

# MLST & SLST: Technical Background

**Dag Harmsen**

University of Münster, Germany

[dharmsen@uni-muenster.de](mailto:dharmsen@uni-muenster.de)

# Typing questions & suitable methods

Questions	Suitable methods	Discriminatory power	Time span
Outbreak investigations Short-term/local surveillance Control of hygiene measures	PFGE, RFLP, AFLP, RA-PCR, VNTR, <b>SLST</b> , micro-array	high	weeks - month
Long-term/global epidemiological studies Population genetics Analysis of population-based interventions, e.g. vaccination	(MLEE), <b>MLST</b> , micro-array, in part SLST	low	years

# Multilocus Sequence Typing (MLST)

ESCMID Online Lecture Library  
© by author

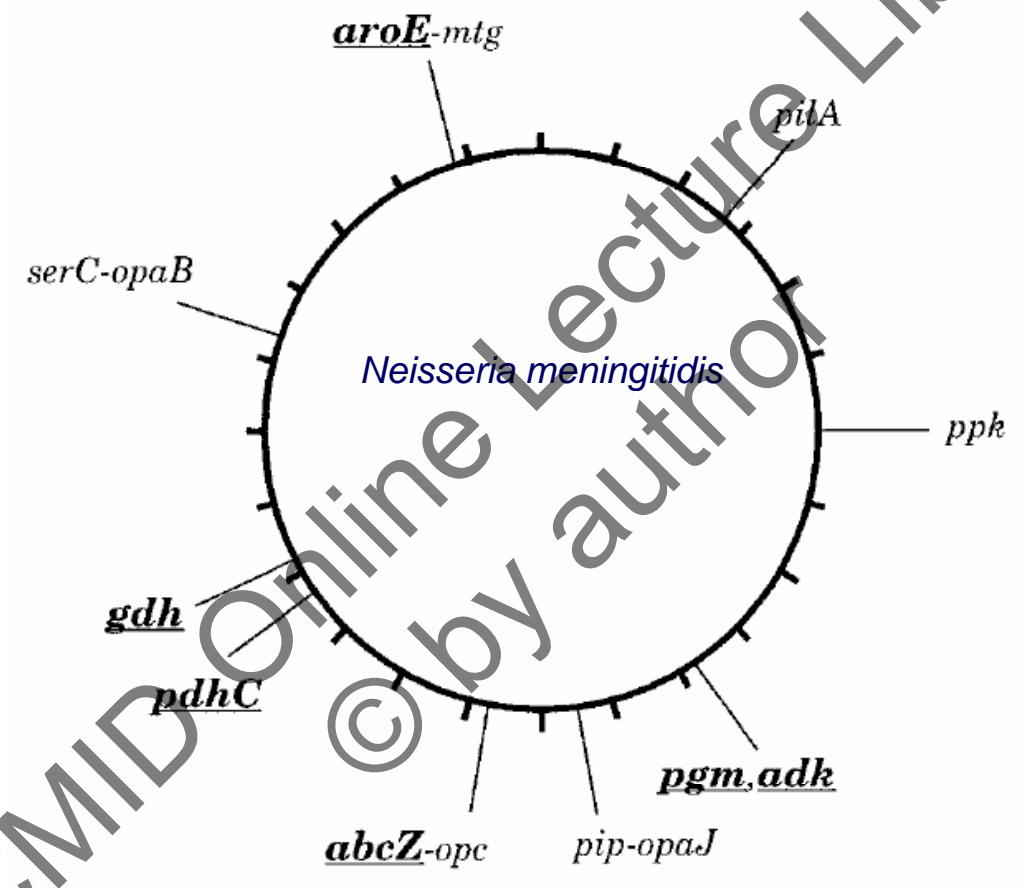
# Multilocus sequence typing (MLST)

- displays long term epidemiology, replaced Multilocus Enzyme Electrophoresis (MLEE) completely
- sequences of 5-7 housekeeping genes distributed over the whole genome (selectively neutral, slowly evolving, recombination independent loci)
- sequence determination of 5-7 internal fragments (one sequencing run) of these genes

# Multilocus sequence typing (MLST)

- identical gene sequences are assigned the same allele number
- for each strain, the combination of alleles at each locus defines its sequence type (ST)
- related STs are grouped into clonal complexes (CC)  $\rightarrow$  population structure
- single expanding central MLST database that can be interrogated electronically but that is curated manually

# Multilocus sequence typing (MLST)



Maiden et al., 1998 (*N. meningitidis*) [[Pubmed](#)]

Today: > 40 MLST schemes for different organisms published  
<http://www.mlst.net>

# Workflow of MLST

Amplification of ~450 bp internal fragment of 7 house keeping genes



Sequencing of seven gene fragments (forward and reverse strand)



Comparison of allelic sequences with the known alleles (<http://www.mlst.net>)



Assignment of all seven alleles



Allelic profile (combination of all alleles) results in sequence type (ST)



Comparison of STs to those of previously typed isolates  
and further analysis (e.g. eBURST)

# MLST database

1. MLST-profiles
2. Isolates
3. Allelic sequences (fastA, Mega)

ST	arcc	aroe	glpf	gmk	pta	tpi	yqil
1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	26
3	1	1	1	9	1	1	12
4	10	10	8	6	10	3	2
5	1	4	1	4	12	1	10
6	12	4	1	4	12	1	3
7	5	4	1	4	4	6	3
8	3	3	1	1	4	4	3
9	3	3	1	1	1	1	10
10	8	7	6	2	9	9	7



# MLST database

1. MLST-profiles
2. Isolates
3. Allelic sequences (fastA, Mega)

strain	ST	country	region	year	Age (a)	sex	disease	source	epidemiology	methicillin
H52	188	UK	Oxford	1997	80	female	unspecified	blood	hospital	S
C38	1	UK	Oxford	1997	60	male	unspecified	blood	endemic	S
C476	1	UK	Oxford	1998	9	male	unspecified	blood	endemic	S
C71	1	UK	Oxford	1997	88	female	unspecified	blood	endemic	S
C98	1	UK	Oxford	1997	76	female	unspecified	blood	endemic	S
H462	1	UK	Oxford	1997	72	female	unspecified	blood	hospital	S
H512	1	UK	Oxford	1998	24	male	unspecified	blood	hospital	S
C162	1	UK	Oxford	1997	2	male	unspecified	blood	endemic	S
H148	3	UK	Oxford	1997	22	male	unspecified	blood	hospital	S

# MLST database

1. MLST-profiles
2. Isolates
3. Allelic sequences (fastA, Mega)

```
>arcc1
```

```
TTATTAATCCAACAAGCTAAATCGAACAGTGACACAACGCCGGCAATGCCATTGGATACTTGTGGTGCAATGTCACAGGGTATGAT  
AGGCTATTGGTTGGAAACTGAAATCAATCGCATTTTAACTGAAATGAATACTGATAGAACTGTAGGCACAATCGTTACACGTGTGG  
AAGTAGATAAAGATGATCCACGATTTCAATAACCCAACCAACCAATTGGTCCTTTTTATACGAAAGAAGAAGTTGAAGAATTACAA  
AAAGAACAGCCAGACTCAGTCTTTAAAGAAGATGCAGGACGTGGTTATAGAAAAGTAGTTGCGTCACCACTACCTCAATCTATACT  
AGAACACCAGTTAATTTCGAACCTTTAGCAGACGGTAAAAATATTGTCATTGCATGCGGTGGTGGCGGTATTCCAGTTATAAAAAAAG  
AAAATACCTATGAAGGTGTTGAAGCG
```

```
>arcc2
```

```
TTATTAATCCAACAAGCTAAATCGAACAGTGACACAACGCCGGCAATGCCATTGGATACTTGTGGTGCAATGTCACAAGGTATGAT  
AGGCTATTGGTTGGAAACTGAAATCAATCGCATTTTAACTGAAATGAATAGTGATAGAACTGTAGGCACAATCGTAACACGTGTGG  
AAGTAGATAAAGATGATCCACGATTTGATAACCCAACCAACCAATTGGTCCTTTTTATACGAAAGAAGAAGTTGAAGAATTACAA  
AAAGAACAGCCAGGCTCAGTCTTTAAAGAAGATGCAGGACGTGGTTATAGAAAAGTAGTTGCGTCACCACTACCTCAATCTATACT  
AGAACACCAGTTAATTTCGAACCTTTAGCAGACGGTAAAAATATTGTCATTGCATGCGGTGGTGGCGGTATTCCAGTTATAAAAAAAG  
AAAATACCTATGAAGGTGTTGAAGCG
```

```
>arcc3
```

```
TTATTAATCCAACAAGCTAAATCGAACAGTGACACAACGCCGGCAATGCCATTGGATACTTGTGGTGCAATGTCACAGGGTATGAT  
AGGCTATTGGTTGGAAACTGAAATCAATCGCATTTTAACTGAAATGAATAGTGATAGAACTGTAGGCACAATCGTTACACGTGTGG  
AAGTAGATAAAGATGATCCACGATTTGATAACCCAACCAACCAATTGGTCCTTTTTATACGAAAGAAGAAGTTGAAGAATTACAA  
AAAGAACAGCCAGACTCAGTCTTTAAAGAAGATGCAGGACGTGGTTATAGAAAAGTAGTTGCGTCACCACTACCTCAATCTATACT  
AGAACACCAGTTAATTTCGAACCTTTAGCAGACGGTAAAAATATTGTCATTGCATGCGGTGGTGGCGGTATTCCAGTTATAAAAAAAG  
AAAATACCTATGAAGGTGTTGAAGCG
```

# Data analysis using the MLST website

**MLST** Multi Locus Sequence Typing  
mlst.net | saureus.mlst.net  
Monday 16th July 2007

**DATA ANALYSIS**

**DATABASIS**

- ▣ *B.cereus*
- ▣ *B.pseudomallei*
- ▣ *C.albicans*
- ▣ *C.glabrata*
- ▣ *C.krusei*
- ▣ *C.tropicalis*
- ▣ *C.jejuni*
- ▣ *C.neoformans var grubii*
- ▣ *E.coli*
- ▣ *E.faecalis*
- ▣ *E.faecium*
- ▣ *H.influenzae*
- ▣ *H.pylori*
- ▣ *M.catarrhalis*
- ▣ *N.meningitidis*
- ▣ *S.agalactiae*
- ▣ *S.aureus*
- ▣ *S.enterica*
- ▣ *S.epidemicus*
- ▣ *S.pneumoniae*
- ▣ *S.pyogenes*
- ▣ *S.suis*
- ▣ *V.vulnificus*

**SUBMISSIONS**

**NEW - MLST-maps - Click here - View global distribution of isolates and STs - NEW**

*Staphylococcus aureus*

- ▣ Organism Specific Information
- ▣ Concatenate Sequences
- ▣ Download Alleles
- ▣ Download ST's
- ▣ Compare profile to reset
- ▣ Draw tree using own MLST data
- ▣ download as MS Excel
- ▣ download as MS Access
- ▣ *S.aureus* Links
- ▣ eBURST V3
- ▣ Contact Curator

**Profile Query**      **Locus Query**      **Batch Query**

Please choose---      Please choose---      Please choose---

This site is hosted at Imperial College and development is funded by the Wellcome Trust.  
For comments, queries, or suggestions please contact David Aanensen

> 40 MLST schemes for different organisms

<http://www.mlst.net>

# Data analysis using the MLST website



The screenshot shows the MLST website interface. The header includes the MLST logo, the text "Multi Locus Sequence Typing", and the URL "mlst.net | saureus.mlst.net". The date "Monday 16th July 2007" is also visible. The main content area is titled "Staphylococcus aureus" and includes a "Please Select Locus:" dropdown menu with "arcc" selected. Below this is a text input field for entering sequences in FASTA or MEGA format. The input field contains the following sequence:

```
>11#2006_arccSa
TTATTAATCCAACAGCTAAATCGAACAGTGACAC
AACGCCGGCAATGCCATTGGATACATGTGGTCCA
TGTACACAGGGTATGATAGGCTATTGGTTGGAACT
GAAATCAATCGCATTTTAACTGAAATGAAATGTA
TAGAACTGTAGGCACAATCGTTACACGTTGGGAG
TAGATAAAGATGATCCACGATTTGATACCCAACT
AAACCAATTGGTCCTTTTACACGAAAGAGAAGT
TGAAGAATTACAAAAGACAGCCCACTCAGTCT
TTAAGAAGATGCAGGACGTGGTTATAGAAAAGTA
GTTGCGTCACCACTACCTCAATGTAATACTAGAACA
CCAGTTAATCGAATTTAGCAGACGGTAAATA
TTGTCATTGCATGCGGTGGTGGCGGTATTCCAGTT
ATAAAAAGGAAATACTATGAAGGTGTTGAAGC
G
```

Below the input field are "Reset" and "Submit" buttons. At the bottom of the page, there are three dropdown menus for "Profile Query", "Locus Query", and "Batch Query", each with "Please choose..." selected.

Copy & paste your allelic sequence into the web interface

<http://www.mlst.net>

# Data retrieval using the MLST website

The screenshot displays the MLST website interface. At the top left, the logo reads "MLST Multi Locus Sequence Typing" with a DNA double helix icon. To the right, there are three small images: a gel electrophoresis pattern, a tree, and a laboratory flask. The URL "mlst.net | saureus.mlst.net" and the date "Monday 16th July 2007" are visible in the top right.

The main content area is titled "Staphylococcus aureus - batch results". Below this, it shows "112006\_arcCSa is allele 14". There are three query options: "Profile Query", "Locus Query", and "Batch Query". Each option has a dropdown menu with the text "Please choose---".

On the left side, there is a "DATA ANALYSIS" section and a "DATABASES" section. The "DATABASES" section lists several species with checkboxes:

- B.cereus*
- B.pseudomallei*
- C.albicans*
- C.glabrata*
- C.krusei*
- C.tropicalis*
- C.jejuni*
- C.neoformans var grubii*
- E.coli*
- E.faecalis*
- E.faecium*
- H.influenzae*

# Data analysis using the MLST website

**MLST** Multi Locus Sequence Typing  
mlst.net | saureus.mlst.net Monday 16th July 2007

**DATA ANALYSIS**

**DATABASES**

- B.cereus*
- B.pseudomallei*
- C.albicans*
- C.glabrata*
- C.krusei*
- C.tropicalis*
- C.jejuni*
- C.neoformans var grubii*
- E.coli*
- E.faecalis*
- E.faecium*
- H.influenzae*
- H.pylori*
- M.catarrhalis*
- N.meningitidis*
- S.agalactiae*
- S.aureus*
- S.enterica*
- S.epidermidis*
- S.pneumoniae*
- S.pyogenes*
- S.suis*
- V.vulnificus*

**SUBMISSIONS**

**NEWS**

**Staphylococcus aureus - Allelic Profile Query** ?

Please enter your query below (the figure below each box represents the number of unique alleles in the current database for that locus).

arcc	aroe	glpf	gmk_	pta	tpi_	yqil
<input type="text" value="1"/>	<input type="text" value="1"/>	<input type="text" value="1"/>	<input type="text" value="1"/>	<input type="text" value="1"/>	<input type="text" value="1"/>	<input type="text" value="1"/>
116	160	126	89	120	123	112

Please select Query type:



Please choose which dataset to interrogate:  
 or

**Profile Query**      **Locus Query**      **Batch Query**

Copy & paste your allelic profiles into the web interface

# Data retrieval using the MLST website

MLST  
Multi Locus Sequence Typing  
mlst.net | saureus.mlst.net  
Monday 16th July 2007

**DATA ANALYSIS**

**DATABASES**

- B.cereus*
- B.pseudomallei*
- C.albicans*
- C.glabrata*
- C.krusei*
- C.tropicalis*
- C.jejuni*
- C.neoformans var grubii*
- E.coli*
- E.faecalis*
- E.faecium*
- H.influenzae*
- H.pylori*
- M.catarrhalis*
- N.meningitidis*
- S.agalactiae*
- S.aureus***
- S.enterica*
- S.epidermidis*
- S.pneumoniae*
- S.pyogenes*
- S.suis*
- V.vulnificus*

**SUBMISSIONS**

**Staphylococcus aureus** Allelic Profiles query results

Your sequence type is 1

Strain	ST	Spa Type	arcc	aroe	glpf	gmk_	pta_	tpi_	yqil
			1	1	1	1	1	1	1
C26	1		1	1	1	1	1	1	1
C478	1		1	1	1	1	1	1	1
C71	1		1	1	1	1	1	1	1
C98	1		1	1	1	1	1	1	1
H462	1		1	1	1	1	1	1	1
H512	1		1	1	1	1	1	1	1
C162	1		1	1	1	1	1	1	1
D525	1		1	1	1	1	1	1	1
D988	1		1	1	1	1	1	1	1
CAN6213-0112	1		1	1	1	1	1	1	1
CAN7819-1066	1		1	1	1	1	1	1	1
CAN7819-1083	1		1	1	1	1	1	1	1
NL011851-2	1		1	1	1	1	1	1	1
Not279	1		1	1	1	1	1	1	1
Not356	1		1	1	1	1	1	1	1
Not382	1		1	1	1	1	1	1	1
UKCARR-1	1		1	1	1	1	1	1	1
E3410	1		1	1	1	1	1	1	1
Btn2164	1		1	1	1	1	1	1	1

<http://www.mlst.net>

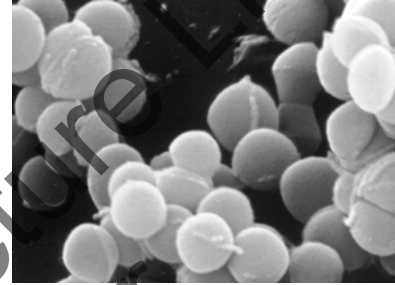
# Singlelocus Sequence Typing (SLST)

ESCMID Online Lecture Library  
© by author



# Examples for SLST usage

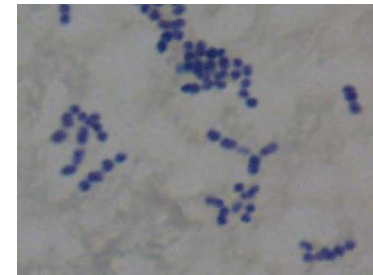
- *S. aureus* protein A gene typing (*spa*)



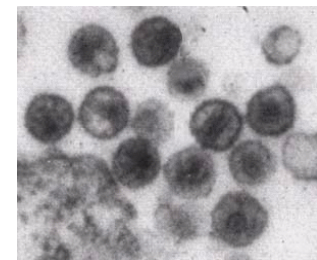
- *C. jejuni* flagellin gene typing (*fla*)

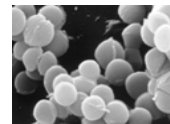


- *S. pyogenes* M protein gene typing (*emm*)

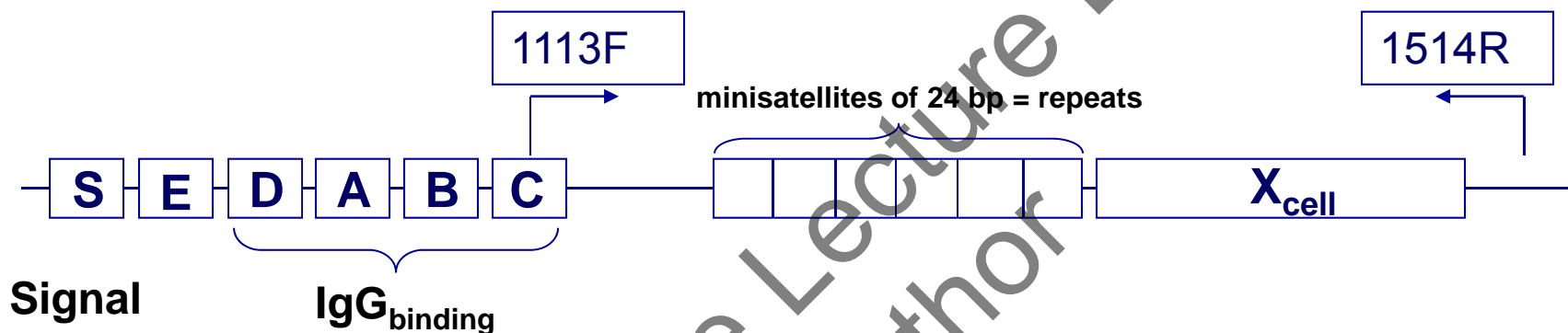


- HIV (protease / reverse transcriptase genes)





# *S. aureus* protein A gene (*spa*)



Frénay et al., *Eur. J. Clin. Microbiol. Infect. Dis.* (1996) [[Pubmed](#)]

Shopsin et al., *J. Clin. Microbiol.* (1999) [[Pubmed](#)]

Repeat succession = *spa* type

VNTR = variable number of tandem repeats (multiple VNTR = MLVA)

- singlelocus sequence typing (SLST)
- fast (< 2 d) and cheap in comparison to MLST or PFGE
- *spa* is stable and highly discriminatory (outbreak detection)

# *spa* sequence data analysis

## StaphType Software

- sequence editor with automated repeat detection
- integrated database
- client / server (via WWW or email)  
synchronization of nomenclature and additional epidemiological data
- calculation of relatedness of *spa* types

Ridom StaphType (0.9.32)

File Tools Windows Help

**Retrieval** Automatic spa-repeat and -type detection

007\_sp\_95

Found 8 repeats: 5' r01 r02 r03 r04 r05 r06 r07 r08 3'

5'- and 3'-signature search

per base estimate of error probability

Spa-typing reliability is good

Integrated expert system

Ridom StaphType version 1.0 (Win 98, NT, 2000, XP; Mac OS X or Linux) <http://www.ridom.de/>

Relational database for personal and other epidemiological relevant data with **strong 128-bit encryption**

Repeat Succession: r01r03r05r03r06r03r07 Spa Type: t003, freq:11.38 %

Cursor Base: 47 with Quality: 27

The screenshot displays the Ridom StaphType software interface. At the top, the title bar reads 'Ridom StaphType (0.9.32)'. Below it is a menu bar with 'File', 'Tools', 'Windows', and 'Help'. The main window is titled '007\_sp\_95' and shows a sequence alignment of a DNA fragment. The sequence is displayed in a grid format with columns numbered from 110 to 245. Above the sequence, a bar indicates the presence of 8 repeats, labeled r01 through r08, with arrows pointing to their positions. A '5'- and 3'-signature search' is indicated by a red oval around a specific sequence motif. Below the sequence, chromatograms show the signal for each base. A 'New Strain' dialog box is open in the center, allowing the user to enter strain information such as 'Strain Source', 'Laboratory ID', 'Isolation date', 'Receipt date', 'Care unit', 'Clinic', 'Ward', 'Room', 'Origin', 'Specimen', 'Isolate deposit', 'ttt', and 'Alternative lab-ID'. The dialog box has 'Save' and 'Cancel' buttons at the bottom. A 'per base estimate of error probability' is shown as a line graph below the chromatograms. On the left side, a table lists 'Repeat Succession' for various strains. A text box titled 'Spa-typing reliability is good' provides information about uncertain positions in the repeats. At the bottom of the window, the 'Repeat Succession' is shown as 'r01r03r05r03r06r03r07' and the 'Spa Type' is 't003, freq:11.38 %'. The 'Cursor Base' is 47 and the 'with Quality' is 27.

Ridom StaphType (0.9.32)

File Tools Options Windows Help

Strain Database Browser

All Strains Search Result "Probe" Search Result "tesst"

Laboratory ID	SpaType ID	Strain Isolation date	Strain Origin	Strain Specimen source	Person Birthday	Person related data Clinic...	Care unit Name	Care unit City	Clinic Name	Ward Name
001_spa_1000_93		07-Jun-2003	person				reference strain	unknown		
001_spa_1026_st239f	t025	10-May-2003	person				reference strain	unknown		
001_spa_1150_93							reference strain	unknown		
001_spa_1155_98							reference strain	unknown		
001_spa_131_96							reference strain	unknown		
001_spa_134_93							reference strain	unknown		
001_spa_1450_94							reference strain	unknown		
001_spa_1492_02f							reference strain	unknown		
001_spa_1869_98							reference strain	unknown		
001_spa_2511_02							reference strain	unknown		
001_spa_2704_02f							reference strain	unknown		
001_spa_2811_02f							reference strain	unknown		
001_spa_3179_02							reference strain	unknown		
001_spa_3253_02							reference strain	unknown		
001_spa_3391_02f							reference strain	unknown		
001_spa_3441_02f							reference strain	unknown		
001_spa_3500_02f							reference strain	unknown		
001_spa_3531_02							reference strain	unknown		
001_spa_3534_02							reference strain	unknown		
001_spa_3600_st2c							reference strain	unknown		
001_spa_405_st236							reference strain	unknown		
001_spa_4209_01							reference strain	unknown		
001_spa_421_03_11							reference strain	unknown		
001_spa_54518_st2							reference strain	unknown		
001_spa_61974_st5							reference strain	unknown		
001_spa_635_93f							reference strain	unknown		
001_spa_96158_st6							reference strain	unknown		
001_spa_96_32010							reference strain	unknown		
001_spa_97118_st2							reference strain	unknown		
001_spa_97151_st6							reference strain	unknown		
001_spa_97393_st2							reference strain	unknown		
001_spa_97s101_st							reference strain	unknown		
001_spa_97s96_st2							reference strain	unknown		
001_spa_97s99_st4							reference strain	unknown		
001_spa_98541_st241	t023	26-May-2003	person				reference strain	unknown		
001_spa_c0941787_st45f	t036	10-May-2003	person				reference strain	unknown		
001_spa_hpv107_st247f	t037	10-May-2003	person				reference strain	unknown		
001_spa_hs2_st247	t038	11-May-2003	person				reference strain	unknown		
001_spa_hsj216_st239	t025	11-May-2003	person				reference strain	unknown		
001_spa_m307_st5	t001	11-May-2003	person				reference strain	unknown		
001_spa_mr1_st5	t039	11-May-2003	person				reference strain	unknown		
001_spa_nctc11939_st239	t025	11-May-2003	person				reference strain	unknown		
007_spa_100_f	tx167	22-Jul-2003	person	skin	03-Nov-1920		UKM	Münster	HAUT	AMB
007_spa_102_f	t003	24-Jul-2003	person	genital	15-Apr-1965		UKM	Münster	CHIRALL	Chirurgie 5
007_spa_1678_96_5	t018	02-Jun-2003	person				reference strain	unknown		
007_spa_2512_02_23	t003	02-Jun-2003	person				reference strain	unknown		
007_spa_281_2002	t003	19-Feb-2003	person				UKM	Münster		
007_spa_288_2002	t001	01-Jan-2002	person				UKM	Münster		
007_spa_802_2003	t002	05-Mar-2003	person	nose	18-Feb-1980		UKM	Münster	MEDA	13B OST
007_spa_82_f	tx168	06-Jun-2003	person	nose	30-Mar-1934		UKM	Münster	THGCH	18B OST
007_spa_88	tx167	20-Jun-2003	person	fluid unknown	09-Sep-1925		UKM	Münster	ORTH	13B WEST
007_spa_89_f	t003	10-Jul-2003	person	nose	26-Sep-1963		UKM	Münster	CHIRALL	Chirurgie 5
007_spa_8_35_2002_f	t003	19-Feb-2003	person				UKM	Münster		
007_spa_90_416_2002	t020	22-Mar-2003	person				Klinikum Lippe	Detmold		
007_spa_92_r	t002	14-Jul-2003	person	nose	08-Oct-1980		UKM	Münster	CHIRALL	Chirurgie 6
007_spa_93_f	t003	14-Jul-2003	person	nose	01-May-1975		UKM	Münster	CHIRALL	Chirurgie 6
007_spa_95	t003	06-Jul-2003	person	nose	19-Jul-1934		UKM	Münster	CHIRALL	Chirurgie 9
007_spa_A9465	t004	05-Jul-2003	person	eye	29-Aug-1931		UKM	Münster	AUGE FRAUEN	
007_spa_K03_2388	t003	10-Jul-2003	person	nose			UKM	Münster	CHIRALL	Chirurgie 6

Ridom SpaServer:  
Relative frequencies of the Spa-Type occurrences:

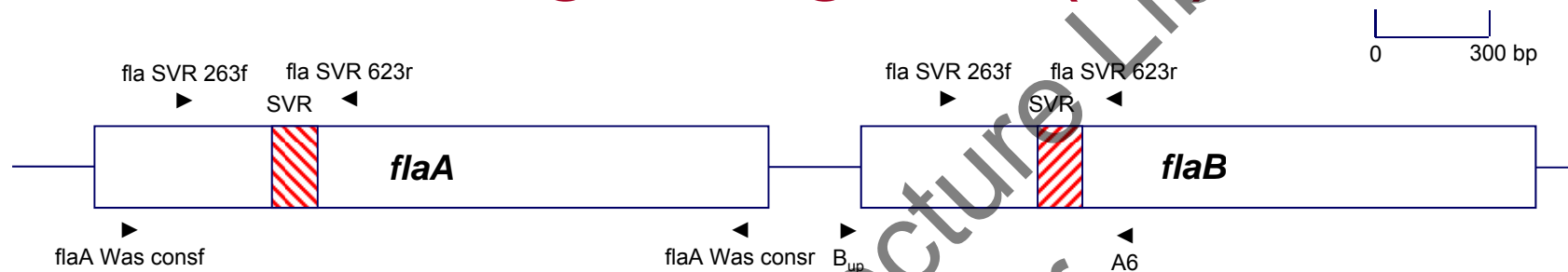
SpaType	Frequency	Country of Origin
t003	13.59 %	Germany
t004	10.38 %	Germany
t001	9.47 %	Germany
t032	7.94 %	Germany
t008	5.57 %	Germany
t037	4.50 %	Germany
t036	4.12 %	Germany
t009	3.74 %	Germany
t002	2.60 %	Germany
t030	2.14 %	Germany
t038	2.06 %	Germany
t015	1.83 %	Germany
t023	1.68 %	Germany

**Synchronization**  
The WWW site (<http://www.ridom.de/spaserver/>) is accessible for everyone and the spa-repeats and the types can be downloaded. In addition it is also possible to submit chromatograms of new spa-repeats and/or -types (open system)

Ridom StaphType  
version 1.0  
(Win 98, NT, 2000, XP; Mac OS X or Linux)  
<http://www.ridom.de/>

Found Strains: 1000+ Strains in DB: 1028

# Flagellin gene (*fla*)



SVR: short variable region

mod. from Meinersmann et al., *J. Clin. Microbiol.* (1997) [[Pubmed](#)]

- *flaA* essential for mobility and pathogenicity
- *flaB* not essential and is thought to be a genetic reservoir for antigenic variation
- long-term typing of *flaA* questioned because of known intra- and inter-genomic recombination within the flagellin genes

# Flagellin gene (*fla*) typing of *C. jejuni*

FlaA

The Campylobacter FlaA database

FlaA

[Database home](#)  
[Sequence query](#)  
[Batch query](#)  
[Download seqs](#)

This website is powered by  
the agdbNet software (Jolley  
& Maiden 2006, BMC  
Bioinformatics, 7:314).

## The *Campylobacter* FlaA Variable Region Database

You can query this database for FlaA variable region alleles and peptide sequences. If an allele or peptide does not exist in the database, a BLAST search will be performed, showing you which sequences it is most like - this also allows you to paste in sequences which are larger than the FlaA variable region.

- [Access database](#)

If you would like to submit FlaA sequences for assignment and addition to the database, please send forward and reverse trace files to [Frances Colles](#) and [Alison Cody](#).

### Related Links

- [Campylobacter MLST database](#)

<http://hercules.medawar.ox.ac.uk/flaA/>

# Flagellin gene (*fla*) typing of *C. jejuni*

FlaA

The Campylobacter FlaA database

FlaA

Database home  
Sequence query  
Batch query  
Download seqs

This website is powered by  
the agdbNet software (Jolley  
& Maiden 2006, BMC  
Bioinformatics, 7:314).

## Single sequence query

Query nucleotide sequence against both nucleotide and peptide tables for flaA.

Locus	Query against	Results
FlaA	nucleotide and peptide	Identical match - nucleotide allele 1. Identical match - peptide(s) 67. (BLAST results)

## Information on FlaA allele 1

id	1
sequence	AAAGCTACTCAAGCGGCTCAAGATGGTCAAAGTCTAAAAACAAGAAGTATGCTTCAAGCA GATATAAACAAACTCATGGAAGAGCTTGATAATATCGCAAATACTACTTCATTTAATGGT AAACAACCTTTAAGCGGTGGTTTCACCAATCAAGAATTTCAAATCGGTGCAAGTTCAAAT CAACAGTAAAAGCAAAGATAGGTGCAACTCAGTCTTCAAAAATCGGTGTAAGTACTAGATTT GAAACAGTTTCACAAAGCTTACTTCAGGTGTAGTGGGACTTACTATTAAAACTATAAT GGTATAGAAGATTTTAAATTT
peptide id	67
sender	Kate Dingle Nuffield Department of Clinical Sciences, Oxford University, UK
dates amp	2003-10-07
curator	Keith Jolley University of Oxford <a href="mailto:keith.jolley@medawar.ox.ac.uk">keith.jolley@medawar.ox.ac.uk</a>

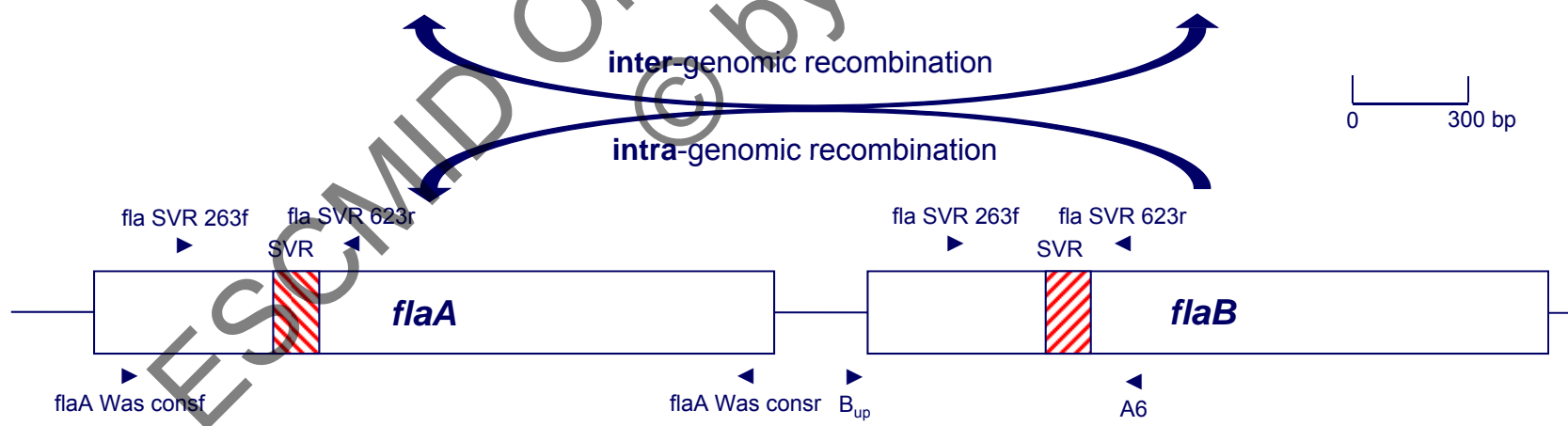
<http://hercules.medawar.ox.ac.uk/flaA/>



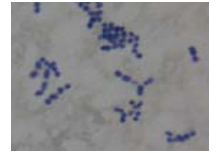


# flaB typing

- In two outbreaks, *flaA* was over-discriminatory in some strains tested compared to the other employed typing schemes.
- These isolates had several nucleotide substitutions distributed throughout the SVR, suggesting that the change was due to **intra-genomic** recombination rather than to spontaneous mutation.



# M protein gene (*emm*) typing of *S. pyogenes*



CDC Home | [About CDC](#) | [Press Room](#) | [A-Z Index](#) | [Contact Us](#)

**CDC** Department of Health and Human Services  
Centers for Disease Control and Prevention

CDC en Español

Search:

## *Streptococcus pyogenes emm* sequence database

**Topic Contents**

- > [Topic Home](#)
- > [CDC \*Streptococcus Lab\*](#)
- > [Introduction to \*emm\* typing](#)
- > [Assigning \*emm\* types and subtypes](#)
- > [Blast-\*emm\*](#)
- > [Browse M protein gene types](#)
- > [\*emm\* type frequencies in invasive U.S. isolates \(2003-2004\)](#)   
55KBS
- > [Protocols](#)
  - [emm typing](#)
  - [PCR deduction of](#)

[Home](#) > [CDC Streptococcus Lab](#) > M protein gene (*emm*) typing *Streptococcus pyogenes*

### M protein gene (*emm*) typing *Streptococcus pyogenes*

#### Introduction

Group A streptococcal mediated disease continues to be a major problem world-wide. There are millions of cases of GAS pharyngitis causing billions of dollars in medical expenses and work stoppage in the United States alone. Approximately 10,000 - 15,000 cases of invasive GAS disease occur annually in the United States, associated with a 10-10% mortality rate. Even though the Lancefield M protein serotyping system over the past 60 years has been very valuable, in recent years the inherent difficulties encountered in expanding this system through conventional serologic procedures have become increasingly evident. Using a less demanding sequence based system that is predictive of Lancefield M serotypes, we have extended the system established decades ago by Dr. Rebecca Lancefield.

#### Links for General Information

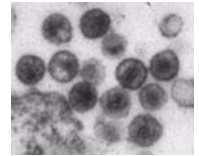
- [Email this page](#)
- [Printer-friendly version](#)

#### Contact Information

For problems, comments, or questions regarding these web pages please contact Bernard Beall ([BBEALL@CDC.GOV](mailto:BBEALL@CDC.GOV)).

[http://www.cdc.gov/ncidod/biotech/strep/M-ProteinGene\\_typing.htm](http://www.cdc.gov/ncidod/biotech/strep/M-ProteinGene_typing.htm)

# Partial protease gene / reverse transcriptase gene typing of HIV



**STANFORD UNIVERSITY**  
**HIV DRUG RESISTANCE DATABASE**  
*A curated public database designed to represent, store, and analyze the divergent forms of data underlying HIV drug resistance.*

**HOME GENOTYPE-RX GENOTYPE-PHENO GENOTYPE-CLINICAL HIVdb PROGRAM**

**Three new programs launched: ART-AiDE, eCARE, and CPR**

Antiretroviral Therapy - Acquisition and Display Engine (ART-AiDE) makes it possible to generate a permanent electronic and graphical record of a patient's antiretroviral treatment (ART) history, plasma HIV-1 RNA levels, ... [More »](#)

**SDRM**  
Susceptibility

**HIVdb PROGRAM** Genotype Resistance Interpretation

This program interprets user-entered mutations to infer the level of resistance to NRTIs, NNRTIs, PIs. Web Service is available.

**ART-AiDE**

Antiretroviral Therapy - Acquisition & Display Engine  
» [Go To Program](#)

**HIVseq Program**

Provides mutation frequencies by subtype.  
» [Go To Program](#)

**HIValg Program**

Compare HIVdb, ANRS, Rega, or create your own algorithm.

**GENOTYPE-TREATMENT CORRELATIONS**

- Retrieve sequences (and/or mutations) from persons receiving selected HIV drugs
- Retrieve sequences and treatments from viruses with specific mutations

**GENOTYPE-CLINICAL CORRELATIONS**

- Summaries of genotype-clinical outcome studies
- Genotype-clinical outcome datasets (download)

**GENOTYPE-PHENOTYPE CORRELATIONS**

- Retrieve drug susceptibility data for isolates with selected mutations
- Download genotype-phenotype research datasets

**REFERENCES**

- Published drug resistance studies in HIVRT&PrDB
- Published studies by Stanford database group

<http://hivdb.stanford.edu>

Rhee et al., *PNAS* (2006) [[Pubmed](#)]

# Conclusions

- MLST/SLST of various microorganisms has proven to be a reliable and discriminatory typing method
- Sequence-based typing greatly simplifies the intra- and inter-laboratory exchange of data
- In future, new targets for DNA-sequencing will supplement or replace today's methods in clinical microbiology

# MLST & SLST: Technical Background

**Dag Harmsen**

University of Münster, Germany

[dharmsen@uni-muenster.de](mailto:dharmsen@uni-muenster.de)