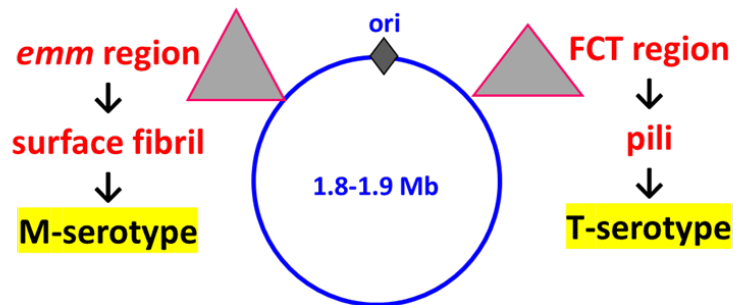


**PUBMLST ISOLATES DATABASE:
RULES-BASED ASSIGNMENT OF *EMM* AND PILIN GENOTYPES
FOR GROUP A STREPTOCOCCI (SPYOGENES)**

Updated February 2024

Contents:

Topic	Page
emm type and subtype	2-3
emm pattern	4-6
emm cluster	6
Pilin alleles	7
pilA and pilB nt80 clusters	8-13
Pilin type	14
FCT-region form	15



Group A streptococci-specific typing resources:

- CDC link (*emm* type deposits):
<https://www2.cdc.gov/vaccines/biotech/strepblast.asp>
- Bessen DE, BW Beall, A Hayes, W Huang, JM DiChiara, S Velusamy, H Tettelin, KA Jolley, JT Fallon, S Chochua, MSA Alobaidallah, C Higgs, TC Barnett, JD Steemson, T Proft and MR Davies, 2024. Recombinational exchange of M-fibril and T-pilus genes generates extensive cell surface diversity in the global group A Streptococcus population. mBio, under review.
- Sanderson-Smith M, De Oliveira DM, Guglielmini J, McMillan DJ, Vu T, Holien JK, Henningham A, Steer AC, Bessen DE, Dale JB, Curtis N, Beall BW, Walker MJ, Parker MW, Carapetis JR, Van Melder L, Sriprakash KS, Smeesters PR; M Protein Study Group. J Infect Dis. 2014, 210:1325-38.

emm type, subtype, cluster and pattern:

1. emm type and subtype assignment

Given the historical importance of *emm* type in group A strep epidemiology, new isolate submissions are **expected** to minimally provide 'emm type' in addition to the 7 core housekeeping alleles.

- The *emm* type format is a simple number of 1 to 3 digits.
- Allowable values are listed in the Isolate Database template (Excel file).

How to establish *emm* type?

- By the CDC BLAST server: <https://www2.cdc.gov/vaccines/biotech/strepblast.asp>
- By www.pubmlst.org/spyogenes: Sequence queries with whole genome sequences identify exact matches with the PubMLST Sequence Typing Database of *emm* subtypes (i.e., partial allele), which is curated by the CDC and updated to PubMLST every 1-2 months.
 - The *emm* subtype format is emm#.#. For e.g., emm5.32 or emm103.1
 - In the example above, the *emm* types are 5 and 103, respectively.
 - The **emm subtype is optional** for new isolate submissions.
 - Some *emm* types (emmCG) are derived from a related streptococcal species (e.g., stG6). Please consult with the Curator as to how this information might best be handled.
 - The development of Sequence Typing Databases is ongoing for *emm*-like (*emmL*) genes. This information may be useful in sequence queries but is not yet available in the Isolate Database.
- For sequence queries with multiple *emm* and *emmL* hits: **Which one is *emm*?**
 - The position of the hit with emmPrime (~50 nt upstream) establishes the identity of the (downstream) *emm* locus.

Locus	Allele	Length	Contig	Start position	End position
emm	53.0	180	sid 463836 accn CP003068__Streptococcus_pyogenes_Alaba49__complete_genome__[Streptococcus_pyogenes_Alaba49]	1656846	1657025
emmL	134.2	180	sid 463836 accn CP003068__Streptococcus_pyogenes_Alaba49__complete_genome__[Streptococcus_pyogenes_Alaba49]	1655354	1655533
emmPrime	2	19	sid 463836 accn CP003068__Streptococcus_pyogenes_Alaba49__complete_genome__[Streptococcus_pyogenes_Alaba49]	1657073	1657091
mga-2	17	32	sid 463836 accn CP003068__Streptococcus_pyogenes_Alaba49__complete_genome__[Streptococcus_pyogenes_Alaba49]	1658718	1658749
SF1	1	26	sid 463836 accn CP003068__Streptococcus_pyogenes_Alaba49__complete_genome__[Streptococcus_pyogenes_Alaba49]	1656004	1656029
SF3	1	30	sid 463836 accn CP003068__Streptococcus_pyogenes_Alaba49__complete_genome__[Streptococcus_pyogenes_Alaba49]	1654688	1654717
SF4	9	54	sid 463836 accn CP003068__Streptococcus_pyogenes_Alaba49__complete_genome__[Streptococcus_pyogenes_Alaba49]	1657468	1657521

- For sequence queries with no *emm* type hits → (possibly) a new *emm* subtype
 - New *emm* subtype assignments are made by the CDC.
 - CDC link: <https://www.cdc.gov/streplab/groupa-strep/index.html>
 - The PubMLST curator can also provide guidance if needed.
 - However - the PubMLST software is not designed to accept new *emm* subtype submissions directly from users.

2. *emm* pattern assignment (optional)

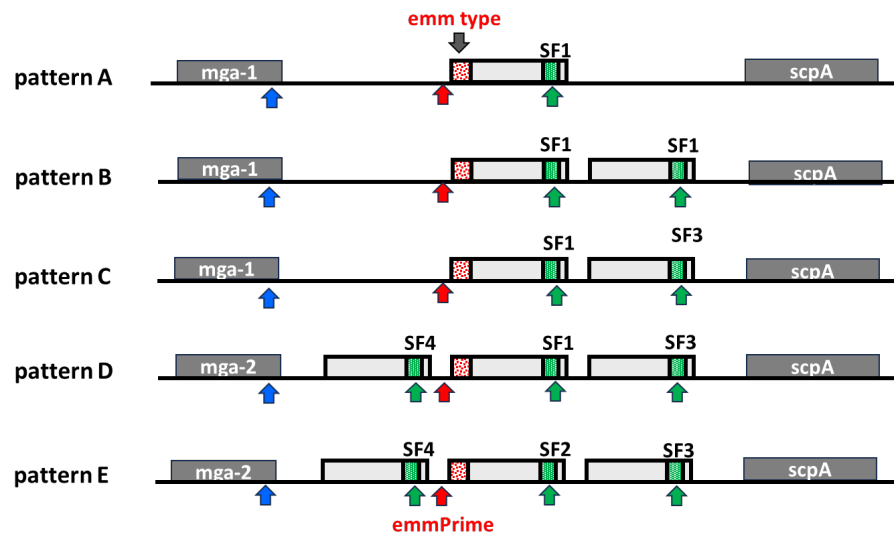
- The Isolate Database has 2 fields for submission of *emm* pattern (optional)
 - emm_pattern_measured*
 - Values** are A-C, D, E, REA (rearranged), A, B, C
 - emm_pattern_predicted*
 - Values** are A-C, D, E, REA (rearranged)

How to establish *emm* pattern (i.e., measured)?

- Combinations of hits with *emm* typing loci are used to measure *emm* pattern.
- Note: The PubMLST databases include nearly all known variants of the (partial) *emm* pattern-typing loci but may not be 100% inclusive. Please contact the Curator if you have new allelic variants and/or wish to submit a direct deposit.

<i>emm</i> pattern	<i>mga-1</i>	<i>mga-2</i>	SF1	SF2	SF3	SF4	<i>emm</i>
A	hit		hit				hit
B	hit		2 hits				hit
C	hit		hit		hit		hit
D		hit	hit		hit	hit	hit
E		hit		hit	hit	hit	hit
REA (example)		hit			hit	hit	hit (or <i>emmL</i> only)

- If the SF hit combination differs from those established for *emm* patterns A through E (above), one can query the genome for each individual SF to identify a new allelic variant (i.e., $\geq 90\%$ id). The Curator can offer further guidance.



Example (from downloaded xls following sequence query): *emm53*-pattern D

Locus	Allele	Length	Contig	Start position	End position
<i>mga-2</i>	17	32	sid 463836 accn CP003068__Streptococcus_pyogenes_Alalab49,_complete_genome._	1658718	1658749
SF4	9	54	sid 463836 accn CP003068__Streptococcus_pyogenes_Alalab49,_complete_genome._	1657468	1657521
<i>emmPrime</i>	2	19	sid 463836 accn CP003068__Streptococcus_pyogenes_Alalab49,_complete_genome._	1657073	1657091
<i>emm</i>	53	180	sid 463836 accn CP003068__Streptococcus_pyogenes_Alalab49,_complete_genome._	1656846	1657025
SF1	1	26	sid 463836 accn CP003068__Streptococcus_pyogenes_Alalab49,_complete_genome._	1656004	1656029
<i>emmL</i>	134.2	180	sid 463836 accn CP003068__Streptococcus_pyogenes_Alalab49,_complete_genome._	1655354	1655533
SF3	1	30	sid 463836 accn CP003068__Streptococcus_pyogenes_Alalab49,_complete_genome._	1654688	1654717

How to predict *emm* pattern?

- Previously established associations between *emm* type and *emm* pattern groupings A-C, D and E are presented in the listing below.
- REA is included only for *emm* types for which an *emm* pattern group remains to be established.
- *emm* types recovered in association with >1 *emm* pattern group are highlighted.
- Data is based on the mBio 2024 study (Table S1).

emm type	emm pattern group	emm type	emm pattern group	emm type	emm pattern group	emm type	emm pattern group
1	A-C	56	D	106	E	208	D
2	E	57	A-C	107	E	209	E
3	A-C	58	E	108	D	213	D
4	E	59	D	109	E	216	D
5	A-C	60	E	110	E	217	D
6	A-C	63	E	111	D	218	A-C
8	E	64	D	112	E	218	D
9	E	65	D	113	E	221	D
11	E	66	E	114	E	222	A-C
12	A-C	67	D	115	D	223	D
13	E	68	E	116	D	224	D
14	A-C	70	D	117	E	225	D
15	E	71	D	118	E	226	A-C
17	A-C	72	D	119	D	227	A-C
18	A-C	73	E	120	D	229	A-C
18	D	74	D	121	D	230	D
19	A-C	75	E	122	D	231	E
22	E	76	E	123	D	232	E
23	A-C	77	E	124	E	233	A-C
24	A-C	78	E	137	A-C	233	D
25	E	79	E	138	REA	238	A-C
26	A-C	80	D	147	D	239	A-C
27	E	81	D	151	E	241	A-C
28	E	81	D	162	D	242	D
29	A-C	82	E	164	REA	stG1750	REA
30	A-C	83	D	165	E	stG653	A-C
30	D	84	E	166	A-C	stG7882	A-C
31	A-C	85	D	166	E	stG866	A-C
32	D	86	D	168	E		
33	D	87	E	169	E		
34	REA	88	E	171	REA		
36	D	89	E	172	D		
37	A-C	90	E	173	E		
38	A-C	91	D	174	REA		
39	A-C	92	E	176	E		
41	D	93	D	177	E		
42	D	94	E	178	D		
43	D	95	D	179	D		
44	E	95	D	180	E		
46	A-C	95	D	182	D		
47	A-C	96	E	183	E		
48	E	97	D	184	D		
49	E	98	D	185	D		
50	E	99	D	186	D		
51	A-C	100	D	191	D		
52	D	101	D	192	D		
53	D	102	E	193	A-C		
54	A-C	103	E	197	A-C		
54	D	104	E	205	D		
55	A-C	105	REA	207	D		

3. *emm* cluster assignment (optional)

The Isolates Database allows for entry of 'emm_cluster' based on:

- Sanderson-Smith M, De Oliveira DM, Guglielmini J, McMillan DJ, Vu T, Holien JK, Henningham A, Steer AC, Bessen DE, Dale JB, Curtis N, Beall BW, Walker MJ, Parker MW, Carapetis JR, Van Melder L, Sriprakash KS, Smeesters PR; M Protein Study Group. J Infect Dis. 2014 Oct 15;210(8):1325-38

Values are in **text** format, as listed below.

- Singletons and outliers should be listed in the isolate template xls file in the format exemplified by: clade_Y_M6 and outlier_M55
- If you are uncertain of the proper format, one can examine the Isolate Database.

clade	<i>emm</i> cluster	<i>emm</i> types represented
X	E1	4, 60, 78, 165, 176
X	E2	13, 27, 50, 66, 68, 76, 90, 92, 96, 104, 106, 110, 117, 166, 168
X	E3	9, 15, 25, 44, 49, 58, 79, 82, 87, 103, 107, 113, 118, 144, 180, 183, 209, 219, 231
X	E4	2, 8, 22, 28, 73, 77, 84, 88, 89, 102, 109, 112, 114, 124, 169, 175, 232
X	E5	34, 51, 134, 137, 170, 174, 205
X	E6	11, 42, 48, 59, 63, 65, 67, 75, 81, 85, 94, 99, 139, 158, 172, 177, 182, 191
X	Singletons	164, 185, 211, 236
Y	D1	36, 54, 207
Y	D2	32, 71, 100, 115, 213
Y	D3	123, 217
Y	D4	33, 41, 43, 52, 53, 56, 64, 70, 72, 80, 83, 86, 91, 93, 98, 101, 108, 116, 119, 120, 121, 178, 186, 192, 194, 208, 223, 224, 225, 230, 242
Y	D5	97, 157, 184
Y	A-C1	46, 142
Y	A-C2	30, 197
Y	A-C3	1, 163, 227, 238, 239
Y	A-C4	12, 39, 193, 228, 229
Y	A-C5	3, 31, 133
Y	Singletons	5, 6, 14, 17, 18, 19, 23, 24, 26, 29, 37, 38, 47, 57, 74, 105, 122, 140, 179, 218, 233, 234
Outliers	Outliers	55, 95, 111, 215, 221, 222


FCT-region form, pilin type, pil nt80 sequence clusters and pilin alleles

1. Pilin alleles (optional)

- New pilin alleles can be directly submitted to the Curator for **integer** assignment.
- A listing of pilin loci in the Typing Databases is provided below.
- Details are presented in the mBio 2024 study. Briefly:
 - pilA, adhesin
 - pilB, backbone
 - pilL, linker
 - Numerals refer to FCT-regions forms: 1, 2, 3 and 4 (34), 5, 6, 9, 10.

pilA1
pilA2
pilA34
pilA5
pilA6
pilB1
pilB10
pilB2
pilB34
pilB5
pilB6
pilB9
pilL1
pilL10
pilL2
pilL34
pilL5
pilL6
pilL9

2. pilA_nt80 and pilB_nt80 clusters (optional).

- nt80 sequence clusters for established pilA and pilB alleles are listed in the charts below.
- nt80 sequence clusters are assigned by the Curator for new pilA and pilB alleles.
- Further details are presented in the mBio 2024 study.
- The Table S1 xls file in the mBio 2024 study lists pilA and pilB alleles and cluster assignments, and data can be readily sorted.
- See example in section #4: 

pilA1

pilA1 allele	pilA_nt80 cluster
1	nt80_4
2	nt80_3
3	nt80_3
4	nt80_3
5	nt80_4
6	nt80_4
7	nt80_3
8	nt80_3
9	nt80_3
10	nt80_4
11	nt80_4
12	nt80_3
13	nt80_3
14	nt80_4
15	nt80_4
16	nt80_4
17	nt80_4
18	nt80_3
19	nt80_3
20	nt80_4
21	nt80_4
22	nt80_4
23	nt80_3
24	nt80_3
25	nt80_3
26	nt80_4
27	nt80_4
28	nt80_4
29	nt80_3
30	nt80_4

pilA2

pilA2 allele	pilA_nt80 cluster
1	nt80_7
2	nt80_7
3	nt80_7
4	nt80_7
5	nt80_7

pilA6

pilA6 allele	pilA_nt80 cluster
1	nt80_5

pilA34

pilA34 allele	pilA_nt80 cluster		pilA34 allele	pilA_nt80 cluster		pilA34 allele	pilA_nt80 cluster		pilA34 allele	pilA_nt80 cluster
1	10		52	12		102	9		152	12
2	13		53	12	✿	103	13		152	12
3	9		54	12		104	12		153	6
4	13		55	16		105	13		154	6
5	9		56	6		106	9		155	9
6	10		57	15		107	8		156	6
7	13		58	12		108	12		157	13
8	13		59	14		109	12		158	14
9	9		60	14		110	6		159	9
10	9		61	14		111	14		160	6
12	13		62	14		112	17		161	6
13	10		63	14		112	17		162	15
14	11		64	14		112	17		163	12
15	6		65	12		113	16		164	6
16	8		66	14		114	8		165	9
17	8		67	17		115	6		166	10
18	8		68	16		116	12		167	8
19	10		69	16		117	17		168	8
20	6		70	17		118	11		169	14
21	8		71	17		119	6		170	9
22	8		72	17		120	14		171	12
23	6		73	17		121	12		172	12
24	8		74	9		122	6		173	13
25	14		75	13		123	12		174	15
26	12		76	13		124	9		175	12
27	16		77	10		125	13		176	13
28	14		78	12		126	12		177	8
29	18		79	16		127	9		178	13
30	18		80	9		130	8		179	15
31	15		81	12		131	8		180	9
32	12		82	9		132	13		181	14
33	14		83	6		133	17		182	12
34	12		84	8		134	14		183	9
35	14		85	9		135	12		184	9
36	14		86	9		136	14		185	12
37	12		87	9		137	8		186	12
38	15		88	6		138	12		187	13
39	17		89	6		139	6		188	13
40	14		90	9		140	12		198	13
41	17		91	13		141	12			
42	16		92	6		142	10			
43	12		93	6		143	6			
44	17		94	6		144	13			
45	13		95	13		145	12			
46	12		96	6		146	9			
47	12		97	6		147	12			
48	12		98	12		148	14			
49	14		99	11		149	12			
50	12		100	8		150	14			
51	12		101	9		151	10			

pilA5

pilA5 allele	pilA_nt80 cluster
1	nt80_1
2	nt80_1
3	nt80_1
4	nt80_1
5	nt80_1
6	nt80_1
7	nt80_1
8	nt80_1
9	nt80_0
10	nt80_0
11	nt80_0
12	nt80_0
15	nt80_1
17	nt80_1
18	nt80_0
19	nt80_1
20	nt80_1
21	nt80_1
22	nt80_1
26	nt80_0
27	nt80_0
28	nt80_0
29	nt80_1
31	nt80_1
32	nt80_1
33	nt80_1
34	nt80_0
35	nt80_0

pilA10

pilA10 allele	pilA_nt80 cluster
13	nt80_2
14	nt80_2
16	nt80_2
23	nt80_2
23	nt80_2
24	nt80_2
25	nt80_2
30	nt80_2

pilB1

pilB1 allele	pilB_nt80 cluster
1	nt80_20
2	nt80_24
3	nt80_21
4	nt80_21
5	nt80_21
6	nt80_25
7	nt80_21
8	nt80_21
9	nt80_22
10	nt80_23
11	nt80_26
12	nt80_25
13	nt80_21
15	nt80_22
16	nt80_20
17	nt80_21
18	nt80_24
20	nt80_24
21	nt80_23
22	nt80_20
23	nt80_25
24	nt80_21


pilB2

pilB2 allele	pilB_nt80 cluster
1	nt80_39
2	nt80_39

pilB6

pilB6 allele	pilB_nt80 cluster
1	nt80_11

pilB34

pilB34 allele	pilB_nt80 cluster		pilB34 allele	pilB_nt80 cluster		pilB34 allele	pilB_nt80 cluster
2	nt80_34		51	nt80_29		99	nt80_34
3	nt80_34		52	nt80_29		100	nt80_37
4	nt80_29		53	nt80_29		101	nt80_34
5	nt80_37		54	nt80_29		102	nt80_33
6	nt80_28		55	nt80_37		103	nt80_36
7	nt80_28		56	nt80_34		104	nt80_34
8	nt80_30		57	nt80_34		105	nt80_31
9	nt80_27		58	nt80_34		106	nt80_27
10	nt80_33		59	nt80_34		107	nt80_33
11	nt80_28		60	nt80_35		108	nt80_38
12	nt80_27		61	nt80_34		109	nt80_36
13	nt80_28		62	nt80_36		110	nt80_30
14	nt80_28		63	nt80_36		111	nt80_33
15	nt80_38		64	nt80_36		112	nt80_31
16	nt80_31		65	nt80_36		113	nt80_30
17	nt80_32		66	nt80_36		114	nt80_36
18	nt80_34		67	nt80_33		115	nt80_37
19	nt80_28		68	nt80_33		116	nt80_38
20	nt80_29		69	nt80_31		117	nt80_36
21	nt80_34		70	nt80_31		118	nt80_29
22	nt80_36		71	nt80_31		119	nt80_28
23	nt80_36		72	nt80_35		119	nt80_28
24	nt80_34		73	nt80_36		120	nt80_30
25	nt80_34		74	nt80_36		121	nt80_32
26	nt80_34		75	nt80_36		122	nt80_34
27	nt80_37		76	nt80_35		123	nt80_33
28	nt80_37		77	nt80_35		124	nt80_35
29	nt80_37		78	nt80_35		125	nt80_30
30	nt80_28		79	nt80_35		126	nt80_34
31	nt80_30		80	nt80_35		127	nt80_36
32	nt80_30		81	nt80_35		128	nt80_36
33	nt80_30		82	nt80_35		129	nt80_28
34	nt80_30		83	nt80_34		130	nt80_34
35	nt80_33		84	nt80_34		131	nt80_31
36	nt80_33		85	nt80_30		132	nt80_37
37	nt80_28		86	nt80_36		133	nt80_28
38	nt80_28		87	nt80_36		134	nt80_32
39	nt80_28		88	nt80_35		135	nt80_36
40	nt80_38		89	nt80_35		136	nt80_29
41	nt80_38		90	nt80_29		137	nt80_29
42	nt80_38		91	nt80_28		138	nt80_33
43	nt80_38		92	nt80_37		139	nt80_37
44	nt80_31		93	nt80_32			
45	nt80_32		94	nt80_35			
46	nt80_32		95	nt80_32			
47	nt80_32		96	nt80_35			
48	nt80_32		97	nt80_36			
49	nt80_34		98	nt80_34			
50	nt80_34						

pilB5

pilB5 allele	pilB_nt80 cluster
1	nt80_1
2	nt80_1
3	nt80_2
4	nt80_1
5	nt80_1
6	nt80_1
6	nt80_1
7	nt80_3
8	nt80_3
9	nt80_2
10	nt80_2
11	nt80_3
11	nt80_3
12	nt80_1
13	nt80_2
14	nt80_2
15	nt80_4
16	nt80_4
17	nt80_2
18	nt80_2
19	nt80_4
20	nt80_4

pilB10


pilB10 allele	pilB_nt80 cluster
1	nt80_5
2	nt80_0
3	nt80_0
4	nt80_0
5	nt80_5
6	nt80_5

pilB9

pilB9 allele	pilB_nt80 cluster
1	nt80_14
2	nt80_14
3	nt80_12
4	nt80_7
5	nt80_7
6	nt80_7
7	nt80_13
8	nt80_15
9	nt80_15
10	nt80_17
11	nt80_12
12	nt80_19
13	nt80_19
14	nt80_15
15	nt80_12
16	nt80_14
17	nt80_19
18	nt80_14
19	nt80_8
20	nt80_15
21	nt80_18
22	nt80_18
23	nt80_19
24	nt80_10
25	nt80_13
26	nt80_10
27	nt80_9
28	nt80_16
29	nt80_6

3. Pilin type (optional)

- Pilin type is defined based on the (unique) combination of pilA_nt80 and pilB_nt80 clusters, as listed in the chart below.
- New pilin types are assigned by the Curator.

pilin type	pilA_nt80 cluster	pilB_nt80 cluster	FCT-region form		pilin type	pilA_nt80 cluster	pilB_nt80 cluster	FCT-region form
pil001	nt80_7	nt80_39	FCT-2		pil066	nt80_9	nt80_38	FCT-4
pil002	nt80_5	nt80_11	FCT-6		pil068	nt80_3	nt80_25	FCT-1
pil003	nt80_14	nt80_35	FCT-3		pil070	nt80_12	nt80_34	FCT-3
pil004	nt80_1	nt80_1	FCT-5		pil072	nt80_14	nt80_38	FCT-3
pil005	nt80_18	nt80_33	FCT-3		pil074	nt80_14	nt80_32	FCT-3
pil005	nt80_18	nt80_33	FCT-3		pil075	n.a.	nt80_8	FCT-9
pil006	nt80_4	nt80_25	FCT-1		pil076	nt80_4	nt80_21	FCT-1
pil007	n.a.	nt80_16	FCT-9		pil077	nt80_6	nt80_29	FCT-4
pil008	n.a.	nt80_19	FCT-9		pil078	nt80_10	nt80_33	FCT-4
pil009	nt80_6	nt80_32	FCT-4		pil080	nt80_14	nt80_36	FCT-3
pil010	n.a.	nt80_6	FCT-9		pil081	nt80_1	nt80_3	FCT-5
pil011	nt80_13	nt80_27	FCT-4		pil083	nt80_17	nt80_37	FCT-3
pil012	nt80_8	nt80_38	FCT-4		pil084	nt80_9	nt80_34	FCT-4
pil013	nt80_10	nt80_29	FCT-4		pil085	nt80_9	nt80_28	FCT-4
pil014	nt80_11	nt80_36	FCT-4		pil087	nt80_8	nt80_34	FCT-4
pil015	nt80_9	nt80_35	FCT-4		pil088	nt80_2	nt80_5	FCT-10
pil016	nt80_1	nt80_4	FCT-5		pil089	nt80_9	nt80_30	FCT-4
pil018	nt80_17	nt80_36	FCT-3		pil090	nt80_9	nt80_31	FCT-4
pil019	n.a.	nt80_12	FCT-9		pil092	nt80_8	nt80_28	FCT-4
pil020	n.a.	nt80_10	FCT-9		pil094	nt80_14	nt80_30	FCT-3
pil021	nt80_10	nt80_31	FCT-4		pil095	nt80_4	nt80_26	FCT-1
pil022	nt80_2	nt80_0	FCT-10		pil096	nt80_13	nt80_35	FCT-4
pil023	nt80_3	nt80_21	FCT-1		pil098	nt80_15	nt80_34	FCT-3
pil024	nt80_12	nt80_38	FCT-3		pil099	nt80_15	nt80_36	FCT-3
pil025	n.a.	nt80_14	FCT-9		pil100	nt80_15	nt80_35	FCT-3
pil027	nt80_11	nt80_33	FCT-4		pil101	nt80_17	nt80_29	FCT-3
 pil028	nt80_13	nt80_34	FCT-4		pil102	nt80_0	nt80_4	FCT-5
pil033	nt80_16	nt80_37	FCT-3		pil104	nt80_8	nt80_27	FCT-4
pil036	nt80_13	nt80_31	FCT-3		pil105	nt80_6	nt80_36	FCT-4
pil038	nt80_4	nt80_23	FCT-1		pil106	nt80_6	nt80_33	FCT-4
pil039	nt80_6	nt80_30	FCT-4		pil107	nt80_6	nt80_28	FCT-4
pil040	nt80_13	nt80_37	FCT-4		pil110	nt80_4	nt80_20	FCT-1
pil042	nt80_8	nt80_35	FCT-4		pil111	nt80_15	nt80_32	FCT-3
pil043	nt80_12	nt80_37	FCT-3		pil114	nt80_8	nt80_31	FCT-4
pil044	nt80_8	nt80_33	FCT-4		pil115	nt80_3	nt80_22	FCT-1
pil045	nt80_13	nt80_38	FCT-4		pil116	nt80_12	nt80_33	FCT-3
pil046	nt80_8	nt80_30	FCT-4		pil117	nt80_18	nt80_28	FCT-3
pil048	nt80_0	nt80_3	FCT-5		pil122	nt80_14	nt80_28	FCT-3
pil049	nt80_11	nt80_38	FCT-4		pil165	n.a.	nt80_18	FCT-9
pil052	nt80_16	nt80_36	FCT-3		pil177	nt80_9	nt80_33	FCT-4
pil053	nt80_12	nt80_36	FCT-3		pil178	nt80_17	nt80_35	FCT-3
pil054	nt80_3	nt80_24	FCT-1		pil183	nt80_9	nt80_37	FCT-4
pil055	n.a.	nt80_15	FCT-9		pil186	nt80_12	nt80_31	FCT-3
pil056	nt80_14	nt80_34	FCT-3		pil217	nt80_12	nt80_28	FCT-3
pil057	n.a.	nt80_17	FCT-9		pil222	nt80_10	nt80_32	FCT-4
pil058	n.a.	nt80_9	FCT-9		pil225	nt80_17	nt80_34	FCT-3
pil060	nt80_1	nt80_2	FCT-5		pil232	nt80_0	nt80_2	FCT-5
pil063	nt80_4	nt80_24	FCT-1		pil238	n.a.	nt80_13	FCT-9
pil064	nt80_12	nt80_35	FCT-3					
pil065	n.a.	nt80_7	FCT-9					

4. FCT-region form (optional)

- The Isolate Database has a field for submission of FCT-region form.
- Allowable **values** are:
 - FCT-1
 - FCT-2
 - FCT-3
 - FCT-4
 - FCT-3/4
 - FCT-5
 - FCT-6
 - FCT-9
 - FCT-10

How to establish FCT-region?

- Combinations of hits are listed in the chart below.
- Note: The PubMLST databases include nearly all known variants of the (partial) *rofA/nra* and full-length pilin loci but may not be 100% inclusive. Please contact the Curator if you have new allelic variants and/or wish to submit a direct deposit.

FCT-region	<i>rofA</i>	<i>nra</i>	<i>pilA1/pilB1</i>	<i>pilA2/pilB2</i>	<i>pilA34/pilB34</i>	<i>pilA5/pilB5</i>	<i>pilA6/pilB6</i>	<i>pilB9</i>	<i>pilA10/pilB10</i>
1	hit		hit						
2	hit			hit					
3		hit			hit				
4	hit				hit				
3/4					hit				
5	hit					hit			
6	hit						hit		
9	hit							hit	
10	hit								hit

❖ Example: (from downloaded xls following sequence query):

- FCT-4
- *pilA34_103* = *pilA nt80_13*; *pilB34_83* = *pilB nt80_34*
- *pil028* = *pilA nt80_13* and *pilB nt80_34*
- Also, *emm28*-pattern E

Locus	Allele	Length	Contig	Start position	End position
<i>rofA</i>	1	23	sid 1822 accn NC_007296__Streptococcus_pyogenes_MGAS6180_complete_genome_	116616	116638
❖ <i>pilA34</i>	103	2244	sid 1822 accn NC_007296__Streptococcus_pyogenes_MGAS6180_complete_genome_	120079	122322
❖ <i>pilB34</i>	83	1038	sid 1822 accn NC_007296__Streptococcus_pyogenes_MGAS6180_complete_genome_	122858	123895
<i>pilL34</i>	57	588	sid 1822 accn NC_007296__Streptococcus_pyogenes_MGAS6180_complete_genome_	124653	125240
SF3	1	30	sid 1822 accn NC_007296__Streptococcus_pyogenes_MGAS6180_complete_genome_	1709892	1709921
SF2	1	30	sid 1822 accn NC_007296__Streptococcus_pyogenes_MGAS6180_complete_genome_	1711041	1711070
<i>emm</i>	28	180	sid 1822 accn NC_007296__Streptococcus_pyogenes_MGAS6180_complete_genome_	1711812	1711991
<i>emmPrime</i>	2	19	sid 1822 accn NC_007296__Streptococcus_pyogenes_MGAS6180_complete_genome_	1712039	1712057
SF4	10	54	sid 1822 accn NC_007296__Streptococcus_pyogenes_MGAS6180_complete_genome_	1712429	1712482
<i>mga-2</i>	8	32	sid 1822 accn NC_007296__Streptococcus_pyogenes_MGAS6180_complete_genome_	1713682	1713713