**Supplementary Figure Legends.**

Figure S1: Photographs of A) AB-end: pH 1.07, 43°C; B) AB-front: pH 0.99, 39°C; and C) UBA: pH 1.1, 41°C.

Figure S2. Diagram of the genome of *Leptospirillum* Group III (outer circle) and orthologs in *Leptospirillum* Group II (inner circle, color-coded by scaffold). Histogram of percent identity between orthologs is black bars on inner ring. Heat maps of protein identification (NSAF) are given by six middle rings (grey: no expression): *Leptospirillum* Group II samples UBA, ABfront, ABend, *Leptospirillum* Group III sample UBA, ABfront and ABend (from inside to outside).

Figure S3. Protein abundance (NSAF) values of orthologous proteins in *Leptospirillum* Groups II and III. Red: overrepresented, green: underrepresented, black: median, grey: no expression. Heat maps were made using Cluster 3.0 software, median-centering clusters by organism dataset, and visualized using TreeView application. Most abundant proteins in four clusters (number of examples in parentheses): Cluster A: Ribosomal proteins (8), transport and secretion (6), electron transport (6), other (25). Cluster B (only part shown): Electron transport (6), vitamin and cofactor synthesis (5), transport and secretion (4), other (27). Cluster C: Ribosomal proteins (6), information processing (6), energy-related (5), amino acid metabolism (5), others (22). Cluster D: Energy-related (9), flagella, chemotaxis and signal transduction (5), information processing (4), others (26).

24 Figure S4. ClustalW (Thompson et al., 1994) phylogram of the 20S proteasome subunits α
25 and β gene tree. Sequences from *Leptospirillum* Groups II and III (red and blue dots,
26 respectively), Actinobacteria (green) and Unassigned bacteria (light grey) from the AMD
27 community genomics datasets are included. Archaeal group includes other unassigned
28 scaffolds form the AMD genomic datasets (dark grey).
29
30 Figure S5. A) Diagram of RNA polymerase, sigma 70 subunit, RpoD family (RpoD) protein
31 with fused Adenine/Guanosine phosphoribosyltransferase (Apt) domain in *Leptospirillum*
32 group II (adapted from Conserved Domains results - NCBI). RpoD and Apt are two proteins
33 in *Leptospirillum* group III. B) Protein sequence showing protein coverage from all sites in
34 blue. Underlined is the region that fuses both genes, with peptide coverage in purple.
35
36

**Supplementary figures**

Figure S1.

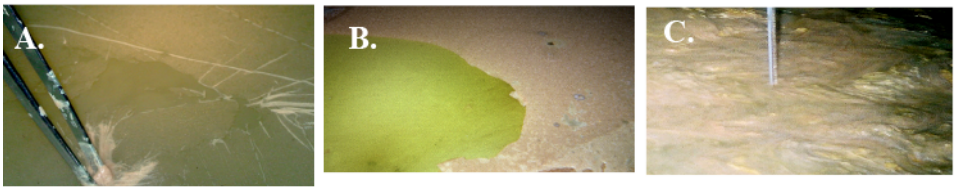


Figure S2.

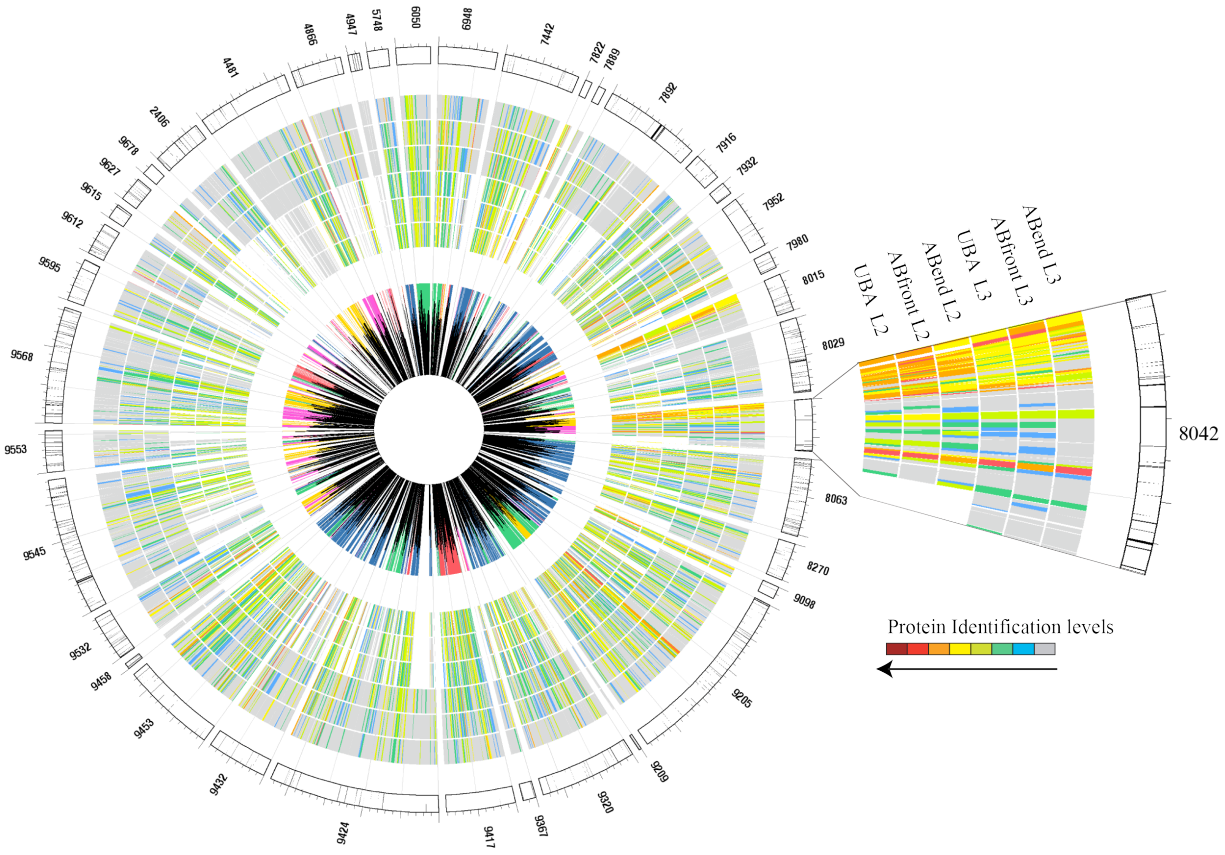

UBA L2
ABfront L2
ABend L2
UBA L3
ABfront L3
ABend L3
8042
Protein Identification levels

Figure S3.

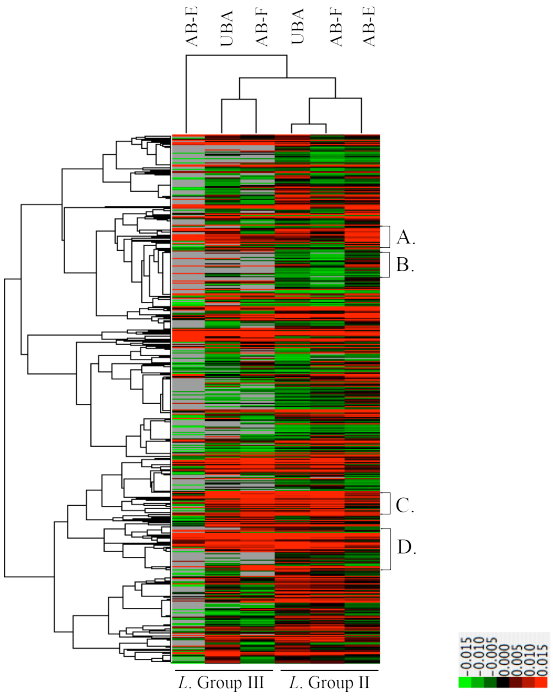

AB-E
UBA
AB-F
UBA
AB-F
AB-E
A.
B.
C.
D.
L. Group III
L. Group II
-0.015
-0.010
-0.005
0.000
0.005
0.010
0.015

Figure S4.

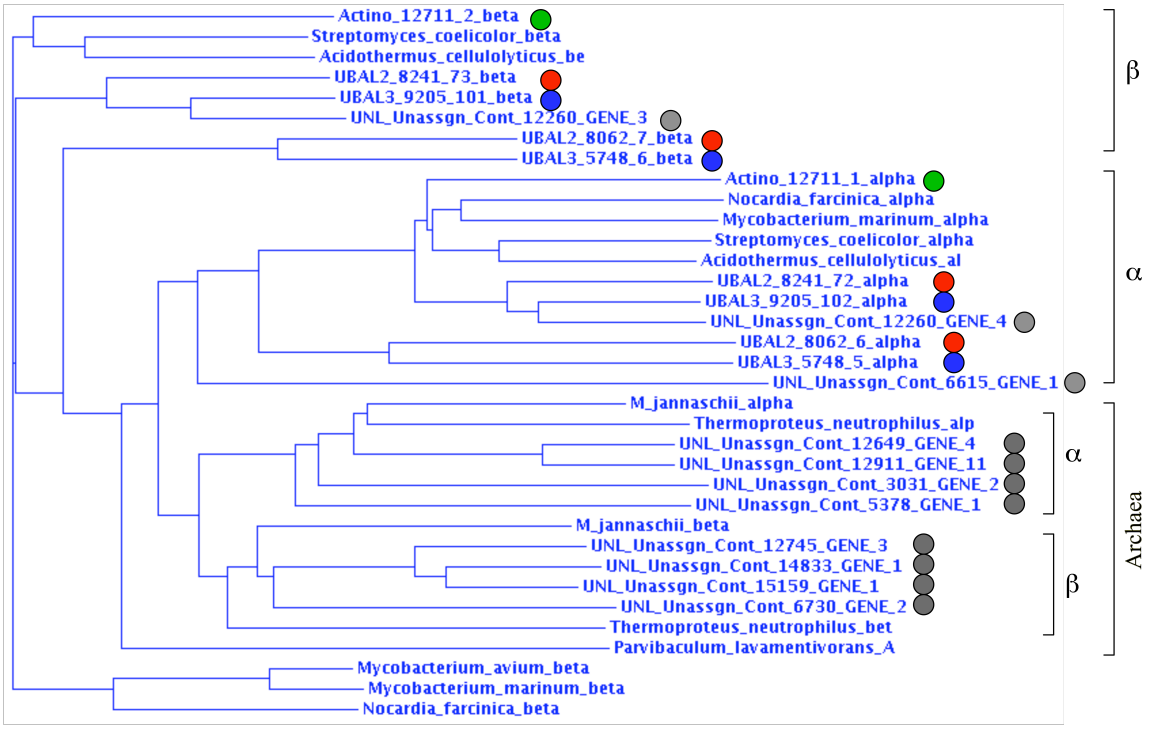
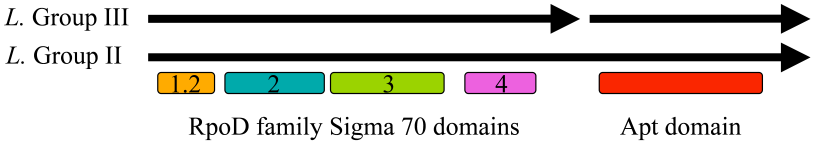

Actino_12711_2_beta
Streptomyces_coelicolor_beta
Acidothermus_cellulolyticus_be
UBAL2_8241_73_beta
UBAL3_9205_101_beta
UNL_Unassgn_Cont_12260_GENE_3
UBAL2_8062_7_beta
UBAL3_5748_6_beta
Actino_12711_1_alpha
Nocardia_farcinica_alpha
Mycobacterium_marinum_alpha
Streptomyces_coelicolor_alpha
Acidothermus_cellulolyticus_al
UBAL2_8241_72_alpha
UBAL3_9205_102_alpha
UNL_Unassgn_Cont_12260_GENE_4
UBAL2_8062_6_alpha
UBAL3_5748_5_alpha
UNL_Unassgn_Cont_6615_GENE_1
M_jannaschii_alpha
Thermoproteus_neutrophilus_alp
UNL_Unassgn_Cont_12649_GENE_4
UNL_Unassgn_Cont_12911_GENE_11
UNL_Unassgn_Cont_3031_GENE_2
UNL_Unassgn_Cont_5378_GENE_1
M_jannaschii_beta
UNL_Unassgn_Cont_12745_GENE_3
UNL_Unassgn_Cont_14833_GENE_1
UNL_Unassgn_Cont_15159_GENE_1
UNL_Unassgn_Cont_6730_GENE_2
Thermoproteus_neutrophilus_bet
Parvibaculum_lavamentivorans_A
Mycobacterium_avium_beta
Mycobacterium_marinum_beta
Nocardia_farcinica_beta
β
α
α
β
Archaea

Figure S5.

A




B

```
>UBA_LeptoII_Scaffold_8135_GENE_0103c
MDGTIPENDSPWFSPDEQTEDSGKSVSRADEANEPLERGDLEDDHLSALGSYLKKLQKSHLLTFEEEQDLAKRVRAGDE
SARQAMIQSNLRLVVSIAKRYIGRGLLLEDLIEEGNLGLMKAVSKFDPDKGFRFSTYASWWIRQAIERGIINQGHMIRL
PVHMMEKVNAYLNAIERQVADGQQPDIETLARKNGLKVRDLQEIQALLKNTVSLDTPIGDREDSTLKDILVDPTSFSPI
STIESREEGDRLRESLALLKDKERRVIHLRFGLPGEGPEEGMTLEAIGQMLGVTRERVRQIEASAMAKLRAYMEDPSEF
RKKHRWREKHNMDFNKWVRQIPDFPKKGILFYDLMPLFGEPTAFHALIDTMVEPYRDQGIRKVVGIEARGFILAAPIAD
RLKAGFVPIRKKGKLPGAVHEISYSLEYGKDSIAIHQGAIRPGDRVLLVDDLLATGGTALAALELIRKDGGQLAGVHFV
AELVGLGGRDRLRQKDIRSVLTYEV
```

53 **REFERENCE**

54

55 Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTAL W: improving the
56 sensitivity of progressive multiple sequence alignment through sequence weighting,
57 positions-specific gap penalties and weight matrix choice. Nucleic Acids Research, 22:4673-
58 4680.