

# Supplementary Information

## Sequence analysis of nonulosonic acid biosynthetic gene clusters in *Vibrionaceae* and *Moritella viscosa*

Marie-Josée Haglund Halsør<sup>1</sup>, Bjørn Altermark<sup>1</sup>, and Inger Lin Uttakleiv Ræder<sup>1,\*</sup>

<sup>1</sup>The Norwegian Structural Biology Centre (NorStruct), Department of Chemistry, UiT- The Arctic University of Norway, N-9037 Tromsø, Norway

\*inger.l.rader@uit.no

**Table S1.** Sequences used for *Vibrio* sp. B9-25K2 phylogeny.

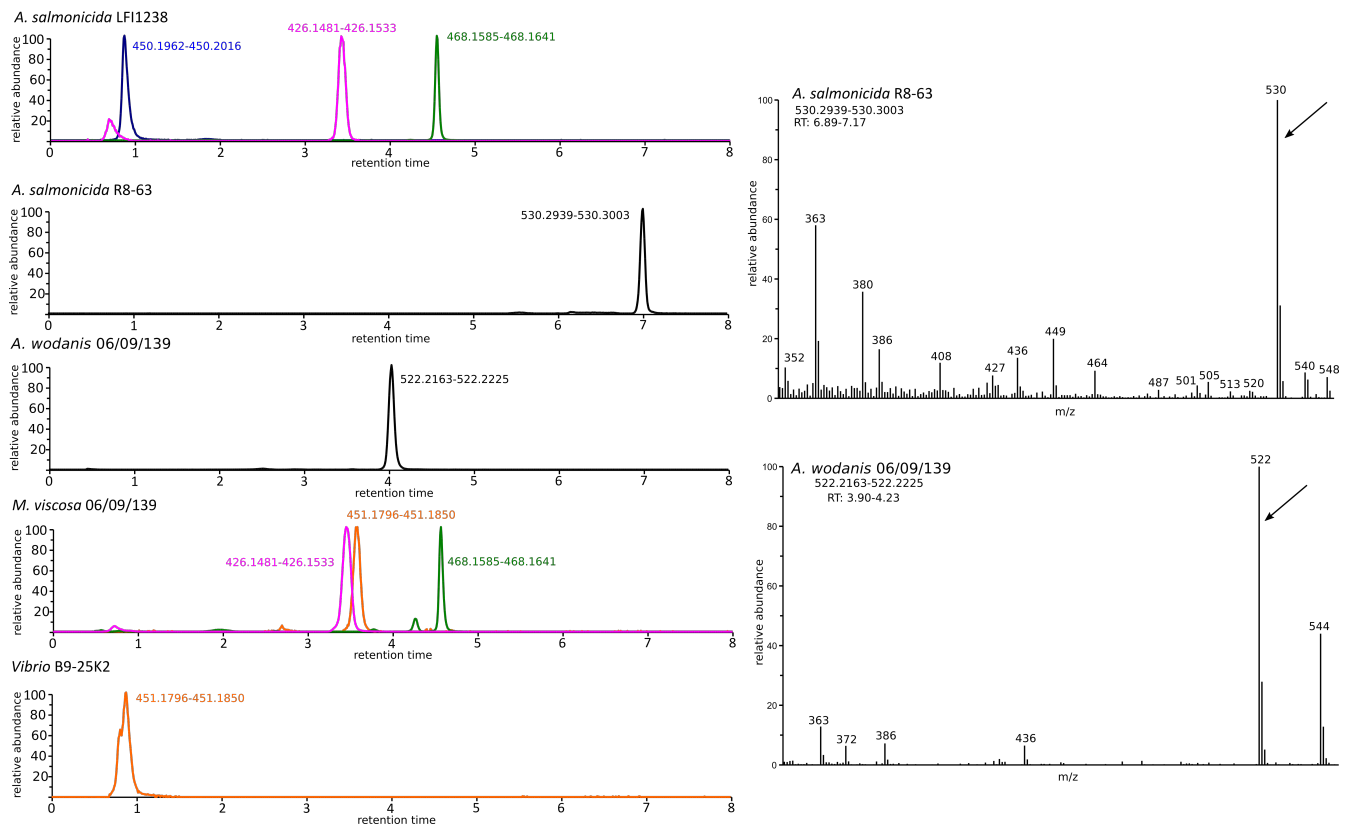
Strain <i>reference</i>	Accession and ranges		
	16SrDNA	<i>ftsZ</i>	<i>gapA</i>
<i>Aliivibrio logei</i> MR17-77 <sup>1</sup>	1434910-1436398	3523605-3524840	2928874-2929790
<i>A. wodanis</i> 06/09/139 <sup>1</sup>	LN554846.1 (506389-507875)	464002-465231	1057819-1058735
<i>A. sp.</i> R8-63 <sup>1</sup>	2991032-2992518	80694-81932	4041528-4042444
<i>A. salmonicida</i> LFI1238 <sup>2</sup>	NC_011312.1 (87128-88616)	2857507-2856272	1984565-1985481
<i>Moritella viscosa</i> 06/09/139 <sup>3</sup>	LN554852.1 (31135-32623)	4800155-4798974	3844953-3845966
<i>M. viscosa</i> LFI 5006 <sup>3</sup>	FPLG <sup>a</sup> (81 <sup>b</sup> ; 98-1586)	88 <sup>b</sup> ; 80018-81199	53 <sup>b</sup> ; 23338-24351
<i>Vibrio anguillarum</i> NB10 <sup>4,5</sup>	LK021130.1 (233633-235121)	2515501-2516718	1112491-1113410
<i>V. sp.</i> B9-25K2 <sup>1</sup>	3642484-3643981	242232-243211	4201830-4202749
<i>V. vulnificus</i> CMCP6 <sup>6,7</sup>	AE016795.3 (475743-477229)	588232-587018	3210958-3211877
	<i>gyrB</i>	<i>mreB</i>	<i>pyrH</i>
<i>Aliivibrio logei</i> MR17-77 <sup>1</sup>	3216464-3217646	169476-170505	3021801-3022387
<i>A. wodanis</i> 06/09/139 <sup>1</sup>	LN554846.1 (11401-12583)	2635881-2634838	2379026-2379612
<i>A. sp.</i> R8-63 <sup>1</sup>	465716-466898	2026880-2027922	1570486-1571072
<i>A. salmonicida</i> LFI1238 <sup>2</sup>	NC_011312.1 (11398-12580)	541014-542057	2595577-2596154
<i>Moritella viscosa</i> 06/09/139 <sup>3</sup>	LN554852.1 (5072328-5073736)	4853906-4854946	4396130-4396716
<i>M. viscosa</i> LFI 5006 <sup>3</sup>	FPLG <sup>a</sup> (94 <sup>b</sup> ; 26583-27765)	88 <sup>b</sup> ; 134953-135993	05 <sup>b</sup> ; 39318-39889
<i>Vibrio anguillarum</i> NB10 <sup>4,5</sup>	LK021130.1 (11313-12710)	357047-358090	2368060-2368788
<i>V. sp.</i> B9-25K2 <sup>1</sup>	5040401-5041803	37279-38250	3344816-3345402
<i>V. vulnificus</i> CMCP6 <sup>6,7</sup>	AE016795.3 (999953-1001349)	1432966-1434009	1856011-1856597
	<i>recA</i>	<i>rpoA</i>	<i>topA</i>
<i>Aliivibrio logei</i> MR17-77 <sup>1</sup>	4960-6003	3796758-3797685	2769318-2771957
<i>A. wodanis</i> 06/09/139 <sup>1</sup>	LN554846.1 (616746-617792)	2765320-2766247	1218882-1221515
<i>A. sp.</i> R8-63 <sup>1</sup>	1850404-1851453	2759195-2760122	2965401-2968040
<i>A. salmonicida</i> LFI1238 <sup>2</sup>	NC_011312.1 (702343-703386)	375824-376751	1266401-1269040
<i>Moritella viscosa</i> 06/09/139 <sup>3</sup>	LN554852.1 (529995-531035)	4908357-4907374	1552072-1554717
<i>M. viscosa</i> LFI 5006 <sup>3</sup>	FPLG <sup>a</sup> (05 <sup>b</sup> ; 60952-61992)	90 <sup>b</sup> ; 852-1835	33 <sup>b</sup> ; 82187-84832
<i>Vibrio anguillarum</i> NB10 <sup>4,5</sup>	LK021130.1 (529112-529908)	2696727-2697657	1281671-1284301
<i>V. sp.</i> B9-25K2 <sup>1</sup>	255724-256681	4035345-4036275	4894973-4897596
<i>V. vulnificus</i> CMCP6 <sup>6,7</sup>	AE016795.3 (1563785-1564703)	752353-753283	2129301-2131925

<sup>a</sup> Whole Genome Sequencing (WGS) project identifier

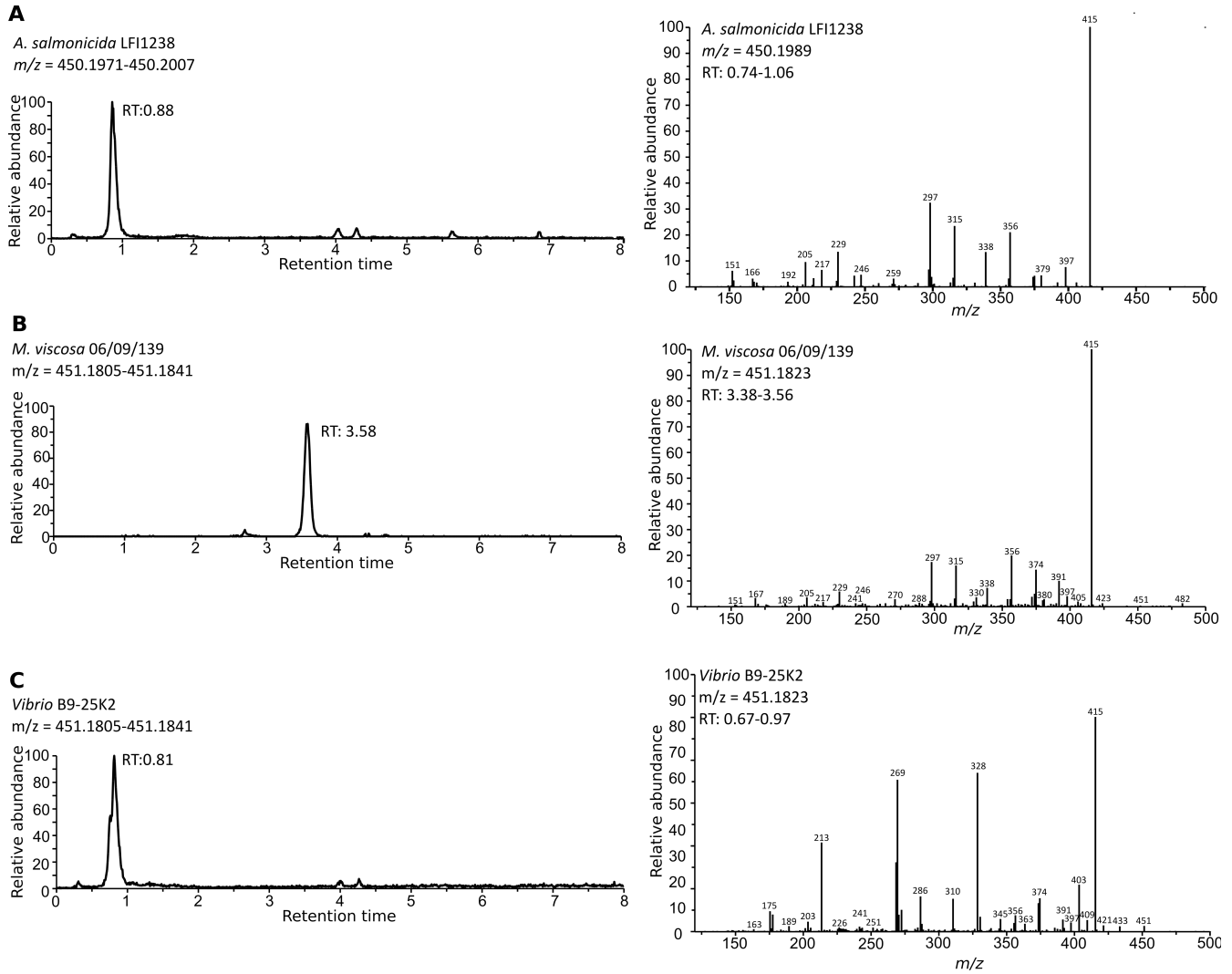
<sup>b</sup> Contig number

A. salmonicida LFI1238 (Neu)	.....MKKIYIVAEIGCNHNGDFNIAKMMVDEAKKAGVDAVKFQTFKADQLISKFAPKAEYQIKVGTGND	64
M. viscosa 06/09/139 (Neu)	MTNPVFEISGRKVLGDYAPLVIAEIGINHEGSLKTAFFEMVDAAIEGCAEIIKHQTH....VIEDEMSSBAKKVIPGNA	74
A. salmonicida LFI1238 (Leg)	.....MTLIIABAGVNHNGDENLAFQLVDAAYHAGADIVKFFQTFKAKNLVTEDAVQAEYQVNTNQKQ	62
A. logei MR17-77	.....MTLIIABAGVNHNGDENLAFQLVDAAYHAGADIVKFFQTFKAKNLVTEDAVQAEYQVNTNQKQ	62
V. vulnificus CMCP6	.....MNNQTFIIABAGVNHNGDIIVLAKKLIDAAANSQVDAVKFQTFWKTELLVTEDAKMAEYQIDNTQRE	65
A. wodanis 06/09/139	.....MKNQTFIIABAGVNHNGDIALAKKLIDVAANSQVDAVKFQTFWKTELLVTEDAKMAEYQIDNTQRE	65
A. magni R8-63	.....MTNKTFFIIABAGVNHNGDIRLAKQLIDAAADAGVDAVKFQTFWKTELLVTEDAKMAEYQVENTQRE	65
M. viscosa 06/09/139 (Pse)	MTEQYITIDGRKIGPNFSPYIIAELSANHNGDINRAFAIMEAAKACADAIKLQTYTQDT.ITMDCDSEFQIKGGLWH	78
Vibrio sp. B9-25K2	.....MQIAGRKIGAGHKPYIIAEMSGNHNGDIKRAIELIKAAKEAGADAVKLTQTYTADT.ITIDHDGDEFLIKGGLWD	73
A. salmonicida LFI1238 (Neu)	ESQ.LEMTRKLELPYDEFIKLEEYAKEIGLDVFSTPFDFSDIFL.ASREQTKTIPSGELLNLEYLEKIIARLPIAGKE	141
M. viscosa 06/09/139 (Neu)	DVSIYEIMDRCSLNEEDEIKLKKYIESKGAIFISTPFSRAALRE.ERMVGSAYKICSGECNNYFLDLIASY...GKP	149
A. salmonicida LFI1238 (Leg)	ESQ.LAMLSRLELSYDIHHQLVKHCESLGIIEFLSTAFDSESLDFVNDLGLTRLKIPSGEITNAPLVLEHART...GCD	137
A. logei MR17-77	ESQ.LAMLSRLELSYDVHQQLVKHCESLGIIEFLSTAFDSESLDFVNDLGLTRLKIPSGEITNAPLVLEHART...GCD	137
V. vulnificus CMCP6	ESQ.FQMLKRLELSYDDFLELKEYCDSKNILFMSTPDEEQSATFL.NALQE.TFKICSGELTNTFLRHIASF...GKP	138
A. wodanis 06/09/139	ESQ.FQMLKRLELSYDNFLELKEYCDSKNILFMSTPDEEQSATFL.NALQD.TFKICSGELTNTFLRHIASF...GKP	138
A. magni R8-63	ETQ.FEMLKRLELSYDDFTELKSYCDDKGITFMSTPDEEQSATFL.DGLQA.VFKICSGELTNTFLRHIASF...AKP	138
M. viscosa 06/09/139 (Pse)	GQSLYQLYTSAHMPWEWHQPLFAKAKELDITIFSSPFDFTAVDLE.EELDAPAYKIASFEIVDLPLIKRVAQT...GKP	153
Vibrio sp. B9-25K2	GSRLYDLYQDAHTPWDPWHKVLDFDEAKKLGITIFSSPFDHTAVDLE.QALDAPAYKIASFELIDLPLIRKVAQT...GKP	148
A. salmonicida LFI1238 (Neu)	IVI <del>STG</del> MATVEEIQ <del>LAL</del> .....DVLAKNGMKPEE.....ITILHCNTEYPTAFEDVNLNSIAGFK	196
M. viscosa 06/09/139 (Neu)	VII <del>STG</del> MNDIPSIRKSV.....EIFRKYKTP.....LCILH <del>TTNL</del> YPTDHLIRIGAMEEMQ	201
A. salmonicida LFI1238 (Leg)	LIV <del>STG</del> MATLSEIEAVLGVIAGFYVADKTEEPSLAFQQAAYASEQQKALKKKV <del>TILH</del> CTTEY <del>PAPM</del> VEINLKAM <del>DTLG</del>	216
A. logei MR17-77	LIV <del>STG</del> MATLSEIEAVLGVIAGFYVADKTEEPSLAFQQAAYASEQQKALKKQV <del>TILH</del> CTTEY <del>PAPM</del> AEINLKAM <del>DTLG</del>	216
V. vulnificus CMCP6	II <del>STG</del> MGYLSEVEHAL.....FTLKEAGIPTEL.....ITV <del>LH</del> ATTDYPTSPEDVNLQAM <del>LTIG</del>	193
A. wodanis 06/09/139	II <del>STG</del> MGYLSEVEHAL.....FTLKEAGIPTES.....ITV <del>LH</del> ATTDYPTSPEDVNLQAM <del>LTIG</del>	193
A. magni R8-63	VII <del>STG</del> MGYLSEVEHAV.....VTLRDAGLSLDM.....ITV <del>LH</del> ATTDYPTAPEDVNL <del>LAM</del> KDTIE	193
M. viscosa 06/09/139 (Pse)	MI <del>ISTG</del> MADQSEIEIAI.....QTA <del>KD</del> NGCDE.....LVV <del>LH</del> CVSGY <del>PAPAA</del> QYNLRTIADIG	206
Vibrio sp. B9-25K2	II <del>ISTG</del> MGANLAEIEEAI.....LAAK <del>GAG</del> AKE.....LVL <del>LH</del> CTSGY <del>PTPAD</del> QANIST <del>M</del> SVMR	201
A. salmonicida LFI1238 (Neu)	ETIKQYKICFSDHSPGYFASIASVYPCITFIEKHF <del>TL</del> DKNFHG <del>PE</del> HKASVTP <del>EB</del> LTL <del>L</del> CQGI <del>RA</del> VEQALGSHDKLVTNS	275
M. viscosa 06/09/139 (Neu)	REFSDVVVGLSDHSIDNLACLGAAAGASVLERHFTDNKARSQPDICCSMDGAECAELISQSKRMAQMRG.GSKGAVKE	279
A. salmonicida LFI1238 (Leg)	RAFE.LPAGYSDHSAGITIPIAAVARGAVLIEKHF <del>TL</del> DQNMEG <del>PE</del> HKASLEP <del>QL</del> TAMVSAIRQVEVALG <del>VGVK</del> SP <del>TVS</del>	294
A. logei MR17-77	RAFE.LPAGYSDHSAGITIPIAAVARGAVLIEKHF <del>TL</del> DKNMEG <del>PE</del> HKASLEP <del>QL</del> TAMVSAIRQVEVALG <del>VGVK</del> SP <del>TVS</del>	294
V. vulnificus CMCP6	YSFPGITICYS <del>DL</del> TLGIEIPVAAVAMCAK <del>VI</del> EKHF <del>TL</del> DNTMEG <del>PE</del> HKASLEP <del>EB</del> L <del>TAM</del> VQAIRNIELALGSGW <del>KV</del> PTAT	272
A. wodanis 06/09/139	HSFPGITICYS <del>DL</del> TLGIEIPVAAVAMCAK <del>VI</del> EKHF <del>TL</del> DNTMEG <del>PE</del> HKASLEP <del>EB</del> L <del>TAM</del> VKAIRNIELALGSGW <del>KV</del> PTPT	272
A. magni R8-63	QAFPGIIVGYS <del>DL</del> TLGTEIPVAAVAMCAK <del>VI</del> EKHF <del>TL</del> DKTMSG <del>PE</del> HKASLEP <del>QL</del> ADMVTAIRNIEQALGNGW <del>KV</del> PTKT	272
M. viscosa 06/09/139 (Pse)	QRFD.VLACLSDHTIDNATAVVSVAFGAC <del>VI</del> EKHF <del>TL</del> DRNAGCA <del>DS</del> FSLEP <del>DEL</del> ARLCRD <del>TY</del> TAWQAMCNVNYERTPA	284
Vibrio sp. B9-25K2	GAFN.CEVGLSDHTMGIGVSI <del>AA</del> VALGAC <del>VI</del> EKHF <del>TL</del> ARADG <del>PE</del> S <del>AF</del> SLEKEBELKSLVDN <del>CAM</del> AFESLQPNFISTEA	279
A. salmonicida LFI1238 (Neu)	ERKNKIVARKSIVAKFDIKKGV <del>FT</del> LDNITTKR <del>PC</del> NG.ISPMSWYEV <del>L</del> GKTAEQD <del>F</del> SE <del>D</del> QLIAHSEFVAQEV	346
M. viscosa 06/09/139 (Neu)	EQVTIDFAYASVVTIKEIKAGEA <del>FT</del> KNLWV <del>KR</del> PC <del>TC</del> DFLAD <del>D</del> YEMLLGKKASQ <del>N</del> IDFVQLKKEFIK....	347
A. salmonicida LFI1238 (Leg)	EVKNKAVARKSLVAAKAIQQCDLITEDNLTIKR <del>PC</del> SG.MSPYHYWALLNQPASQDYKAGDLIVE.....	357
A. logei MR17-77	EVKNKAVARKSLVAAKAIQQGELITEDNLTIKR <del>PC</del> SG.MSPYHYW <del>TV</del> LNQPASQDYKAGDLIVE.....	357
V. vulnificus CMCP6	EKQNRNIVRKSIVAAQAIPEGTV <del>FT</del> KDMLS <del>IKR</del> PC <del>DC</del> .ISPTRWDEVIGSIAKKDYRDGELI.....	333
A. wodanis 06/09/139	EKQNRNIVRKSIVAAQTISEGTV <del>FT</del> KDMLS <del>IKR</del> PC <del>DC</del> .ISPTRWDEVIGSIAKKDYRDGELI.....	333
A. magni R8-63	EQENRNIVRKSIVAGKPI <del>M</del> AGSICAD <del>M</del> LEIKR <del>PC</del> NG.ISPTRWDEVVGSIAKKDYQIGELI.....	333
M. viscosa 06/09/139 (Pse)	EQGNVKF.RRSLYAVKDI <del>A</del> AGEL <del>L</del> TADN <del>V</del> RSI <del>R</del> PC <del>FG</del> .LEPKYYDQV <del>L</del> GKANV <del>I</del> ISKGTALSPGLIS....	350
Vibrio sp. B9-25K2	ESQTKPH.RRSLYIVKDI <del>A</del> KCEL <del>F</del> TDEH <del>V</del> RSI <del>R</del> PC <del>NG</del> .ILPKYLD <del>D</del> VIG <del>S</del> TATEDLTFGTPLKFGHFK....	345

**Figure S1.** Multiple alignment of unique NeuB sequences. The sequences from *A. salmonicida* LFI1238 (WP\_012551408.1 and WP\_012549051.1), *A. magni* R8-63 and *A. logei* MR17-77 were aligned with that of *M. viscosa* 06/09/139 (WP\_045111757.1 and WP\_045111735.1) and *A. wodanis* 06/09/139 (WP\_045100955.1) as well as *Vibrio* B9-25K2. The alignment was done using MUSCLE<sup>8,9</sup>.



**Figure S2.** LC-MS analysis of bacterial NuO content derivatized with DMB. Left: FTMS+ pESI full MS spectra. The spectra are represented for each species as a function of relative abundance versus retention time (in minutes). When several mass ranges are considered, the spectra are combined in a single graph using different colors for each curve (blue, pink, green, and orange). The mass range ( $m/z$ ) corresponding to each peak is indicated next to it. Right: Mass spectra for the peaks from *A. salmonicida* R8-63 and *A. wodanis*.



**Figure S3.** Mass spectrometry analysis of putative LegAcAm and PseAc2 compounds. A. Analysis for *A. salmonicida* LFI1238. Left: FTMS+ pESI full MS spectra at  $m/z = 450.1971 - 450.2007$ . Right: ITMS + cESI with full MS2 spectra of mass 450.1989 at time 0.74-1.06 minutes. B. Analysis for *M. viscosa* 06/09/139. Left: FTMS+ pESI full MS spectra at  $m/z = 451.1805 - 451.1841$ . Right: ITMS + cESI with full MS2 spectra of mass 451.1823 at time 3.38-3.56 minutes. C. Analysis for *Vibrio* B9-25K2. Left: FTMS+ pESI full MS spectra at  $m/z = 451.1805 - 451.1841$ . Right: ITMS + cESI with full MS2 spectra of mass 451.1823 at time 0.67-1.06 minutes.

## References

1. Purohit, A. A. *et al.* Presence of acyl-homoserine lactones in 57 members of the *Vibrionaceae* family. *J. Appl. Microbiol.* **115**, 835–847, DOI: <https://doi.org/10.1111/jam.12264> (2013).
2. Hjerde, E. *et al.* The genome sequence of the fish pathogen *Aliivibrio salmonicida* strain LFI1238 shows extensive evidence of gene decay. *BMC Genomics* **9**, DOI: <https://doi.org/10.1186/1471-2164-9-616> (2008).
3. Karlsen, C., Hjerde, E., Klemetsen, T. & Willassen, N. P. Pan genome and CRISPR analyses of the bacterial fish pathogen *Moritella viscosa*. *BMC genomics* **18**, 313, DOI: <https://doi.org/10.1186/s12864-017-3693-7> (2017).
4. Eiler, A. & Bertilsson, S. Detection and quantification of *Vibrio* populations using denaturant gradient gel electrophoresis. *J. Microbiol. Methods* **67**, 339–348, DOI: <https://doi.org/10.1016/j.mimet.2006.04.002> (2006).
5. Holm, K. O., Nilsson, K., Hjerde, E., Willassen, N. P. & Milton, D. L. Complete genome sequence of *Vibrio anguillarum* strain NB10, a virulent isolate from the Gulf of Bothnia. *Standards Genomic Sci.* **10**, DOI: <https://doi.org/10.1186/s40793-015-0060-7> (2015).
6. Kim, Y. R. *et al.* Characterization and Pathogenic Significance of *Vibrio vulnificus* Antigens Preferentially Expressed in Septicemic Patients. *Infect. Immun.* **71**, 5461–5471, DOI: [10.1128/IAI.71.10.5461-5471.2003](https://doi.org/10.1128/IAI.71.10.5461-5471.2003) (2003).
7. Kim, H. U. *et al.* Integrative genome-scale metabolic analysis of *Vibrio vulnificus* for drug targeting and discovery. *Mol. Syst. Biol.* **7**, 460, DOI: [10.1038/msb.2010.115](https://doi.org/10.1038/msb.2010.115) (2011).
8. Edgar, R. C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797, DOI: <https://doi.org/10.1093/nar/gkh340> (2004).
9. Edgar, R. C. MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinforma.* **5**, DOI: <https://doi.org/10.1186/1471-2105-5-113> (2004).