	England	392	300-400
SURREY	England	56	50-100
SUSSEX	England	1	1-10
SWANSEA	Wales	200	200-250
SWINDON	England	0	0
TAMESIDE	England	0	0
TELFORD AND WREKIN	England	0	0
THURROCK	England	0	0
TORBAY	England	0	0
TORFAEN	Wales	71	50-100
TRAFFORD	England	0	0
TYNE AND WEAR	England	37	10-50
TYRONE	Northern Ireland	13	10-50
VALE OF GLAMORGAN	Wales	118	100-150
WARRINGTON	England	0	0
WARWICKSHIRE	England	9	1-10
WEST DUNBARTONSHIRE	Scotland	0	0
WEST LOTHIAN	Scotland	88	50-100
WEST MIDLANDS	England	87	50-100
WEST YORKSHIRE	England	19	10-50
WIGAN	England	0	0
WILTSHIRE	England	150	150-200
WORCESTERSHIRE	England	7	1-10
WREXHAM	Wales	64	50-100
YORK	England	0	0

There are some sequences with locations that are not matched to real Admin2 regions, some manual curation

required.

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