

SUPPORTING ONLINE MATERIAL:

1
2 **Further description of the sampling sites.** Hawaiian soil samples were collected at
3 Laupahoehoe National Forest on Mauna Kea, Big Island, at an altitude of 1200 m. The volcanic
4 ash parent material had an age of 10-30,000 yrs. The soil was sampled 15 cm below a heavy
5 organic root and litter layer (1). This location has a history of high and sustained impact by feral
6 swine. Samples from the Caribbean region consisted of weathered soil accumulated within
7 cavities near cliffs located at the south east coast of Mona Island, a tectonically uplifted Mio-
8 Pliocene limestone island of a semi-arid climate located at 47 miles from the west coast of Puerto
9 Rico (Lat.: 18.10° N Long.: 67.90° W). Mona Island has been under the protection of Puerto
10 Rico's Department of Natural Resources since 1975 (2).

11 Soil from Uruguay was retrieved from an agricultural field growing an endemic variety
12 of maize (Cangüe) at the South Experimental Station of the Agronomy Faculty, University of the
13 Republic of Uruguay (Lat.: 34° 40' S, Long. 56° 15' W) Uruguay. Samples from South Africa
14 were obtained from an experimental field located at Bien Donne (Lat.: 33° 50'S, Long. 18 °
15 59'E). The dominating vegetation was plantago. The Slovenian soil was collected from a peat
16 field that received its last manure application in 1999 (three years before sampling). Since then
17 mineral fertilizers (NPK) and compost material have been used to increase organic matter
18 content. The field is located at University of Ljubljana, Slovenia (Lat.: 46°02'55" N, Long.
19 14°28'15E).

20 Arctic marine sediment samples were collected at two different locations off the north
21 coast of Alaska: Barrow Canyon station 180 m (Lat. 71:36.24N, Long.156:12.49W, Depth 160
22 m) Barrow Canyon station 2000 m (Lat. 72:12.33N, Long.154:05.5W, Depth 2000 m), and East

23 Hanna Shoal station 160 m (Lat. 72.637N, Long.158.667W). The depths are of the overlying
24 water column.

25 The Puget Sound samples came from surface sediment at Carr Inlet under a water depth
26 of 84 m while sediments from the Washington coast were retrieved under a water depth of 1000
27 m (3). The former is the only sample that might have been moderately impacted by fecal material
28 although *E. coli* counts and ammonium concentrations were low. Sediments were raised from the
29 sea floor using an Ocean Instruments MC-800 multicore, which collects eight individual cores
30 with well-preserved interfaces. Individual cores were extruded and sampled at a depth interval
31 ranging from 2-6 cm, immediately frozen and stored at -80°C .

32 The biosolids were the dried sludge from a domestic sewage treatment plant, East
33 Lansing Michigan. The composted manure as well as the paired set of manured and non-
34 manured but adjacent soils were provided by anonymous swine producers from Ohio and
35 Michigan, respectively. Swine feed was supplemented daily with chlorotetracycline on the
36 Michigan farm. At this site the manure was collected in pits and applied in June of 2001 as
37 slurries to fields in which row crops were grown annually. Manured fields were sampled at three
38 days (sample AMS4), one week (sample AMS23) and four weeks (sample AMS25) after manure
39 application, respectively. The manure was incorporated into the soil before the 4 week sampling
40 took place. The temperature and moisture conditions would not have been lethal too most of the
41 microbial populations in the manure. The adjacent, non-agricultural soils used as controls
42 (NMS5 and NMS24) had never been manured, had the same soil type (Capac loam) and were
43 sampled on the same dates.

44

- 45 1. **Nüsslein, K., and J. M. Tiedje.** 1999. Soil bacterial community shift correlated with change
46 from forest to pasture vegetation in a tropical soil. *Appl. Environ. Microbiol.* **65**:3622-3626.
47
- 48 2. **Frank, F. J., C. Wicks, J. Mylroie, J. Troester, E. Calvin Alexander, and J. L. Carew.**
49 1998. Geology of Isla de Mona, Puerto Rico. *J. Cave Karst. Stud.* **60**:69-72.
50
- 51 3. **Grüntzig, V., S. C. Nold, J. Zhou, and J. M. Tiedje.** 2001. *Pseudomonas stutzeri* nitrite
52 reductase gene abundance in environmental samples measured by real-time PCR. *Appl. Environ.*
53 *Microbiol.* **67**:760-768.

76 75 74 73 72 71 70 69 68 67 66 65 64 63 62 61 60 59 58 57 56 55 54

Supplementary Table 1. Characterization of bacterial phylotypes detected in fraction F7 (40-45% G+C) according to the RDP-II Seqmatch tool.

| Phylotypes | Phylogenetic affiliation and closest isolate or 16S rRNA | | | |
|---------------------------------------|--|-----------------------|------------|-------------------------------|
| Clone ID | phylotype match | S _{ab} score | GenBank ID | Source |
| <i>Firmicutes (Clostridia)</i> | | | | |
| F7-A07* | Uncultured bacterium, clone 8 | 0.95 | DQ238608 | Piglet gastrointestinal tract |
| F7-B01* | Uncultured bacterium, clone p-406-o3 | 0.98 | AF371835 | Swine intestine |
| F7-C02* | Uncultured bacterium, clone B016 | 0.95 | DQ325554 | Human stool |
| F7-C09 | <i>Veillonellaceae</i> , bacterium FCF9A | 0.87 | DQ767881 | Arctic environmental sample |
| F7-C12* | Uncultured bacterium, clone 8 | 0.94 | DQ238608 | Piglet gastrointestinal tract |
| F7-D02* | Uncultured rRNA phylotype OTU-151 | 0.98 | AF371835 | Swine intestine |
| F7-D04* | Uncultured bacterium, clone OTU8 | 0.98 | AY753318 | Giant panda feces |
| F7-D10* | Uncultured bacterium, clone 8 | 0.95 | DQ238608 | Piglet gastrointestinal tract |
| F7-E03* | Uncultured bacterium, clone p-406-o3 | 0.96 | AF371835 | Swine intestine |
| F7-E11 | Uncultured bacterium, clone Elev_16S_1359 | 0.96 | EF019973 | Trembling aspen rhizosphere |
| F7-F07* | Uncultured bacterium, clone 4 | 0.95 | DQ238611 | Piglet gastrointestinal tract |
| F7-F12* | Uncultured bacterium, clone A-2D | 0.97 | AY953222 | Anaerobic swine waste lagoon |
| F7-G05* | Uncultured bacterium, clone RL183_aao02f08 | 0.94 | DQ800727 | Human stool |
| F7-G10* | Uncultured bacterium, clone B054 | 0.96 | DQ325569 | Human stool |
| F7-H02* | Uncultured bacterium, clone A-3B | 0.95 | AY953232 | Anaerobic swine lagoon |
| F7-H04* | Uncultured bacterium, clone B054 | 0.99 | DQ325569 | Human stool |
| F7-H05* | Uncultured bacterium, clone 8 | 0.97 | DQ238608 | Piglet gastrointestinal tract |
| F7-H10* | Uncultured bacterium, clone B054 | 0.99 | DQ325569 | Human stool |
| 18 phylotypes | | | | |
| <i>Firmicutes (Bacilli)</i> | | | | |
| F7-A08 | <i>Bacillales, Bacillus asahii</i> | 0.98 | AB109209 | Soil |
| F7-A10 | <i>Bacillales, Bacillus sp. 27-1</i> | 0.75 | AB043845 | N.A. |
| F7-A12* | Uncultured bacterium, clone p-66-a5 | 0.98 | AF371535 | Swine intestine |
| F7-B07 | <i>Lactobacillales, Weissella thailandensis</i> | 0.96 | AB023839 | Fermented food |
| F7-B11 | <i>Bacillales, Bacillus asahii</i> | 0.98 | AB109209 | Soil |
| F7-C01 | Uncultured bacterium, clone FCPT497 | 0.94 | EF515973 | Grassland soil |
| F7-C05* | <i>Bacillales, Bacillus sp. No.61</i> | 0.99 | AB066338 | Animal feces compost |
| F7-D07 | <i>Bacillales, Bacillus pycnus</i> | 0.97 | AF169531 | Soil |
| F7-D08 | <i>Bacillales, Bacillus sp. isolate IDA624</i> | 0.60 | AJ236893 | Grassland soil |
| F7-D11 | <i>Bacillales, Bacterium JL-74</i> | 0.96 | AY745842 | Eastern Chinese Sea |
| F7-E09 | <i>Bacillales, Bacillus pycnus</i> | 0.94 | AF169531 | Soil |

99 98 97 96 95 94 93 92 91 90 89 88 87 86 85 84 83 82 81 80 79 78 77

Supplementary Table 1 (continued).

| Phylotypes | Phylogenetic affiliation and closest isolate or 16S rRNA | | | |
|-----------------------------|--|----------------|------------|-------------------------------------|
| Clone ID | phylotype match | S_{ab} score | GenBank ID | Source |
| F7-E10 | <i>Bacillales, Bacillaceae</i> , bacterium SM17 | 0.93 | DQ195828 | Marine sediments |
| F7-E12* | Uncultured bacterium, clone p-66-a5 | 0.98 | AF371535 | Swine intestine |
| F7-F01 | Uncultured bacterium, clone AKIW1000 | 0.94 | DQ129527 | Urban aerosol |
| F7-F04 | <i>Bacillales, Bacillus drentensis</i> | 0.88 | DQ870699 | Clean room environment |
| F7-F09 | <i>Bacillales, Bacillus asahii</i> | 0.96 | AB109209 | Soil |
| F7-G01 | <i>Bacillales, Bacillus senegalensis</i> | 0.93 | AF519465 | Tropical rice soil |
| F7-H03 | <i>Bacillales, Bacillus megaterium</i> -MO31 | 1.00 | AY553118 | Great Salt Plains, Oklahoma |
| F7-H12 | <i>Bacillales, Bacillus lentus</i> NCIMB8773 | 0.86 | AB021189 | N.A. |
| 19 phylotypes | | | | |
| <i>Bacteroidetes</i> | | | | |
| F7-A02 | Uncultured <i>Sphingobacteriales, Chitinophaga</i> | 0.98 | AB234246 | Sediment and soil |
| F7-A03 | Uncultured <i>Sphingobacteriales, Chitinophaga</i> | 0.86 | EF018366 | Trembling aspen rhizosphere |
| F7-A05* | Uncultured <i>Flavobacteriales</i> , unclassified <i>Flavobacteriaceae</i> | 0.76 | DQ824738 | Human feces |
| F7-B04 | Uncultured <i>Bacteroidetes</i> , clone 9NBGBact_75 | 0.52 | DQ070811 | Seafloor basalts, East Pacific Rise |
| F7-B05 | Uncultured, unclassified <i>Bacteroidales</i> | 0.90 | EF559220 | Anaerobic solid waste digester |
| F7-B06 | Uncultured, unclassified <i>Bacteroidales</i> | 1.00 | EF559197 | Anaerobic solid waste digester |
| F7-B08 | Uncultured <i>Sphingobacteriales, Chitinophaga</i> | 0.97 | AB234246 | Sediment and soil |
| F7-B10* | Uncultured <i>Flavobacteriales</i> , unclassified <i>Cryomorphaceae</i> | 0.76 | AB117716 | Enriched wastewater plant effluents |
| F7-B12 | Uncultured <i>Sphingobacteriales, Chitinophaga</i> | 0.83 | AY922017 | Farm soil |
| F7-C08* | Uncultured <i>Flavobacteriales</i> , unclassified <i>Flavobacteriaceae</i> | 0.90 | AY218549 | Penguin droppings sediments |
| F7-D01 | <i>Flavobacterium sp.</i> GH29-5 | 0.76 | DQ222428 | Soil |
| F7-D05 | Uncultured <i>Sphingobacteriales, Sphingobacterium</i> | 0.92 | AY838480 | Soil fungus endophyte |
| F7-D09 | Uncultured <i>Sphingobacteriales, Chitinophaga</i> | 0.96 | AB234246 | Sediment and soil |
| F7-E01 | Uncultured, unclassified <i>Bacteroidales</i> | 0.84 | EF559083 | Anaerobic solid waste digester |
| F7-E02 | Uncultured <i>Sphingobacteriales, Sphingobacterium</i> | 0.95 | AY838458 | Soil fungus endophyte |
| F7-E04 | Uncultured, unclassified <i>Bacteroidetes</i> | 0.85 | AY947969 | Temperate river |
| F7-E06 | Uncultured <i>Sphingobacteriales, Chitinophaga</i> | 0.96 | EF018756 | Trembling aspen rhizosphere |
| F7-E07 | Uncultured <i>Sphingobacteriales, Chitinophaga</i> | 0.65 | AJ318177 | Waste gas biofilter |
| F7-F02 | Uncultured <i>Sphingobacteriales, Chitinophaga</i> | 0.77 | EF516017 | Grassland soil |
| F7-F10 | Uncultured, unclassified <i>Bacteroidales</i> | 0.99 | EF559073 | Anaerobic solid waste digester |
| F7-F11 | Uncultured <i>Sphingobacteriales, Sphingobacterium</i> | 0.98 | AY838466 | Soil fungus endophyte |
| F7-G02 | Uncultured <i>Sphingobacteriales, Chitinophaga</i> | 0.87 | EF516601 | Grassland soil |
| F7-G06* | Uncultured, unclassified <i>Flavobacteria</i> | 0.83 | AY218574 | Penguin droppings sediments |
| F7-G07 | Uncultured <i>Sphingobacteriales, Chitinophaga</i> | 0.96 | AY921856 | Farm soil |

122
121
120
119
118
117
116
115
114
113
112
111
110
109
108
107
106
105
104
103
102
101
100

Supplementary Table 1 (continued).

| Phylotypes Clone ID | Phylogenetic affiliation and closest isolate or 16S rRNA phylotype match | S_{ab} score | GenBank ID | Source |
|---|---|----------------|------------|---|
| F7-G08 | Uncultured <i>Sphingobacteriales</i> , <i>Chitinophaga</i> | 0.87 | EF018366 | Trembling aspen rhizosphere |
| F7-H01 | Uncultured <i>Sphingobacteriales</i> , <i>Sphingobacterium</i> | 0.99 | AY838458 | Soil fungus endophyte |
| F7-H06* | Uncultured <i>Sphingobacteriales</i> , <i>Chitinophaga</i> | 0.78 | AY945881 | Sludge-seeded bioreactor |
| F7-H07 | Uncultured <i>Sphingobacteriales</i> , <i>Chitinophaga</i> | 0.98 | DQ297978 | Hydrocarbon contaminated soil |
| F7-H08* | Uncultured, unclassified <i>Flavobacteria</i> | 0.84 | AY218574 | Penguin droppings sediments |
| 29 phylotypes | | | | |
| γ-Proteobacteria | | | | |
| F7-A06 | Uncultured <i>Methylococcales</i> , unclassified <i>Methylococcaceae</i> | 0.77 | EF019533 | Trembling aspen rhizosphere |
| F7-B02 | Uncultured <i>Methylococcales</i> , unclassified <i>Methylococcaceae</i> | 0.76 | EF019533 | Trembling aspen rhizosphere |
| F7-C04 | Uncultured <i>Legionellales</i> , <i>Aquicella</i> | 0.79 | EF540406 | Soil |
| F7-C07* | Uncultured <i>Pseudomonadales</i> , unclassified <i>Pseudomonadaceae</i> | 0.92 | DQ337540 | Swine effluent |
| F7-C11 | Uncultured <i>Legionellales</i> , <i>Aquicella</i> | 0.82 | EF019160 | Trembling aspen rhizosphere |
| F7-E08 | Uncultured <i>Legionellales</i> , <i>Aquicella</i> | 0.81 | EF540406 | Soil |
| F7-G04 | Uncultured, unclassified γ -Proteobacteria | 0.84 | DQ413093 | Anaerobic/Aerobic SBR Reactor |
| F7-G12 | <i>Legionellales</i> , <i>Legionella anisa</i> | 0.93 | AY744776 | N.A. |
| F7-H09 | Uncultured <i>Legionellales</i> , <i>Aquicella</i> | 0.82 | EF540406 | Soil |
| 9 phylotypes | | | | |
| α-Proteobacteria | | | | |
| F7-B09 | Uncultured α -proteobacterium | 0.91 | AB274850 | Medium with aromatic and saturated hydrocarbons |
| F7-C10 | Uncultured α -proteobacterium | 0.86 | DQ123621 | Hospital water network (amoebal co-culture) |
| 2 phylotypes | | | | |
| δ-Proteobacteria | | | | |
| F7-F03 | Uncultured δ -proteobacterium | 0.73 | AY792293 | Humic lake |
| 1 phylotype | | | | |
| Actinobacteria | | | | |
| F7-A01 | <i>Actinomycetales</i> , <i>Mycobacterium nonchromogenicum</i> | 0.91 | DQ058406 | Clinical isolate |
| F7-A09 | <i>Rubrobacterales</i> , <i>Solirubrobacter</i> | 0.85 | AY234700 | Soil |
| F7-B03 | <i>Actinomycetales</i> , <i>Mycobacterium nonchromogenicum</i> | 0.92 | DQ058406 | Clinical isolate |
| 3 phylotypes | | | | |

145
144
143
142
141
140
139
138
137
136
135
134
133
132
131
130
129
128
127
126
125
124
123

Supplementary Table 1 (continued).

| Phylotypes Clone ID | Phylogenetic affiliation and closest isolate or 16S rRNA phylotype match | S_{ab} score | GenBank ID | Source |
|------------------------------|---|----------------|------------|-----------------------------|
| <i>Acidobacteria</i> | | | | |
| F7-C06 | Uncultured bacterium, clone Elev_16S_1653, genus Gp6 | 0.96 | EF020191 | Trembling aspen rhizosphere |
| F7-G11 | Uncultured bacterium, clone AKAU3910, genus Gp6 | 0.82 | DQ125778 | Uranium contaminated soil |
| 2 phylotypes | | | | |
| <i>Planctomycetes</i> | | | | |
| F7-G09 | Uncultured <i>Planctomycetales</i> , <i>Isosphaera</i> | 0.81 | EF516309 | Grassland soil |
| 1 phylotype | | | | |
| Genera Incertae | | | | |
| F7-A04 | Uncultured bacterium, genera incertae sedis TM7 | 0.78 | AM116752 | Uranium contaminated soil |
| 1 phylotype | | | | |
| Unclassified Bacteria | | | | |
| F7-C03* | Uncultured bacterium, clone KD6-72 | 0.88 | AY218768 | Penguin droppings sediments |
| F7-E05 | Uncultured bacterium, clone 300I-B02 | 0.75 | AY661981 | Contaminated groundwater |
| F7-F05 | Uncultured soil bacterium, clone C06 | 0.51 | AF507679 | Soil |
| F7-H11 | Uncultured eubacterium, vadinCA02 | 0.57 | U81706 | Wine distillery waste |
| | | <u>ID%</u> | | |
| F7-D03** | Uncultured bacterium, clone Amb_16S_1804 | 97 | EF019157 | Trembling aspen rhizosphere |
| F7-D06** | Uncultured candidate division OP11, bacterium clone DA7 | 93 | AY193205 | Contaminated aquifer |
| 6 phylotypes | | | | |

N.A. = not available from GenBank record.

* = clone related to bacterial phylotype detected in fecal/intestinal environment.

** = sequences with S_{ab} scores <0.50, classified in terms of identity percentage relative to best GenBank matching sequence as determined with BLAST.

146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
Supplementary Table 2. Characterization of bacterial phylotypes detected in fraction F12 (60-65% G+C) according to the RDP-II Seqmatch tool.

| Phylotypes | Phylogenetic affiliation and closest isolate or 16S rRNA | | | |
|--------------------------------|---|-----------------------|------------|-------------------------------|
| Clone ID | phylotype match | S _{ab} score | GenBank ID | Source |
| <i>α-Proteobacteria</i> | | | | |
| F12-A02 | Uncultured <i>Sphingomonadales, Sphingomonas</i> | 0.92 | AY439192 | Soil |
| F12-A03 | Uncultured <i>Rhizobiales, Rhodoplanes</i> | 0.95 | EF020212 | Trembling aspen rhizosphere |
| F12-A04 | Uncultured <i>Rhizobiales, Rhodoplanes</i> | 0.92 | EF020062 | Trembling aspen rhizosphere |
| F12-A05 | Uncultured <i>Rhizobiales, Rhodoplanes</i> | 0.99 | DQ453806 | Soil |
| F12-A06 | Uncultured <i>Sphingomonadales, Sphingomonas</i> | 0.89 | AY921836 | Farm soil |
| F12-B01 | Uncultured <i>Rhodospirillales, unclassified Acetobacteraceae</i> | 0.95 | Y12596 | Grassland soil |
| F12-B05 | Uncultured, unclassified <i>Rhizobiales</i> | 0.95 | AB240333 | Rhizosphere biofilm |
| F12-B09 | Uncultured <i>Sphingomonadales, Sphingomonas</i> | 0.94 | EF516245 | Grassland soil |
| F12-B10 | Uncultured <i>Sphingomonadales, Sphingomonas</i> | 0.95 | DQ521501 | Lake ice cover |
| F12-B12 | <i>Rhizobiales, Rhizobium mongolense</i> | 0.88 | U89816 | Alfalfa roots nodule |
| F12-C02 | Uncultured, unclassified <i>Rhodospirillales</i> | 0.84 | AY493936 | Soil |
| F12-C04 | Uncultured <i>Sphingomonadales, Sphingomonas</i> | 0.71 | EF018252 | Trembling aspen rhizosphere |
| F12-C09 | Uncultured <i>Sphingomonadales, Sphingomonas</i> | 0.82 | EF019685 | Trembling aspen rhizosphere |
| F12-D01 | Uncultured <i>Sphingomonadales, Sphingomonas</i> | 0.93 | EF018252 | Trembling aspen rhizosphere |
| F12-D06 | Uncultured <i>Rhizobiales, Balneimonas</i> | 0.84 | DQ125646 | Uranium contaminated soil |
| F12-D10 | Uncultured <i>Sphingomonadales, Sphingomonas</i> | 0.91 | AY102315 | Heavy metal contaminated soil |
| F12-E01 | Uncultured <i>Sphingomonadales, Sphingomonas</i> | 0.94 | EF018252 | Trembling aspen rhizosphere |
| F12-E03 | Uncultured <i>Sphingomonadales, Sphingomonas</i> | 0.89 | EF018252 | Trembling aspen rhizosphere |
| F12-E08 | Unclassified <i>α-Proteobacteria</i> | 0.94 | AB245344 | Ginseng field |
| F12-E10 | Uncultured <i>Sphingomonadales, Sphingomonas</i> | 0.95 | EF018252 | Trembling aspen rhizosphere |
| F12-E11 | Uncultured, unclassified <i>Rhizobiales</i> | 0.95 | AB240333 | Rhizosphere biofilm |
| F12-F06 | Uncultured <i>Rhizobiales, Rhodoplanes</i> | 0.99 | EF020212 | Trembling aspen rhizosphere |
| F12-F08 | Uncultured <i>Sphingomonadales, Sphingomonas</i> | 0.92 | AY102315 | Heavy metal contaminated soil |
| F12-F11 | Uncultured, unclassified <i>Rhizobiales</i> | 0.90 | EF018463 | Trembling aspen rhizosphere |
| F12-G01 | Uncultured <i>Sphingomonadales, Sphingomonas</i> | 0.68 | EF018252 | Trembling aspen rhizosphere |
| F12-G07 | Uncultured <i>Rhizobiales, unclassified Hyphomicrobiaceae</i> | 0.90 | EF516797 | Grassland soil |
| F12-G08 | <i>Rhizobiales, Mesorhizobium sp.</i> | 0.91 | AF156710 | Oxygen limited bioreactor |
| F12-G09 | Uncultured, unclassified <i>Rhizobiales</i> | 0.97 | EF020226 | Trembling aspen rhizosphere |
| F12-G10 | Uncultured <i>Sphingomonadales, Novosphingobium</i> | 0.97 | DQ129610 | Urban aerosol |
| F12-G11 | Uncultured <i>Rhizobiales, unclassified Hyphomicrobiaceae</i> | 0.88 | EF516797 | Grassland soil |
| F12-G12 | Uncultured <i>Sphingomonadales, Sphingomonas</i> | 0.89 | EF540399 | Solid waste from oil shale |
| F12-H05 | Uncultured <i>Sphingomonadales, Sphingomonas</i> | 0.94 | EF018252 | Trembling aspen rhizosphere |
| F12-H06 | Uncultured <i>Rhizobiales, Devosia</i> | 0.86 | DQ125710 | Uranium contaminated soil |
| F12-H10 | Uncultured <i>Rhodospirillales, unclassified Acetobacteraceae</i> | 0.91 | EF019851 | Trembling aspen rhizosphere |

34 phylotypes

| 191 | 190 | 189 | 188 | 187 | 186 | 185 | 184 | 183 | 182 | 181 | 180 | 179 | 178 | 177 | 176 | 175 | 174 | 173 | 172 | 171 | 170 | 169 | |
|---|--|--|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----------------------|------------|------------------------------|
| Supplementary Table 2 (continued). | | | | | | | | | | | | | | | | | | | | | | | |
| Phylotypes | | Phylogenetic affiliation and closest isolate or 16S rRNA | | | | | | | | | | | | | | | | | | | | | |
| Clone ID | phylogroup | phylogroup match | | | | | | | | | | | | | | | | | | | S _{ab} score | GenBank ID | Source |
| β-Proteobacteria | | | | | | | | | | | | | | | | | | | | | | | |
| F12-B03 | Uncultured, unclassified β-Proteobacteria | | | | | | | | | | | | | | | | | | | | 0.90 | EF032745 | Cyanobacterial mat lava cave |
| F12-B04 | Uncultured, unclassified β-Proteobacteria | | | | | | | | | | | | | | | | | | | | 0.94 | EF516025 | Grassland soil |
| F12-D03 | Uncultured, unclassified β-Proteobacteria | | | | | | | | | | | | | | | | | | | | 0.95 | AY921787 | Farm soil |
| F12-D04 | <i>Burkholderiales</i> , unclassified <i>Oxalobacteraceae</i> | | | | | | | | | | | | | | | | | | | | 0.98 | AB252072 | Deforested tropical peatland |
| F12-F01 | Uncultured <i>Burkholderiales</i> , <i>Thiobacter</i> | | | | | | | | | | | | | | | | | | | | 0.86 | EF516234 | Grassland soil |
| F12-F09 | Uncultured <i>Burkholderiales</i> , unclassified <i>Comamonadaceae</i> | | | | | | | | | | | | | | | | | | | | 0.93 | EF018197 | Trembling aspen rhizosphere |
| F12-H02 | Uncultured, unclassified β-Proteobacteria | | | | | | | | | | | | | | | | | | | | 0.94 | EF018790 | Trembling aspen rhizosphere |
| F12-H03 | Uncultured, unclassified β-Proteobacteria | | | | | | | | | | | | | | | | | | | | 0.91 | EF520946 | Soil |
| 8 phylotypes | | | | | | | | | | | | | | | | | | | | | | | |
| γ-Proteobacteria | | | | | | | | | | | | | | | | | | | | | | | |
| F12-A07 | <i>Xanthomonadales</i> , <i>Rhodanobacter lindaniclasticus</i> | | | | | | | | | | | | | | | | | | | | 0.98 | AB245365 | Ginseng field |
| F12-C07 | Uncultured <i>Xanthomonadales</i> , unclassified <i>Xanthomonadaceae</i> | | | | | | | | | | | | | | | | | | | | 0.75 | AJ863337 | Poplar tree rhizosphere |
| F12-E02 | <i>Xanthomonadales</i> , <i>Thermomonas dokdonensis</i> | | | | | | | | | | | | | | | | | | | | 0.89 | EF100698 | Soil |
| F12-H01 | Uncultured <i>Xanthomonadales</i> , <i>Rhodanobacter sp.</i> | | | | | | | | | | | | | | | | | | | | 0.99 | AY922074 | Farm soil |
| F12-H11 | Uncultured <i>Xanthomonadales</i> , unclassified <i>Xanthomonadaceae</i> | | | | | | | | | | | | | | | | | | | | 0.88 | DQ123670 | PAH-contaminated soil |
| 5 phylotypes | | | | | | | | | | | | | | | | | | | | | | | |
| Actinobacteria | | | | | | | | | | | | | | | | | | | | | | | |
| F12-A10 | Unclassified <i>Actinobacteria</i> | | | | | | | | | | | | | | | | | | | | 0.83 | AM157320 | Soil |
| F12-A12 * | <i>Actinomycetales</i> , <i>Rhodococcus sp.</i> | | | | | | | | | | | | | | | | | | | | 0.99 | DQ337561 | Swine effluent applied soil |
| F12-B08 * | <i>Actinomycetales</i> , <i>Rhodococcus sp.</i> | | | | | | | | | | | | | | | | | | | | 0.98 | DQ337561 | Swine effluent applied soil |
| F12-C11 | Uncultured <i>Actinomycetales</i> , <i>Pseudonocardia sp.</i> | | | | | | | | | | | | | | | | | | | | 0.92 | AY921961 | Farm soil |
| F12-D02 | <i>Actinomycetales</i> , <i>Arthrobacter sp.</i> | | | | | | | | | | | | | | | | | | | | 1.00 | AY452081 | Lead-zinc mine tailings |
| F12-D07 | <i>Actinomycetales</i> , <i>Streptomyces violaceolatus</i> | | | | | | | | | | | | | | | | | | | | 0.99 | AF503497 | N.A. |
| F12-D08 | <i>Actinomycetales</i> , <i>Streptomyces ferralitis</i> | | | | | | | | | | | | | | | | | | | | 0.95 | AY262826 | Soil |
| F12-E06 | Uncultured, unclassified <i>Actinomycetales</i> | | | | | | | | | | | | | | | | | | | | 0.71 | DQ017947 | Upland stream |
| F12-F04 | <i>Actinomycetales</i> , <i>Nocardioideis sp.</i> | | | | | | | | | | | | | | | | | | | | 0.79 | D87974 | Marine environments |
| F12-G02 | <i>Actinomycetales</i> , <i>Kribbella</i> | | | | | | | | | | | | | | | | | | | | 0.96 | EF472955 | Soil |
| F12-G04 | <i>Actinomycetales</i> , <i>Nocardioideis aquiterrae</i> | | | | | | | | | | | | | | | | | | | | 0.83 | AF529063 | Groundwater |
| F12-G06 | <i>Actinomycetales</i> , <i>Corynebacterium sp.</i> | | | | | | | | | | | | | | | | | | | | 0.87 | DQ219354 | Coastal seaweed |
| F12-H08 | <i>Actinomycetales</i> , <i>Streptomyces sp.</i> | | | | | | | | | | | | | | | | | | | | 0.97 | DQ663193 | N.A. |
| 13 phylotypes | | | | | | | | | | | | | | | | | | | | | | | |
| Acidobacteria | | | | | | | | | | | | | | | | | | | | | | | |
| F12-A01 | Uncultured <i>Acidobacterium</i> | | | | | | | | | | | | | | | | | | | | 0.66 | EF515883 | Grassland soil |
| F12-A11 | Uncultured, unclassified <i>Acidobacteriaceae</i> | | | | | | | | | | | | | | | | | | | | 0.86 | EF516240 | Grassland soil |
| F12-C03 | Uncultured, unclassified <i>Acidobacteriaceae</i> | | | | | | | | | | | | | | | | | | | | 0.97 | AJ532723 | Uranium mining waste pile |
| F12-D09 | Uncultured <i>Acidobacterium</i> | | | | | | | | | | | | | | | | | | | | 0.98 | EF516558 | Grassland soil |
| F12-D11 | Uncultured <i>Acidobacterium</i> | | | | | | | | | | | | | | | | | | | | 0.82 | AY922076 | Farm soil |
| 5 phylotypes | | | | | | | | | | | | | | | | | | | | | | | |

192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214

Supplementary Table 2 (continued).

| Phylotypes | Phylogenetic affiliation and closest isolate or 16S rRNA | | | |
|--------------------------------|--|----------------|------------|-----------------------------|
| Clone ID | phylotype match | S_{ab} score | GenBank ID | Source |
| <i>Gemmatimonadetes</i> | | | | |
| F12-A09 | Uncultured <i>Gemmatimonadales</i> , <i>Gemmatimonas</i> | 1.00 | AY921714 | Farm soil |
| F12-B06 | Uncultured <i>Gemmatimonadales</i> , <i>Gemmatimonas</i> | 0.97 | EF516879 | Grassland soil |
| F12-C01 | Uncultured <i>Gemmatimonadales</i> , <i>Gemmatimonas</i> | 0.73 | AY921800 | Farm soil |
| F12-C06 | Uncultured <i>Gemmatimonadales</i> , <i>Gemmatimonas</i> | 0.87 | AY922107 | Farm soil |
| F12-E04 | Uncultured <i>Gemmatimonadetes</i> | 0.90 | AKYG871 | Farm soil |
| F12-E07 | Uncultured <i>Gemmatimonadales</i> , <i>Gemmatimonas</i> | 0.76 | AY921994 | Farm soil |
| F12-H04 | Uncultured <i>Gemmatimonadales</i> , <i>Gemmatimonas</i> | 0.94 | EF516206 | Grassland soil |
| 7 phylotypes | | | | |
| <i>Chloroflexi</i> | | | | |
| F12-A08* | Uncultured <i>Chloroflexales</i> , <i>Roseiflexus</i> | 0.56 | X84565 | Activated sludge |
| F12-C08 | Uncultured, unclassified Chloroflexi | 0.89 | AY922047 | Farm soil |
| F12-C12 | Uncultured <i>Chloroflexales</i> , <i>Roseiflexus</i> | 0.56 | DQ129248 | Urban aerosol |
| F12-D12 | Uncultured <i>Chloroflexales</i> , <i>Roseiflexus</i> | 0.74 | AF507694 | Forest soil |
| 4 phylotypes | | | | |
| <i>Planctomycetes</i> | | | | |
| F12-B02 | Uncultured <i>Planctomycetales</i> , unclassified <i>Planctomycetaceae</i> | 0.62 | EF020287 | Trembling aspen rhizosphere |
| F12-B07 | Uncultured <i>Planctomycetales</i> , <i>Isosphaera</i> | 0.80 | EF516255 | Grassland soil |
| F12-G05 | Uncultured <i>Planctomycetales</i> , <i>Isosphaera</i> | 0.79 | EF516786 | Grassland soil |
| 3 phylotypes | | | | |
| Unclassified bacteria | | | | |
| F12-C10 | Uncultured bacterium clone Amb_16S_854 | 0.85 | EF018585 | Trembling aspen rhizosphere |
| 1 phylotype | | | | |

NA = not available from GenBank record.

* = clone related to bacterial phylotype detected in fecal/intestinal environment.

215 SUPPLEMENTARY FIGURE LEGENDS

216 **Supplemental Figure 1.** Agarose gel electrophoresis (inverted ethidium bromide coloration) of
217 gradient PCR amplification of *intl* genes performed at temperatures ranging from 45 to 56°C.
218 Panel (a) depicts results using ~50 ng per reaction of genomic DNA from a reference *E. coli*
219 strain (SK1592, pDU202) harboring a class 1 integron whereas Panel (b) illustrates results
220 obtained when using ~20 ng per reaction of genomic DNA from *E. coli* strain J53.3::Tn7
221 carrying a class 2 integron. Panel (c) shows the amplification of *intl* genes from metagenomic
222 DNA extracted from manure supplemented agricultural soil (AMS4, ~20ng of template DNA per
223 reaction). Numbers on top of each lane refer to the temperature gradient used while arrows at the
224 right depict unspecific products smaller or larger than the expected amplicon (~491bp). M=DNA
225 marker, 100 bp ladder (Invitrogen, Carlsbad, CA)

226
227 **Supplemental Figure 2.** Alignment of partial sequences of environmental integrases having an
228 extension of the additional domain downstream the patch III region relative to functional
229 integrases (classes 1-4, IntI4 numbering system). Blue columns indicate conserved positions of
230 catalytic function whereas the red lines illustrate the position of insertions located within the IntI-
231 specific additional domain.

232
233 **Supplemental Figure 3.** Partial alignment of environmental integrases over 89% identical to the
234 class 2 integrase (GenBank accession no. AAT72891). The red arrow shows the presence of a
235 glutamine residue among 11 IntI2-like phylotypes at position 179 which corresponds to the ochre
236 codon present in conventional class 2 integrases. The degree of conservation in each column in

237 the alignment is emphasized relative to the class 2 integrase sequence as follows: black (100%
238 identical), gray (100% similar). Gaps are indicated by “-” characters.

239

240 **Supplemental Figure 4.** Alignment of three partial sequences of IntI2-like integrases detected in
241 fraction F7 (40-45% G+C). Red arrows indicate the position of distinctive, new features relative
242 to canonical class 2 integrases (GenBank accession nos. BAB12601 and ACI32876). The degree
243 of conservation in each column is emphasized as follows: black (100% identical), gray (100%
244 similar). Gaps are indicated by “-” characters.

245

246

247 **Supplemental Figure 5.** Consensus Neighbor-Joining tree depicting the relationship of novel,
248 “non-clinical” IntI elements ($\leq 98\%$ identical) retrieved from fraction F7 (40-45% G+C, green
249 squares) and fraction F12 (60-65% G+C, blue squares) relative to their closest matches among
250 IntI elements detected in bulk DNA and those previously described (red and black triangles,
251 respectively). The number of compared positions was 121. The percentage of bootstrap
252 replications (1000) supporting the branching order are shown at the nodes. Values of bootstrap
253 support less than 30% are not shown. GenBank accession nos. follow the designation of
254 reference and outgroup sequences (XerC and XerD recombinases).

255

256

257

258

259

260

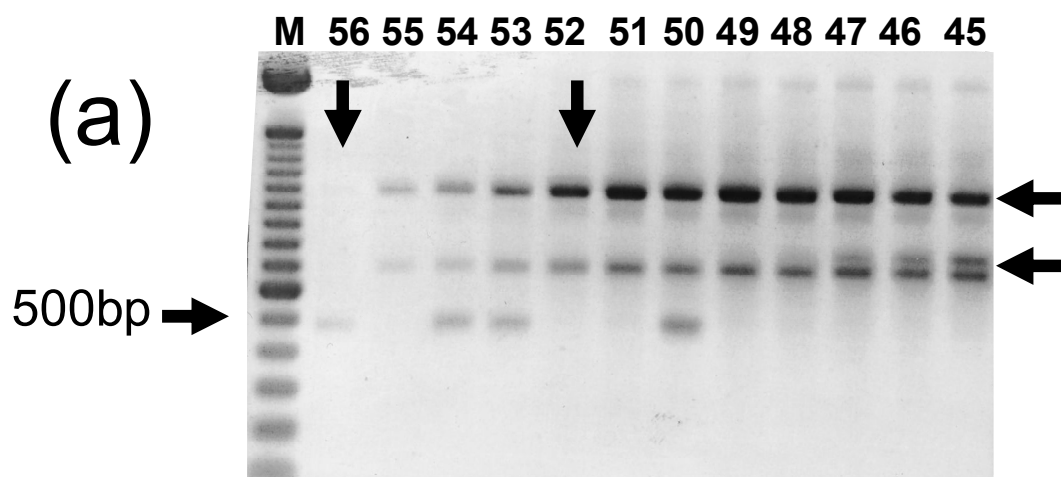
261

262

263

264

265



266

267

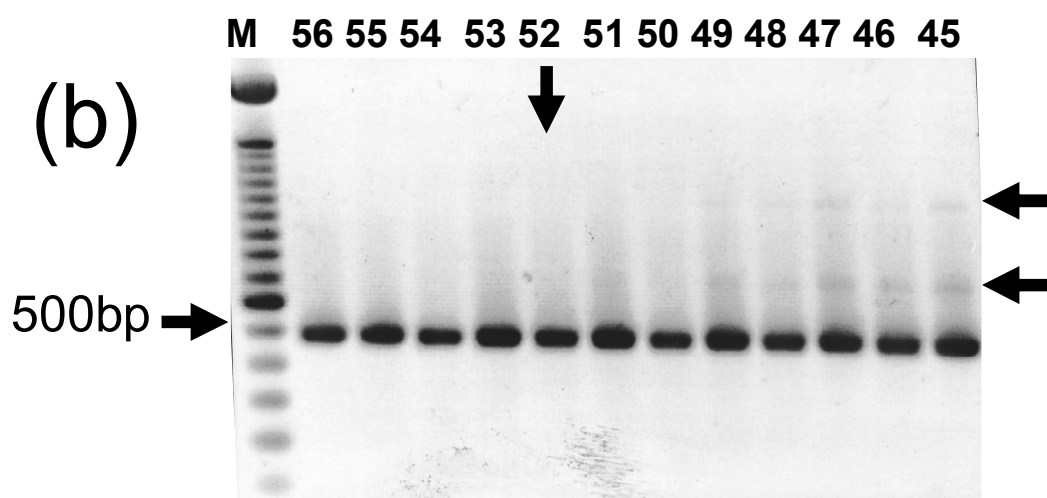
268

269

270

271

272



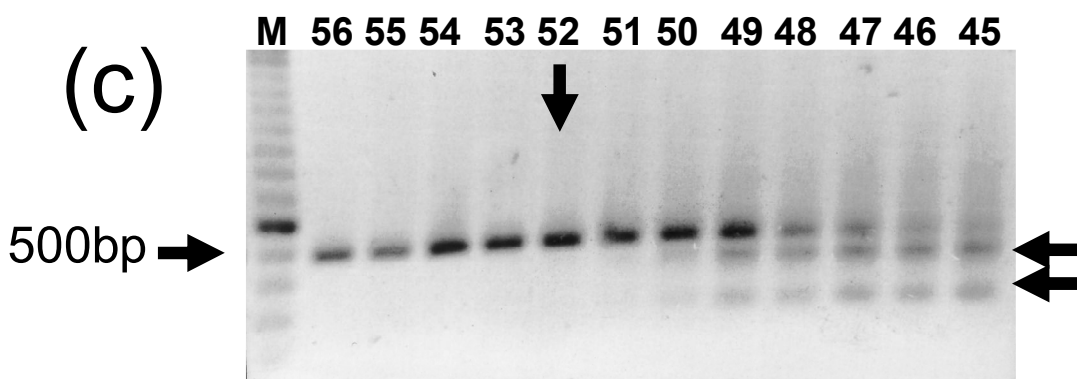
273

274

275

276

277



278
279
280
281
282
283

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---------|-------------------------------|----|-----|-----|-----|-----|-----|-----|-----|----|----|----|----|-----|-------|----|----|-------|-------|-------|----|---|---|---|---|---|---|---|
| | 175 | | 205 | 209 | 219 | 223 | 232 | | 241 | | | | | | | | | | | | | | | | | | | |
| IntI4 | PHLKEQIALAKRYYDRDLHQKNYGGVWLP | AL | KE | KY | PN | AP | YEF | RWH | YLF | PS | FQ | LS | LD | PES | ----- | VM | RR | HH | M | NET | VL | | | | | | | |
| IntI2 | PAIKXLIEQARLIQQDDN-LQGVG-PSLP | F | LD | HK | Y | PS | AY | RQA | AM | FV | F | PS | SL | CN | HP | Y | N | ----- | K | LC | R | H | H | L | H | S | V | |
| IntI3 | PRLRAQLIQVRAVWGQDR-ATGRGGVYL | PH | A | L | E | R | K | Y | P | R | A | G | S | W | A | F | W | F | ----- | V | E | R | R | H | L | F | E | E |
| IntI1 | PSLREQLSRARAWWLKDQ-AEGRSGVAL | P | D | A | L | E | R | K | Y | P | R | A | G | H | S | W | P | F | W | ----- | V | V | R | R | H | H | M | Y |
| MS8-F05 | EGLKSILKARRITYEEDL-AAGIGEASL | P | T | P | I | S | N | K | Y | P | N | A | G | K | E | W | I | W | Q | ----- | H | V | F | A | G | S | R | L |
| SAF-F5 | APLQRHLERVKVQHEEDL-EAGFGRVYL | P | F | A | L | D | R | K | Y | K | N | A | A | R | E | W | C | W | Q | ----- | V | E | F | V | R | R | R | H |
| HI-G2 | GLFERQIEKRRILHEEDL-AAGFGTVFL | P | T | A | L | A | R | K | Y | P | N | A | A | R | E | F | L | W | Q | ----- | D | P | P | K | Q | S | R | H |
| SAF-A8 | ASLEGLVVKMKAQHEQDL-EEGFGSVYL | P | E | A | L | A | R | K | Y | K | S | A | R | E | W | C | W | Q | ----- | P | S | T | R | L | S | M | D | |
| UR-C9 | ASLERHLTKIKAQHEQDL-EEGLGEVYL | P | E | A | L | A | R | K | Y | P | N | A | A | R | E | W | G | W | Q | ----- | A | E | - | Q | R | R | H | A |
| MONA-G2 | EPLRRHVARMKAQHEQDL-EEGFGAVS | I | P | G | A | L | A | R | K | Y | P | N | A | Q | R | E | W | A | W | ----- | A | W | D | Q | L | P | C | R |
| XerC | AWIEHWLDL-R-----DL----- | | | | | | | | | | | | | | | | | | ----- | | | | | | | | | |
| XerD | YWLETYLEHGR---PWLL----- | | | | | | | | | | | | | | | | | | | ----- | | | | | | | | |

<-----Additional Domain-----><-----><-----Additional Domain-->

Patch III

284

285

286

287

288

289

290

291

292

293

294

| | 143 | 150 | 160 | 170 | 180 | 190 | 200 | |
|-----------|-------|-------------|-----------|-----------------|--------|---------|------------|-----|
| IntI2 | RVKDF | DFDNGCITVHD | GKGGKSRNS | SLLPTRLIPAIKXLI | EQARLI | IQQDDN- | LQGVGPSLP | 201 |
| CM-A3 | RVKDF | DFDNGCITVHD | GKGGKSRNS | SLLPTRLIPAIKXLI | EQARLI | IQQDDN- | LQGVGPSLP | 201 |
| CM-F9 | RVKDV | DFDNGCITVHD | GKGGKSRNS | SLLPTRLIPAIKXLI | EQARLI | IQQDDN- | LQGVGPSLP | 201 |
| CM-B9 | RVKDV | DFDNGCITVHD | GKGGKSRNS | SLLPTRLIPAIKXLI | EQARLI | IQQDDN- | LQGVGPSLP | 201 |
| HI-B9 | RVKDV | DFDNGCITVHD | GKGGKSRNS | SLLPTRLIPAIKXLI | EQARLI | IQQDDN- | LQGVGPSLP | 201 |
| SAF-D5 | RVKDV | DFDNGCITVHD | GKGGKSRNS | SLLPTRLIPAIKXLI | EQARLI | IQQDDN- | LQGVGPSLP | 201 |
| SLV-E10 | RVKDV | DFDNGCITVHD | GKGGKSRNS | SLLPTRLIPAIKXLI | EQARLI | IQQDDN- | LQGVGPSLP | 201 |
| SLV-F2 | RVKDL | DFDNGCITVHD | GKGGKSRNS | LLPTGLIAAIKQLI | DRVLVI | IQQEDN- | AQGVGPSLP | 201 |
| BIO-A1 | RVKDV | DFDNGCITVHD | GKGGKSRNS | SLLPTRLIPAIKQLI | EQARLI | IQQDDN- | LQGVGPSLP | 201 |
| BIO-A4 | RVKDL | DFDNGCITVHD | GKGGKSRNS | SLLPTRLIPAIKQLI | EQARLI | IQQDDN- | LQGVGPSLP | 201 |
| BIO-H9 | RVKDV | DFDNGCITVHD | GKGGKSRNS | SLLPTRLIPAIKQLI | EQARLI | IQQDDN- | LQGVGPSLP | 201 |
| BIO-A10 | RVKDV | DFDNGCITVHD | GKGGKSRNS | SLLPTRLIPAIKQLI | EQARLI | IQQDDN- | LQGVGPSLP | 201 |
| AMS4-A2 | RVKDL | DFDNGCITVHD | GKGGKSRNS | SLLPTRLIPAIKQLI | EQARLI | IQQDDN- | LQGVGPSLP | 201 |
| AMS4-D02 | RVKDL | DFDNGCITVHD | GKGGKSRNS | SLLPTRLIPAIKQLI | EQARLI | IQQDDN- | YLDKHRPSLP | 202 |
| AMS4-E02 | RVKDV | DFDNGCITVHD | GKGGKSRNS | SLLPTRLIPAIKQLI | EQARLI | IQQDDN- | LQGVGPSLP | 201 |
| AMS23-C12 | RVKDV | DFDNGCITVHD | GKGGKSRNS | LLPTGLIAAIKQLI | DRVLVI | IQQEDN- | AQGVGPSLP | 201 |
| AMS23-E01 | RVKDV | DFDNGCITVHD | GKGGKSRNS | SLLPTRLIPAIKQLI | EQARLI | IQQDDN- | LQGVGPSLP | 201 |
| AMS25-D03 | RVKDL | DFDNGCISVHD | GKGGKSRNS | SLLPTRLIPAIKQLI | DRVLVI | IQQEDN- | AQGVGLSLP | 201 |



179

308

309

310

