ESGLI ESCMID STUDY GROUP FOR LEGIONELLA INFECTIONS

European Society of Clinical Microbiology and Infectious Diseases

NESTED Sequence-based Typing (SBT) protocol for epidemiological typing of *Legionella pneumophila* directly from clinical samples

VERSION 2.0

SUMMARY

This procedure describes the use of nested Sequence-Based Typing (SBT) PCR to obtain typing data from *Legionella pneumophila* in samples with low genomic DNA content. Typically DNA extracts from clinical samples previously demonstrated to be positive for the presence of *L. pneumophila* DNA by PCR are used as the starting material. This protocol is a modification to the one recently published by Ginevra C *et al* (JCM 2009). The first round of PCR is carried out with 7 primer pairs (flaA, pilE, asd, mip, mompS, proA, neuA), which bind externally to those used in the second round. Approximately five microlitres from each of the resulting PCR mixes from the first round is used as the starting template for a second round of PCR with the internal primers. Following purification, the resulting amplicons are then sequenced and assigned allele numbers and sequence types (ST) in the standard way.

DNA TEMPLATE

Previously extracted genomic DNA from clinical or environmental samples using standard DNA extraction methods: e.g. the Instagene Matrix (BIO-RAD), MagNaPure Compact (Roche), Qiagen.

SBT TARGETS

All 7 loci (i.e., *flaA*, *pilE*, *asd*, *mip*, *mompS*, *proA, neuA*) should be determined if possible.

First Round NESTED SBT amplification primers:

Gene	Primer name	Primer sequence (5'-3')	Annealing temperature
flaA	flaA-L-N [1]	TAT GCG TGA GCT TTC CGT TC	50 °C
	flaA-960R [2]	CCA TTA ATC GTT AAG TTG TAG G	
pilE	pilE-L-N [1]	CGT TGG AAT CGG CTT GTC	50 °C
	pilE-R-N [1]	CGC ATT GGC AGA GGA ATC TA	
asd	asd-1-N [1]	CCC TGG AAG TGA ATC CTC AT	50 °C
	asd-2-N [1]	TTG CAG TAT TTC AGC GAT CTG T	
mip	mip-1-N [1]	TGA AGA TGA AAT TGG TGA CTG C	50 °C
	mip-2-N [1]	AAT AGG TCC GCC AAC GCT AC	
mompS	mompS-450F [2]	TTG ACC ATG AGT GGG ATT GG	50 °C
	mompS-R-N [1]	TGG ATA AAT TAT CCA GCC GGA CTT C	
proA	proA-L-N [1]	CCG CTT CTC CAA CCA ATG A	50 °C
	proA-R-N [1]	CAC TCA ACA TAC CGC AAC CA	
neuA	neuA-F-N [1]	CCT TGC AGT CGT CTT GTT GT	50 °C
	neuA-R-N [1]	TTT CTG TTA GAG CCC AAT CG	

[1] Ginevra C et al JCM 2009
[2] The number in the primer name is the position in the reference sequence where the 3'-terminus of the oligonucleotide binds

Second round NESTED SBT PCR amplification primers:

Gene	Primer name	Primer sequence (5'-3')	Annealing temperature
flaA	flaA-587F [2]	GCG TAT TGC TCA AAA TAC TG	55 °C
	flaA-R-N [1]	GGT ATC ACC TGC GGT TCC A	
pilE	pilE-35F [2]	CAC AAT CGG ATG GAA CAC AAA CTA	55 °C
	pilE-453R [2]	GCT GGC GCA CTC GGT ATC T	
asd	asd-511F [2]	CCC TAA TTG CTC TAC CAT TCA GAT G	62°C
	asd-1039R [2]	CGA ATG TTA TCT GCG ACT ATC CAC	
mip	mip-74F [2]	GCT GCA ACC GAT GCC AC	60 °C
	mip-595R [2]	CAT ATG CAA GAC CTG AGG GAA C	
mompS	mompS-509F [2]	GAC ATC AAT GTG AAC TGG	55 °C
	mompS-1015R [2]	CAG AAG CTG CGA AAT CAG	
proA	proA-1107F [2]	GAT CGC CAA TGC AAT TAG	55 °C
	proA-1553R [2]	ACC ATA ACA TCA AAA GCC	
neuA	neuA-196F [2]	CCG TTC AAT ATG GGG CTT CAG	55 °C
	neuA-634R [2]	CGA TGT CGA TGG ATT CAC TAA TAC	

[1] Ginevra C et al JCM 2009
[2] The number in the primer name is the position in the reference sequence where the 3'-terminus of the oligonucleotide binds

Sequencing primers

The primers used in the second round of the NESTED SBT are also used for the forward and reverse sequencing of the PCR products. In the event of using the M13 modified primers in the second round of amplification, M13 forward (5′-TGTAAAACGACGGCCAGT-3′) and M13 reverse (5′-CAGGAAACAGCTATGACC-3′) only will be used for the sequencing reactions (refer to **APPENDIX 1**).

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.

First round PCR reaction

PCR amplification is performed in a total volume of 20 µl. Place PCR plate or tubes on a cold block. Start with adding 4 µl of Taq DNA polymerase Mix to each PCR tube or 96 well plate, then add 11 µl of the Master Mix, finally add the DNA template (5 µl /tube or /well). Gently spin tubes/plates briefly in an appropriate centrifuge before placing them in a thermocycler heating block. Include a negative (5 µl of nuclease-free water) and a positive control (1pg or 50 copies of *Legionella pneumophila* DNA in 5 µl of nuclease-free water). Pipette solutions in the order of: test samples, negative control, positive control.

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			5.2
PCR buffer	10 x	1 x	1.0
$MgCl_2$	50 Mm	2.5 mM	2.0
Nested Primer F	10 pmol/µl	10 pmoles	1.0
Nested Primer R	10 pmol/µl	10 pmoles	1.0
DNTPs	5 mM	200µM	0.8
Volume/reaction			11.0

Thermal cycler parameters

Step	**Temperature (ºC)**	**Time**		**No. cycles**
1	95	5 min	Initial denaturation	1
2	95	30 sec	Denaturation	40
	50	30 sec	Annealing	
	72	60 sec	Extension	
3	72	10 min	Final Extension	1
4	12	∞	Hold	

Second round PCR reaction

PCR amplification is performed in a total volume of 20 µl. Place PCR plate or/tubes on a cold block. Start with adding 4 µl of Taq DNA polymerase (1 unit/reaction) to each PCR tube/or 96 well plate, then add 11 µl of the master mix, finally add 5 µl from each of the completed PCR reaction mixes from the first round as starting template. Gently spin tubes/plates briefly in an appropriate centrifuge before placing them in a thermocycler heating block. Pipette solutions in the order of: test samples, negative control, positive control.

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			5.2
PCR buffer	10 x	1 x	2.0
$MgCl_2$	50 Mm	2.5 mM	1.0
Primer F	10 pmol/µl	10 pmoles	1.0
Primer R	10pmol/µl	10 pmoles	1.0
dNTPs	5 mM	200µM	0.8
Volume/reaction			11.0

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	60 sec	Extension	
3	72	10 min	Final Extension	1
4	12	∞	Hold	

Analysis of PCR products by gel electrophoresis

Run PCR products (5 μl + 15 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

Ginevra, C., Lopez, M., Forey, F., Reyrolle, M., Meugnier, H., Vandenesch, F., Etienne, J., Jarraud, S. and Molmeret, M. (2009). Evaluation of a nested-PCR-derived sequence-based typing method applied directly to respiratory samples from patients with Legionnaires' disease. J Clin Microbiol. **47**:981-7.

Ratzow, S., Gaia, V., Helbig, J., Fry, N.K. and Lück, P.C. (2007). Addition of *neuA*, the gene encoding N-acylneuraminate cytidylyl transferase, increases the discriminatory ability of the consensus sequence-based scheme for typing *Legionella pneumophila* serogroup 1 strains. *J Clin Microbiol* **45**: 1965-1968

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, N.K., Afshar, B., Luck, P.C., 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*. *J Clin Microbiol* **43**: 2047-52.

Gaia, V., N.K. Fry, Harrison, T.G. 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.

APPENDIX 1

M13 forward (5′-TGTAAAACGACGGCCAGT-3′) and M13 reverse (5′-CAGGAAACAGCTATGACC-3′) primers can be attached to the NESTED SBT amplification primers used in the second round of amplification in order to simplify and accelerate the sequencing process. M13F and M13R are then used for forward and reverse sequencing of all PCR products.

Gene	Primer name	Primer sequence (5'-3')	Annealing temperature
flaA	flaA-587F [2] (M13F)	TGTAAAACGACGGCCAGT GCG TAT TGC TCA AAA TAC TG	55 °C
	flaA-R-N [1] (M13R)	CAGGAAACAGCTATGACC GGT ATC ACC TGC GGT TCC A	
pilE	pilE-35F [2] (M13F)	TGTAAAACGACGGCCAGT CAC AAT CGG ATG GAA CAC AAA CTA	55 °C
	pilE-453R [2] (M13R)	CAGGAAACAGCTATGACC GCT GGC GCA CTC GGT ATC T	
asd	asd-511F [2] (M13F)	TGTAAAACGACGGCCAGT CCC TAA TTG CTC TAC CAT TCA GAT G	55 °C
	asd-1039R [2] (M13R)	CAGGAAACAGCTATGACC CGA ATG TTA TCT GCG ACT ATC CAC	
mip	mip-74F [2] (M13F)	TGTAAAACGACGGCCAGT GCT GCA ACC GAT GCC AC	55 °C
	mip-595R [2] (M13R)	CAGGAAACAGCTATGACC CAT ATG CAA GAC CTG AGG GAA C	
mompS	mompS-509F [2] (M13F)	TGTAAAACGACGGCCAGT GAC ATC AAT GTG AAC TGG	55 °C
	mompS-1015R [2] (M13R)	CAGGAAACAGCTATGACC CAG AAG CTG CGA AAT CAG	
proA	proA-1107F [2] (M13F)	TGTAAAACGACGGCCAGT GAT CGC CAA TGC AAT TAG	55 °C
	proA-1553R [2] (M13R)	CAGGAAACAGCTATGACC ACC ATA ACA TCA AAA GCC	
neuA	neuA-196F [2] (M13F)	TGTAAAACGACGGCCAGT CCG TTC AAT ATG GGG CTT CAG	55 °C
	neuA-634R [2] (M13R)	CAGGAAACAGCTATGACC CGA TGT CGA TGG ATT CAC TAA TAC	

[1] Ginevra C *et al* JCM 2009

[2] The number in the primer name is the position in the reference sequence where the 3'-terminus of the oligonucleotide binds