ESGLI ESCMID STUDY GROUP FOR LEGIONELLA INFECTIONS

European Society of Clinical Microbiology and Infectious Diseases

Sequence-Based Typing protocol for epidemiological typing of *Legionella pneumophila*

VERSION 5.0

SUMMARY

This procedure describes the European Legionnaires' Disease Surveillance Network (ELDSNet) method for Sequence-Based Typing (SBT) of isolates of *Legionella pneumophila.* Genomic DNA is extracted and then amplified using primers targeting seven specific gene loci (i.e., *flaA*, *pilE*, *asd*, *mip*, *mompS*, *proA, neuA*). Following purification, amplicons are sequenced directly with forward and reverse primers, and the resulting consensus sequences trimmed and compared to previously assigned allele numbers using the *L. pneumophila* database. According to a pre-determined order (i.e., *flaA*, *pilE*, *asd*, *mip*, *mompS*, *proA, neuA*), the combination of alleles is defined as 7-digit allelic profile (e.g. 1,4,3,1,1,1,1) and a Sequence Type (ST) represented by a number (e.g., ST1). Putative new allele types can be submitted following the instruction on the website (???). This method can be used in the epidemiological typing of *L. pneumophila*.

SBT TARGETS

All 7 loci (i.e., *flaA*, *pilE*, *asd*, *mip*, *mompS*, *proA, neuA*) should be determined if possible. However it has been described that the *neuA* target is not present in some non-serogroup 1 strains of *L. pneumophila* but a *neuA* homolog (*neuAh*) can be amplified and sequenced (Farhat *et al.*, 2011). In case of failed amplification of *neuA* in non sg-1 strains, refer to the *neuAh* protocol.

Legionella cultures

Procedures which involve handling of live microorganisms in liquid suspensions (likely to create aerosols) must be carried out in a Microbiological Safety Cabinet wearing latex gloves. Prior to SBT, *Legionella pneumophila* growth must be harvested from plates or slopes using a sterile loop and gently transferring 1-2 loopfuls into a labelled 1.5 ml microcentrifuge tube containing ca. 1 ml sterile water. Cap tube. Using a centrifuge with sealed rotor, centrifuge the tube at 15,000 $\times g$ for ca. 10 minutes to pellet cells. In the Class 1 Safety Cabinet, open tube, remove and discard supernatant. Cell pellets maybe processed directly or stored at -70 to -80ºC.

DNA EXTRACTION

The following methods have been found to be reliable and the first two are rapid:

1. BIO-RAD Instagene Matrix (Catalogue n. 732-6030), using 1-5 µl of supernatant as template DNA in the primary PCR amplification.
2. Heating emulsified colonies in 0.5 ml of sterile water at 100°C for 8 min, using 1-5 µl of supernatant as template DNA in the primary PCR amplification.
3. Nucleon BACC2 DNA extraction kit (Gen-Probe), using 10-100 ng DNA as template DNA in the primary PCR amplification.

Amplification primers:

The positions of the primers with respect to the Genebank accession number of the reference sequence are shown below:

Gene	Primer name	Position	Primer sequence (5'-3')	Annealing temperature	Fragment size (bp)
flaA	flaA-587F	568-587	GCG TAT TGC TCA AAA TAC TG	55 °C	414
	flaA-960R	981-960	CCA TTA ATC GTT AAG TTG TAG G		
pilE	pilE-35F	12-35	CAC AAT CGG ATG GAA CAC AAA CTA	55 °C	460
	pilE-453R	471-453	GCT GGC GCA CTC GGT ATC T		
asd	asd-511F	487-511	CCC TAA TTG CTC TAC CAT TCA GAT G	55 °C	576
	asd-1039R	1062-1039	CGA ATG TTA TCT GCG ACT ATC CAC		
mip	mip-74F	58-74	GCT GCA ACC GAT GCC AC	55 °C	559
	mip-595R	616-595	CAT ATG CAA GAC CTG AGG GAA C		
mompS	mompS-450F	430-450	TTG ACC ATG AGT GGG ATT GG	55 °C	711
	mompS-1116R	1140-1116	TGG ATA AAT TAT CCA GCC GGA CTT C		
proA	proA-1107F	1090-1107	GAT CGC CAA TGC AAT TAG	55 °C	481
	proA-1553R	1570-1553	ACC ATA ACA TCA AAA GCC		
neuA	neuA-196F	176-196	CCG TTC AAT ATG GGG CTT CAG	55 °C	459
	neuA-634R	634-611	CGA TGT CGA TGG ATT CAC TAA TAC		

Sequencing primers:

As above except that mompS-1015R instead of mompS-1116R is used for the reverse sequencing reaction of the mompS target.

Primer name	Position	Primer sequence (5'-3')
mompS-1015R	1032-1015	CAG AAG CTG CGA AAT CAG

mompS primary amplification:
mompS-450F: 5' - TTG ACC ATG AGT GGG ATT GG - 3'
mompS-1116R: 5' - TGG ATA AAT TAT CCA GCC GGA CTT C - 3'

mompS sequencing:
mompS-450F: 5' - TTG ACC ATG AGT GGG ATT GG - 3'
mompS-1015R: 5' - CAG AAG CTG CGA AAT CAG - 3'

In the event of using the M13 modified SBT protocol, refer to APPENDIX 1.

Reference sequences and regions used for allele assignment:

Target	Amplicon size	GenBank accession no. of reference sequence	Target region included for allele asignment	Size of target region (nt)
flaA	414	X83232	653-834	182
pilE	460	AF048690	103-435	333
asd	576	AF034213	538-1010	473
mip	559	AJ496269	117-518	402
mompS	711	AF078136	523-874	352
proA	481	M32884	1134-1538	405
neuA	459	AJ007311	229-582	354

Oligonucleotide storage and handling recommendations

Appropriate storage of oligonucleotides will extend their shelf life. Most oligonucleotides are supplied dry (lyophilised) and manufacturers usually supply detailed information on appropriate resuspension and storage conditions. Information on data-sheets and instructions from the manufacturers provided with the oligonucleotides should be read, noted and followed. To maximise the shelf-life of oligonucleotides it is recommended that “stock solutions” (e.g., 100 pmol/µl) are made in 1x nuclease-free TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 7.5-8.0) and stored below -20ºC, and working solutions (e.g., 10 pmol/µl) are made in 1x nuclease-free Tris (10 mM Tris pH 7.5-8.0). Working solutions are frozen in 10-50 µl aliquots and should **NOT** be subjected to repeat freeze-thaw cycles and must be kept for **no longer than one week at 4ºC** before discarding. Working solutions for primers used for sequencing reactions should be made up fresh on the day of use, then discarded.

Primary amplification

PCR amplification is performed in a final volume of 20 µl. Place PCR plate on a cold block and start with adding to each well 4 µl of the Taq DNA polymerase Mix, then add 14 µl of the relevant Reaction Mix. Pipette solutions in the order of: negative

control, test samples, positive control. Gently centrifuge plates briefly in an appropriate centrifuge before placing them in the thermal cycler.

Taq DNA polymerase Mix:

Nuclease-free water	3.6 µl
PCR buffer (x10)	0.2 µl
Taq DNA polymerase (5U/µl)	0.2 µl
Volume/reaction	4.0 µl

Reaction Mix (prepare 7 Reaction Mixes, one for each target):

Reagent	Stock concentration	Final concentration	Volume/reaction (µl)
Nuclease-free water			9.4
PCR buffer	10 x	1 x	2.0
$MgCl_2$	50 mM	2.5 mM	1.0
Nested Primer F	10 pmol/µl	4 pmoles	0.4
Nested Primer R	10 pmol/µl	4 pmoles	0.4
dNTPs	5 mM	200µM	0.8
Volume/reaction			14.0

Controls

For each PCR run include a negative (2 µl of nuclease-free water) and a positive control (2 ul of known *L. pneumophila* DNA template). Pipette solutions in the order of: negative control, test samples, positive control.

Thermal cycler parameters

Step	Temperature (ºC)	Time		No. cycles
1	95	5 min	Initial denaturation	1
2	95	30 sec	Denaturation	35
	55	30 sec	Annealing	
	72	45 sec	Extension	
3	72	10 min	Final Extension	1
4	12	∞	Hold	

Analysis of PCR products by gel electrophoresis

Run PCR products (2 µl + 18 µl of 1x loading buffer) with a low range quantitative DNA ladder on 1-2% agarose gel, then analyse results using a UV light system. If positive and negative controls as well as the test sample results are satisfactory, proceed with the DNA sequencing procedure.

DNA Sequencing

PCR products can be purified and sequenced in house according to the instructions provided by the manufacturer of the Sequencing Instrument or sent to a commercial sequencing provider.

INTERPRETATION OF RESULTS

Sequence Quality Tool

Forward and reverse sequence trace files (.scf or .abi) are submitted to the "Sequence Quality Tool" (www.hpa-bioinformatics.org.uk/cgi-bin/legionella/sbt/seq_assemble_legionella1.cgi). This tool assembles contig(s) from the traces, finds start and end (reference) positions in the contig, trims the contig using these positions and finally matches the trimmed sequence against those in the SBT database. The tool identifies individual alleles as well as an allelic profile and a sequence type (ST). Sequences with <100% match are identified as -1, and the position of mismatches are also indicated. Sequences of poor quality should either be repeated or further analysed using a DNA analysis software e.g., BioNumerics (Applied Maths).

SBT Database
(www.hpa-bioinformatics.org.uk/legionella/legionella_sbt/php/sbt_homepage.php)

- Data from both forward and reverse sequencing reactions are combined and aligned to produce a consensus sequence. Reference sequences trimmed to the correct length can be downloaded from the website in order to aid contig assembly. Consensus sequences trimmed to the correct length are submitted to the SBT database as flat text file. The database returns an allele type eg., 1, when submitted sequence shows a 100% match to a pre-designated allele type. Sequences with <100% match are identified as the closest match to a pre-existing allele type with the number of mismatches specified. The SBT database also returns an alignment with mismatches highlighted.

- For each isolate, the combination of alleles at each of the loci is defined as the allelic profile using a pre-determined order, i.e., *flaA*, *pilE*, *asd*, *mip*, *mompS*, *proA*, *neuA*. For example, for strain EUL no. 120, the allelic profile is 4,7,11,3,11,12,9.

- If an individual allele number has not been determined, a zero is entered into the allelic profile, thus maintaining its integrity. For example, if the *proA* allele number was not determined for the examples above, the profile would be 4,7,11,3,11,0,9 and if the *mompS* allele was not determined, it would be 4,7,11,3,0,12,9.

Instructions for submitting sequence data to the EWGLI Sequence Quality Tool

- Under Query Functions, click on the "**Sequence Quality Tool**"

- Choose the number of sequence trace files to be uploaded.

- Click on the "**Browse**" button to download sequence files.

- Forward and reverse sequences must be downloaded for each allele in standard file format (*.scf) or ABI trace file format (*.abi).

- Click the "**Analyse Trace Files**" button.

- The tool also produces a sequence quality report for each uploaded contig and generates a seven-figure allelic profile (e.g., 3,4,1,1,14,9,1), as well as indicating whether the obtained allelic profile is of a novel combination. The tool also identifies a single designated sequence type (ST) for a pre-existing allelic profile.

- You can print a summary sheet by clicking the link: "**Print Friendly Summary**"

- For further details on the sequence quality tool please click on this link: "**explanation and frequently asked questions**"

- Putative novel alleles must be submitted for verification following the link "**New Allele Submission**".

- To submit a novel allelic profile click on the link "**Strain Data Submission**".

Instructions for submitting sequence data to the EWGLI SBT Database

- Under Query Functions, click on the "Check a sequence for it's allele number"

- Before pasting in a query sequence the user must ensure that the consensus sequence is trimmed to the correct length.

- Reference sequences trimmed to the correct position and length can be downloaded to aid contig assembly.

- Consensus sequences of the correct length can be submitted for more than one target, alternatively you can submit consensus sequences one at a time before hitting the "upload sequence" button.

- Please note that sequences of incorrect length will be rejected.

- Sequences of correct length and with a 100% match to a sequence in the database will return an allele number e.g., "1".

- Sequences of correct length and with <100% match to a sequence in the database will return an alignment with the closest matches showing where differences are.

- To enter consensus sequences for another strain, click the "Reset form" button and proceed as before.

New Allele Submission

A new allele must be submitted via the New Allele Submission link from the EWGLI SBT database (http://www.hpa-bioinfotools.org.uk/legionella/php/sbt_query1.php). This link allows you to submit forward and reverse sequence trace files to an automated sequence quality tool. If your sequences successfully pass the criteria, you will be asked to submit your data to the curators via the same link.

New Profile Submission

Once a new allele has been submitted the complete profile of the strain must also be submitted via the Submit a new profile link, from the EWGLI SBT database page (http://www.hpabioinfotools.org.uk/legionella/php/sbt_query1.php).

N.B. The website is under constant development, so please see the website for any changes, new tools and to ensure you are using the latest version of this protocol etc.

References

Amemura-Maekawa, J., F. Kura, B. Chang, and Watanabe, H. (2005). *Legionella pneumophila* serogroup 1 isolates from cooling towers in Japan form a distinct genetic cluster. Microbiol Immunol. **49**:1027-1033.

Benson, R. F., C. E. Lucas, E. W. Brown, K. D. Cowgill, and Fields BS. (2006). Molecular comparison of isolates from a recurring outbreak of Legionnaires' dieases spanning 22 years, p. 139-145. *In* N. P. Cianciotto, K. Y. Abu, P. H. Edelstein, B. S. Fields, D. F. Geary, T. G. Harrison, C. A. Joseph, R. M. Ratcliff, J. E. Stout, and M. S. Swanson (eds.), *Legionella*: State of the Art 30 Years after Its Recognition. ASM Press, Washington, DC.

Farhat, C., Mentasti, M., Jacobs, E.,Fry, N.K. and C. Lück. (2011). The N-Acylneuraminate Cytidyltransferase Gene, neuA, Is Heterogenous in Legionella pneumophila Strains but Can Be Used as a Marker for Epidemiological Typing in the Consensus Sequence-Based Typing Scheme. Journal of Clinical Microbiology, 49:4052-8.

Fry, N.K., B. Afshar, G. Wewalka, and Harrison TG. (2006). Epidemiological typing of *Legionella pneumophila* in the absence of isolates, p. 152-155. *In* N. P. Cianciotto, K. Y. Abu, P. H. Edelstein, B. S. Fields, D. F. Geary, T. G. Harrison, C. A. Joseph, R. M. Ratcliff, J. E. Stout, and M. S. Swanson (eds.), *Legionella*: State of the Art 30 Years after Its Recognition. ASM Press, Washington, DC.

Gaia, V., Fry, NK., Afshar, B., Luck, PC., Meugnier, H., Etienne, J., Peduzzi, R. and Harrison, TG. (2005). Consensus sequence-based scheme for epidemiological typing of clinical and environmental isolates of *Legionella pneumophila*. *Journal of Clinical Microbiololgy* 43, 2047-52.

Gaia, V., NK. Fry, Harrison, TG. and Peduzzi, R. (2003). Sequence-based typing of *Legionella pneumophila* serogroup 1 offers the potential for true portability in legionellosis outbreak investigation. J Clin Microbiol **41**:2932-2939.

Scaturro, M., M. Losardo, G. De Ponte, and Ricci, **ML. (2005).** Comparison of three molecular methods used for subtyping of *Legionella pneumophila* strains isolated during an epidemic of legionellosis in Rome. J Clin Microbiol **43**:5348-5350.

Wong, S., K. Pabbaraju, V. F. Burk, G. C. Broukhanski, J. Fox, T. Louie, M. W. Mah, K. Bernard, and Tilley, PAG. (2006). Use of sequence-based typing for investigation of a case of nosocomial legionellosis. J Med Microbiol **55**:1707-1710.

Young M, Smith H, Gray B, Huang B, Barten J, Towner C, Plowman S, Afshar B, Fry N, Blair B, Savill J, McCall B. (2005). The public health

APPENDIX 1

M13 forward (5′ - TGTAAAACGACGGCCAGT - 3′) and M13 reverse (5′ - CAGGAAACAGCTATGACC - 3′) primers can be attached to the SBT amplification primers in order to simplify and accelerate the sequencing process. M13F and M13R are used for forward and reverse sequencing of all PCR products, except that mompS-1015R (instead of M13R) is used for the reverse sequencing reaction of the *mompS* PCR product.

Gene	Primer name	Position	Primer sequence (5'-3')	Annealing temperature	Fragment size* (bp)
flaA	flaA-587F (M13F)	568-587	TGTAAAACGACGGCCAGT GCG TAT TGC TCA AAA TAC TG	55 °C	450*
	flaA-960R (M13R)	981-960	CAGGAAACAGCTATGACC CCA TTA ATC GTT AAG TTG TAG G		
pilE	pilE-35F (M13F)	12-35	TGTAAAACGACGGCCAGT CAC AAT CGG ATG GAA CAC AAA CTA	55 °C	496*
	pilE-453R (M13R)	471-453	CAGGAAACAGCTATGACC GCT GGC GCA CTC GGT ATC T		
asd	asd-511F (M13F)	487-511	TGTAAAACGACGGCCAGT CCC TAA TTG CTC TAC CAT TCA GAT G	55 °C	612*
	asd-1039R (M13R)	1062-1039	CAGGAAACAGCTATGACC CGA ATG TTA TCT GCG ACT ATC CAC		
mip	mip-74F (M13F)	58-74	TGTAAAACGACGGCCAGT GCT GCA ACC GAT GCC AC	55 °C	595*
	mip-595R (M13R)	616-595	CAGGAAACAGCTATGACC CAT ATG CAA GAC CTG AGG GAA C		
mompS	mompS-450F (M13F)	430-450	TGTAAAACGACGGCCAGT TTG ACC ATG AGT GGG ATT GG	55 °C	729*
	mompS-1116R	1140-1116	TGG ATA AAT TAT CCA GCC GGA CTT C		
proA	proA-1107F (M13F)	1090-1107	TGTAAAACGACGGCCAGT GAT CGC CAA TGC AAT TAG	55 °C	517*
	proA-1553R (M13R)	1570-1553	CAGGAAACAGCTATGACC ACC ATA ACA TCA AAA GCC		
neuA	neuA-196F (M13F)	176-196	TGTAAAACGACGGCCAGT CCG TTC AAT ATG GGG CTT CAG	55 °C	495*
	neuA-634R (M13R)	634-611	CAGGAAACAGCTATGACC CGA TGT CGA TGG ATT CAC TAA TAC		

* including M13 primers