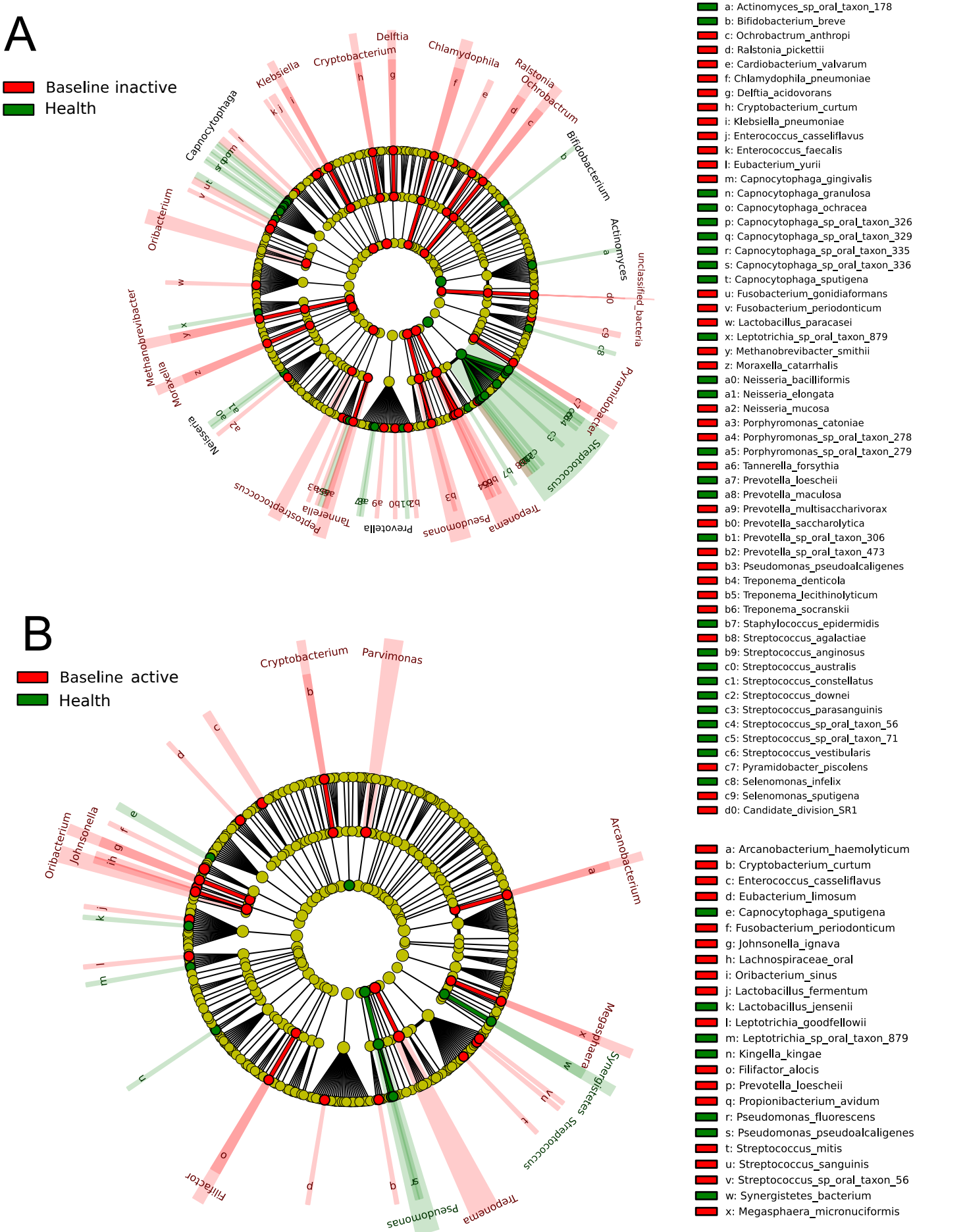
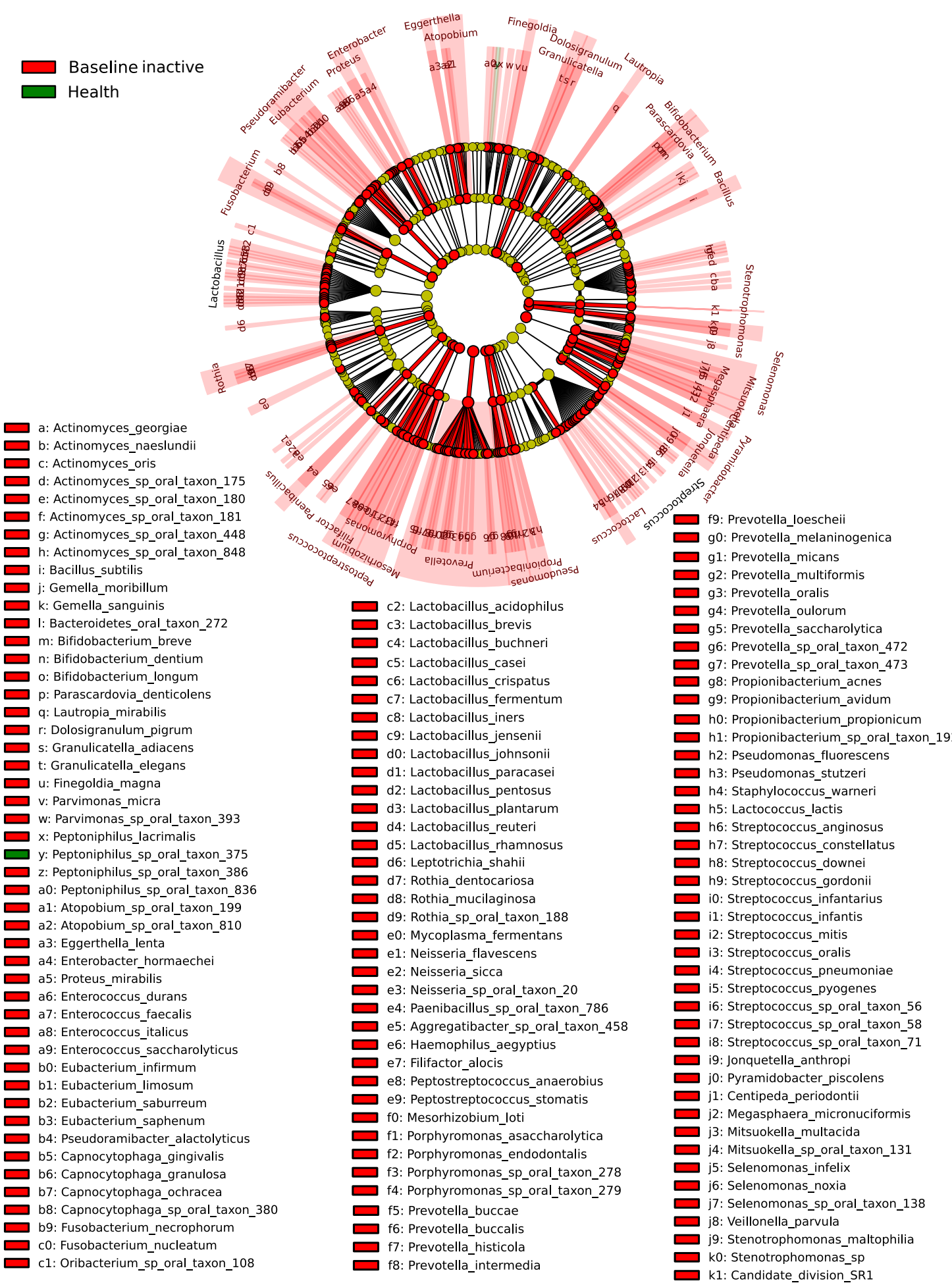


Figure S1. Statistical differences in metagenome composition. Metagenome hit counts were first normalized using GASIC. Normalized counts were then analyzed using LEfSe with default parameters, to identify significant differences at species level between the microbial communities compared.

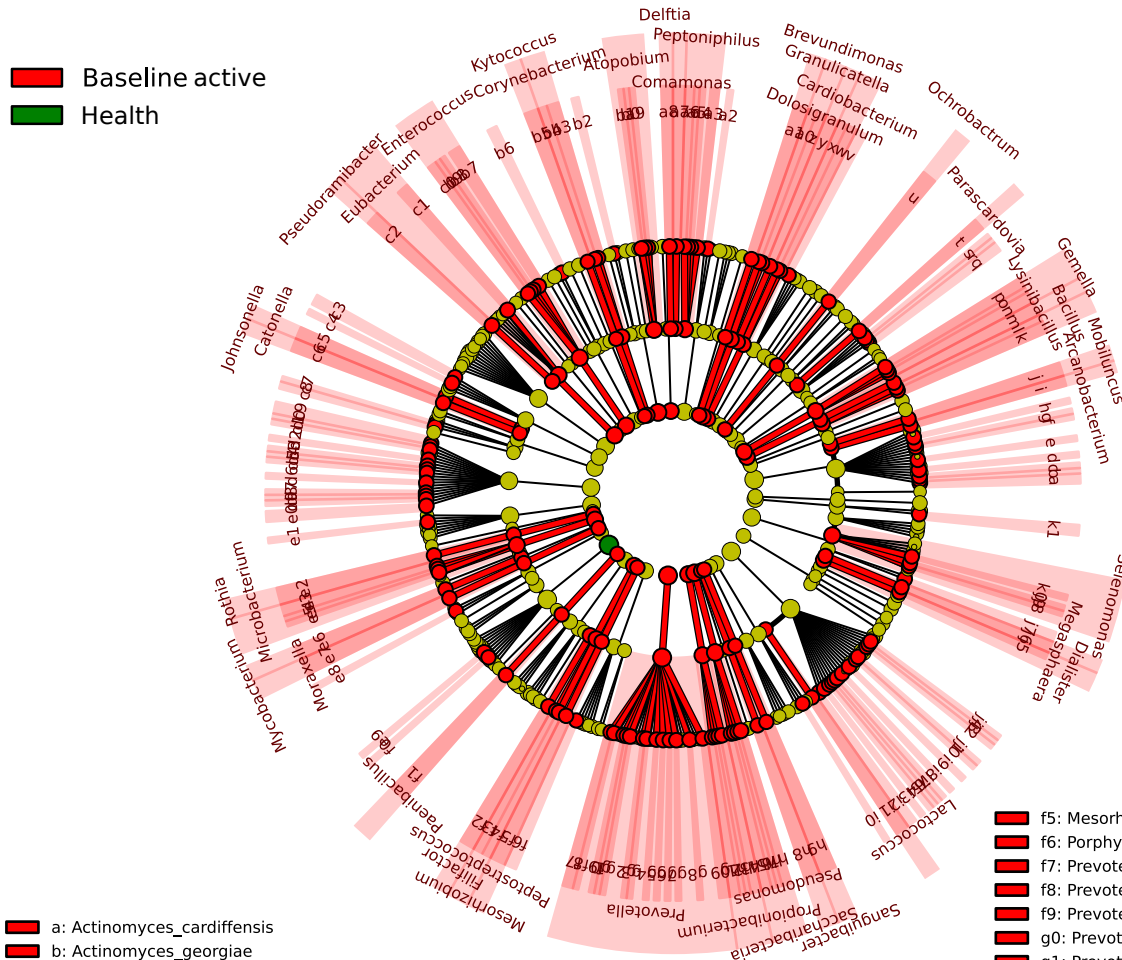
a) Comparison non-progressing site baselines to healthy sites of healthy patients

b) Comparison progressing site baselines to healthy sites of healthy patients



a: Actinomyces_georgiae
b: Actinomyces_naeslundii
c: Actinomyces_oris
d: Actinomyces_sp_oral_taxon_175
e: Actinomyces_sp_oral_taxon_180
f: Actinomyces_sp_oral_taxon_181
g: Actinomyces_sp_oral_taxon_448
h: Actinomyces_sp_oral_taxon_848
i: Bacillus_subtilis
j: Gemella_moribillum
k: Gemella_sanguinis
l: Bacteroidetes_oral_taxon_272
m: Bifidobacterium_breve
n: Bifidobacterium_dentium
o: Bifidobacterium_longum
p: Parascardovia_denticolens
q: Lautropia_mirabilis
r: Dolosigranulum_pigrum
s: Granulicatella_adiacens
t: Granulicatella_elegans
u: Finegoldia_magna
v: Parvimonas_micra
w: Parvimonas_sp_oral_taxon_393
x: Peptoniphilus_lacrimalis
y: Peptoniphilus_sp_oral_taxon_375
z: Peptoniphilus_sp_oral_taxon_386
a0: Peptoniphilus_sp_oral_taxon_836
a1: Atopobium_sp_oral_taxon_199
a2: Atopobium_sp_oral_taxon_810
a3: Eggerthella_lenta
a4: Enterobacter_hormaechei
a5: Proteus_mirabilis
a6: Enterococcus_durans
a7: Enterococcus_faecalis
a8: Enterococcus_italicus
a9: Enterococcus_saccharolyticus
b0: Eubacterium_infirmum
b1: Eubacterium_limosum
b2: Eubacterium_saburreum
b3: Eubacterium_saphenum
b4: Pseudoramibacter_alactolyticus
b5: Capnocytophaga_gingivalis
b6: Capnocytophaga_granulosa
b7: Capnocytophaga_ochracea
b8: Capnocytophaga_sp_oral_taxon_380
b9: Fusobacterium_necrophorum
c0: Fusobacterium_nucleatum
c1: Oribacterium_sp_oral_taxon_108
c2: Lactobacillus_acidophilus
c3: Lactobacillus_brevis
c4: Lactobacillus_buchneri
c5: Lactobacillus_casei
c6: Lactobacillus_crispatus
c7: Lactobacillus_fermentum
c8: Lactobacillus_iners
c9: Lactobacillus_jensenii
d0: Lactobacillus_johnsonii
d1: Lactobacillus_paracasei
d2: Lactobacillus_pentosus
d3: Lactobacillus_plantarum
d4: Lactobacillus_reuteri
d5: Lactobacillus_rhamnosus
d6: Leptotrichia_shahii
d7: Rothia_dentocariosa
d8: Rothia_mucilaginosa
d9: Rothia_sp_oral_taxon_188
e0: Mycoplasma_fermentans
e1: Neisseria_flavescens
e2: Neisseria_sicca
e3: Neisseria_sp_oral_taxon_20
e4: Paenibacillus_sp_oral_taxon_786
e5: Aggregatibacter_sp_oral_taxon_458
e6: Haemophilus_aegyptius
e7: Filifactor_alocis
e8: Peptostreptococcus_anaerobius
e9: Peptostreptococcus_stomatis
f0: Mesorhizobium_loti
f1: Porphyromonas_asaccharolytica
f2: Porphyromonas_endodontalis
f3: Porphyromonas_sp_oral_taxon_278
f4: Porphyromonas_sp_oral_taxon_279
f5: Prevotella_buccae
f6: Prevotella_buccalis
f7: Prevotella_histicola
f8: Prevotella_intermedia
f9: Prevotella_loescheii
g0: Prevotella_melaninogenica
g1: Prevotella_micans
g2: Prevotella_multiformis
g3: Prevotella_oralis
g4: Prevotella_oulorum
g5: Prevotella_saccharolytica
g6: Prevotella_sp_oral_taxon_472
g7: Prevotella_sp_oral_taxon_473
g8: Propionibacterium_acnes
g9: Propionibacterium_avidum
h0: Propionibacterium_propionicum
h1: Propionibacterium_sp_oral_taxon_192
h2: Pseudomonas_fluorescens
h3: Pseudomonas_stutzeri
h4: Staphylococcus_warneri
h5: Lactococcus_lactis
h6: Streptococcus_anginosus
h7: Streptococcus_constellatus
h8: Streptococcus_downei
h9: Streptococcus_gordonii
i0: Streptococcus_infantarius
i1: Streptococcus_infantis
i2: Streptococcus_mitis
i3: Streptococcus_oralis
i4: Streptococcus_pneumoniae
i5: Streptococcus_pyogenes
i6: Streptococcus_sp_oral_taxon_56
i7: Streptococcus_sp_oral_taxon_58
i8: Streptococcus_sp_oral_taxon_71
i9: Jonquetella_anthropi
j0: Pyramidobacter_piscolens
j1: Centipeda_periodontii
j2: Megasphaera_micronuciformis
j3: Mitsuokella_multacida
j4: Mitsuokella_sp_oral_taxon_131
j5: Selenomonas_infelix
j6: Selenomonas_noxia
j7: Selenomonas_sp_oral_taxon_138
j8: Veillonella_parvula
j9: Stenotrophomonas_maltophilia
k0: Stenotrophomonas_sp
k1: Candidate_division_SR1

Figure S2. Statistical differences in normalized metatranscriptome composition comparing non-progressing site baselines to healthy sites of healthy patients. Metagenome hit counts were first normalized using GASIC. Metatranscriptome normalized counts were then analyzed using LEfSe with default parameters, to identify significant differences in activity at species level between the microbial communities compared.



a: Actinomyces_cardiffensis
b: Actinomyces_georgiae
c: Actinomyces_graevenitzii
d: Actinomyces_odontolyticus
e: Actinomyces_sp_oral_taxon_171
f: Actinomyces_sp_oral_taxon_180
g: Actinomyces_sp_oral_taxon_181
h: Actinomyces_sp_oral_taxon_848
i: Arcanobacterium_haemolyticum
j: Mobiluncus_mulieris
k: Bacillus_clausii
l: Bacillus_subtilis
m: Lysinibacillus_fusiformis
n: Gemella_haemolysans
o: Gemella_moribillum
p: Gemella_sanguinis
q: Bifidobacterium_animalis
r: Bifidobacterium_breve
s: Bifidobacterium_dentium
t: Parascardovia_denticolens
u: Ochrobactrum_anthropi
v: Campylobacter_showae
w: Cardiobacterium_hominis
x: Cardiobacterium_valvarum
y: Dolosigranulum_pigrum
z: Granulicatella_adiacens
a0: Granulicatella_elegans
a1: Brevundimonas_diminuta
a2: Parvimonas_sp_oral_taxon_110
a3: Peptoniphilus_indolicus
a4: Peptoniphilus_sp_oral_taxon_375
a5: Peptoniphilus_sp_oral_taxon_386
a6: Peptoniphilus_sp_oral_taxon_836
a7: Comamonas_testosteroni
a8: Delftia_acidovorans
a9: Atopobium_parvulum
b0: Atopobium_rimae
b1: Atopobium_sp_oral_taxon_199
b2: Olsenella_uli
b3: Corynebacterium_matruchotii
b4: Corynebacterium_urealyticum
b5: Kytococcus_sedentarius
b6: Enterobacter_cancerogenus
b7: Enterococcus_casseliflavus
b8: Enterococcus_faecalis
b9: Enterococcus_italicus
c0: Enterococcus_saccharolyticus
c1: Eubacterium_infirmum
c2: Pseudoramibacter_alactolyticus
c3: Fusobacterium_gonidiaformans
c4: Fusobacterium_nucleatum
c5: Catonella_morbi
c6: Johnsonella_ignava
c7: Oribacterium_sp_oral_taxon_108
c8: Oribacterium_sp_oral_taxon_78
c9: Lactobacillus_acidophilus
d0: Lactobacillus_brevis
d1: Lactobacillus_buchneri
d2: Lactobacillus_crispatus
d3: Lactobacillus_gasseri
d4: Lactobacillus_iners
d5: Lactobacillus_jensenii
d6: Lactobacillus_oris
d7: Lactobacillus_pentosus
d8: Lactobacillus_plantarum
d9: Lactobacillus_reuteri
e0: Lactobacillus_salivarius
e1: Leptotrichia_hofstadii
e2: Microbacterium_sp_oral_taxon_186
e3: Rothia_aeria
e4: Rothia_dentocariosa
e5: Rothia_mucilaginosa
e6: Moraxella_catarrhalis
e7: Mycobacterium_tuberculosis
e8: Mycoplasma_hominis
e9: Neisseria_polysaccharea
f0: Neisseria_sp_oral_taxon_14
f1: Paenibacillus_sp_oral_taxon_786
f2: Filifactor_alocis
f3: Peptostreptococcus_anaerobius
f4: Peptostreptococcus_stomatis
f5: Mesorhizobium_loti
f6: Porphyromonas_asaccharolytica
f7: Prevotella_bivia
f8: Prevotella_buccae
f9: Prevotella_denticola
g0: Prevotella_histicola
g1: Prevotella_intermedia
g2: Prevotella_melaninogenica
g3: Prevotella_micans
g4: Prevotella_nigrescens
g5: Prevotella_oris
g6: Prevotella_pallens
g7: Prevotella_salivae
g8: Prevotella_sp_oral_taxon_317
g9: Prevotella_veroralis
h0: Propionibacterium_acnes
h1: Propionibacterium_avidum
h2: Propionibacterium_propionicum
h3: Propionibacterium_sp_oral_taxon_192
h4: Pseudomonas_aeruginosa
h5: Pseudomonas_fluorescens
h6: Pseudomonas_pseudoalcaligenes
h7: Pseudomonas_stutzeri
h8: Candidate_division_TM7
h9: Sanguibacter_keddieii
i0: Staphylococcus_epidermidis
i1: Lactococcus_lactis
i2: Streptococcus_agalactiae
i3: Streptococcus_australis
i4: Streptococcus_cristatus
i5: Streptococcus_downei
i6: Streptococcus_gordonii
i7: Streptococcus_infantarius
i8: Streptococcus_intermedius
i9: Streptococcus_oligofermentans
j0: Streptococcus_peroris
j1: Streptococcus_pneumoniae
j2: Streptococcus_sobrinus
j3: Streptococcus_sp_oral_taxon_56
j4: Streptococcus_sp_oral_taxon_58
j5: Dialister_microaerophilus
j6: Megasphaera_micronuciformis
j7: Mitsuokella_multacida
j8: Selenomonas_flueggei
j9: Selenomonas_infelix
k0: Selenomonas_sp_oral_taxon_137
k1: Stenotrophomonas_maltophilia

Figure S3. Statistical differences in normalized metatranscriptome composition comparing progressing site baselines to healthy sites of healthy patients. Metagenome hit counts were first normalized using GASIC. Metatranscriptome normalized counts were then analyzed using LEfSe with default parameters, to identify significant differences in activity at species level between the microbial communities compared.

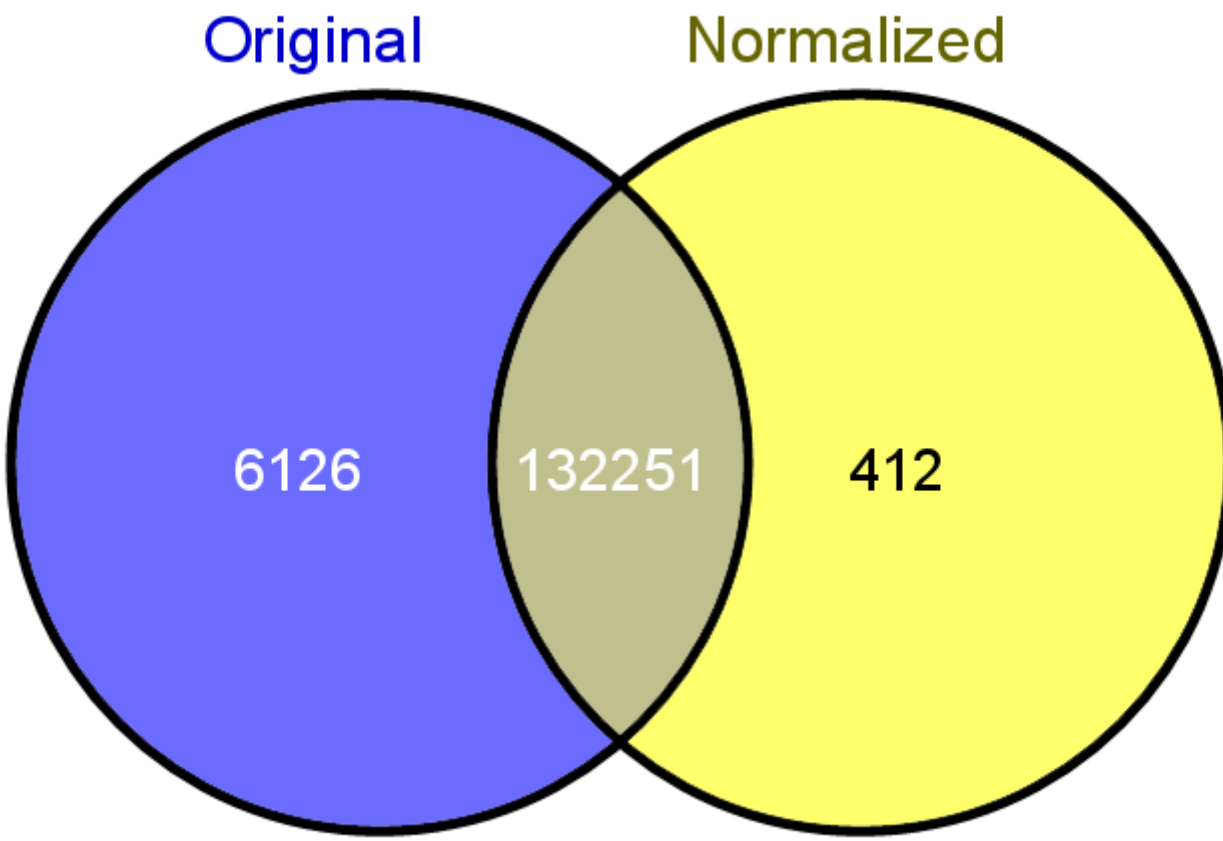


Figure S4. Venn Diagram showing overlaping differentiallly expressed genes comparing baseline and progression with and without normalization. We normalized the expression hit counts against species frequencies estimated using GASIC. Original, differentially expressed genes using raw hits of the metatranscriptome. Venn diagram was obtained using the Venny webpage tool (http://bioinfogp.cnb.csic.es/tools/venny/).

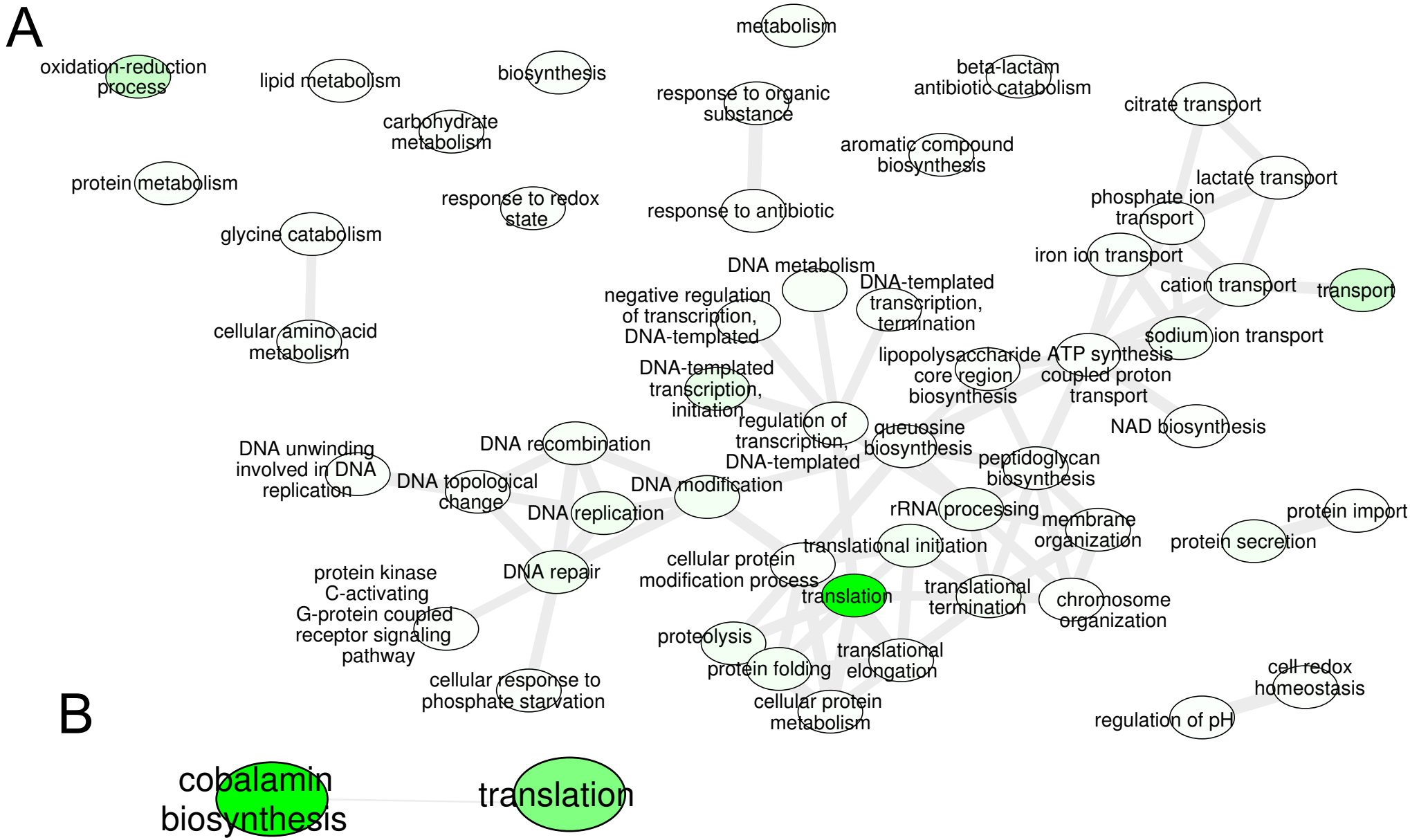


Figure S5. GO terms associated with changes in gene expression profiles in major periodontal pathogens members of the red complex during periodontal disease progression. GO terms were assigned to differentially expressed genes in progression and summaryzed using REVIGO. a) GO terms associated with up-regulated genes in active sites b) GO terms associated with down-regulated genes in active sites.

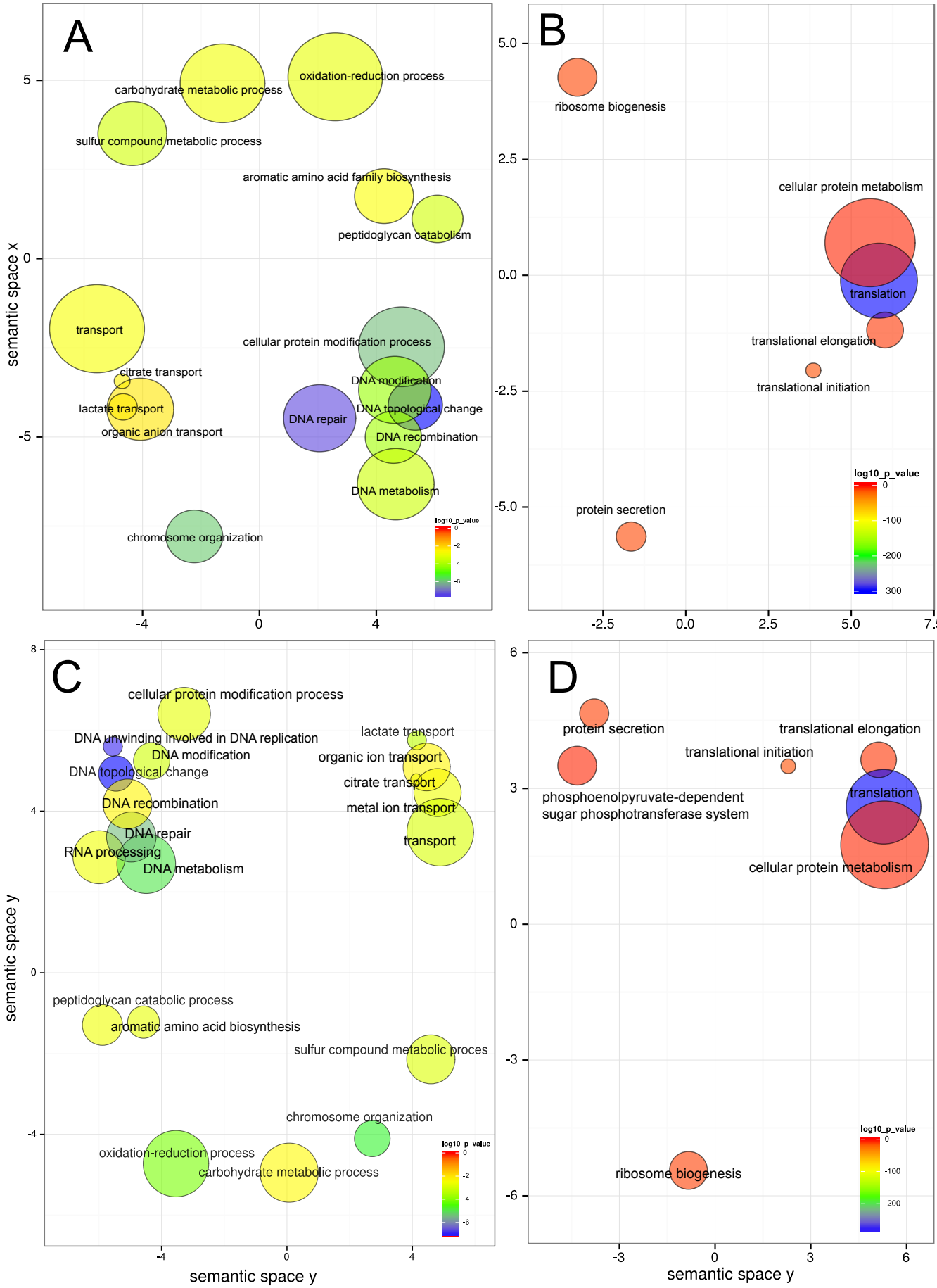


Figure S6. GO enrichment analysis comparing healthy sites from healthy individuals and baselines in progressing and non-progressing sites.. Enriched terms obtained using 'GOseq' were summarized and visualized as a scatter plot using REVIGO. a) Summarized GO terms related to biological processes in inactive baselines.
b) Summarized GO terms related to biological processes in health when compared with inactive baselines.
c) Summarized GO terms related to biological processes in active baselines. d) Summarized GO terms related to biological processes in health when compared with active baselines. Circle size is proportional to the frequency of the GO term, while color indicates the log10 p value (red higher, blue lower). q-value=0.9.

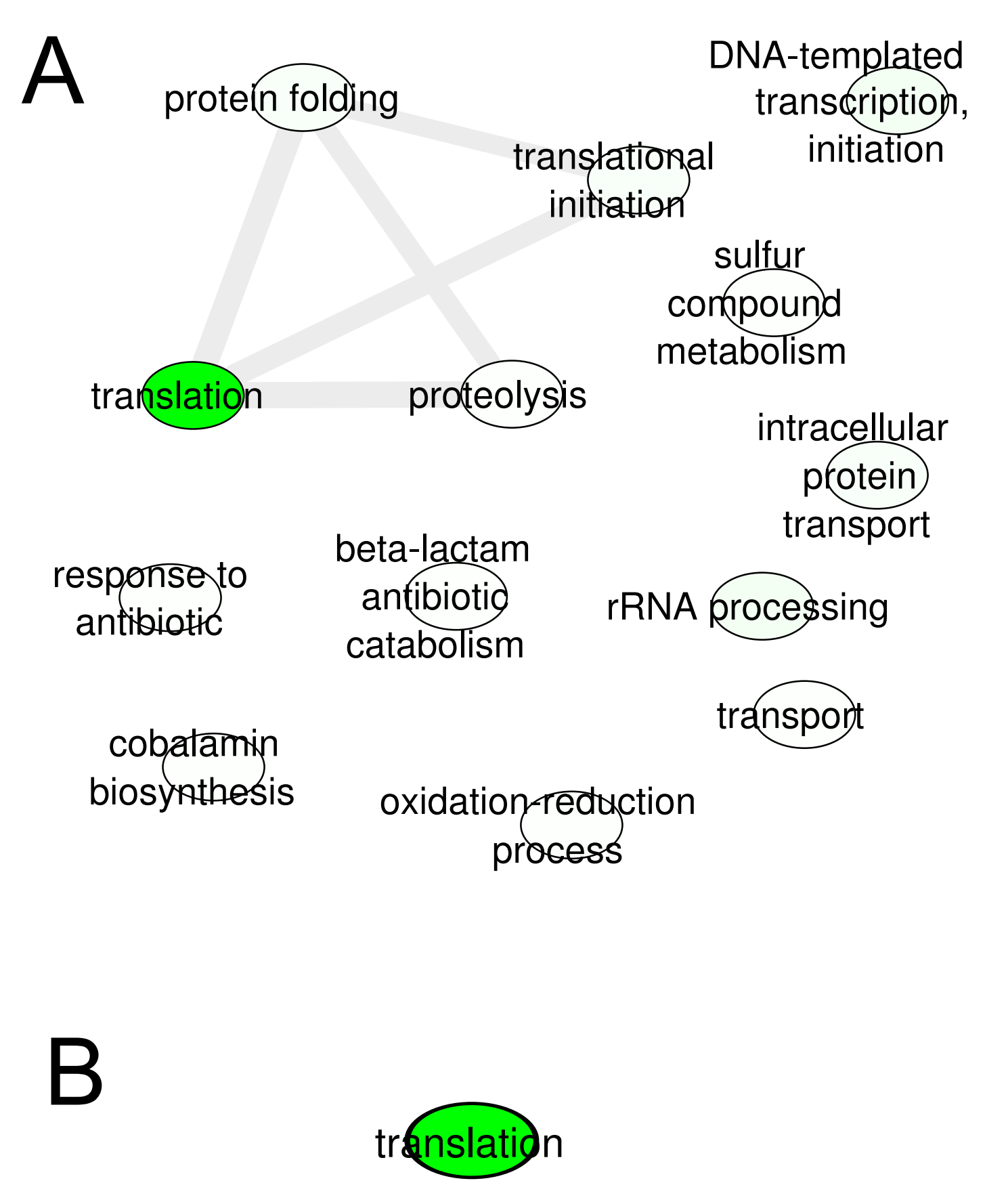


Figure S7. GO terms associated with changes in gene expression profiles in major periodontal pathogens members of the red complex when comparing baselines of active and inactive sites. GO terms were assigned to differentially expressed genes in progression and summaryzed using REVIGO. a) GO terms associated with up-regulated genes in active sites baselines b) GO terms associated with down-regulated genes in active sites baselines.

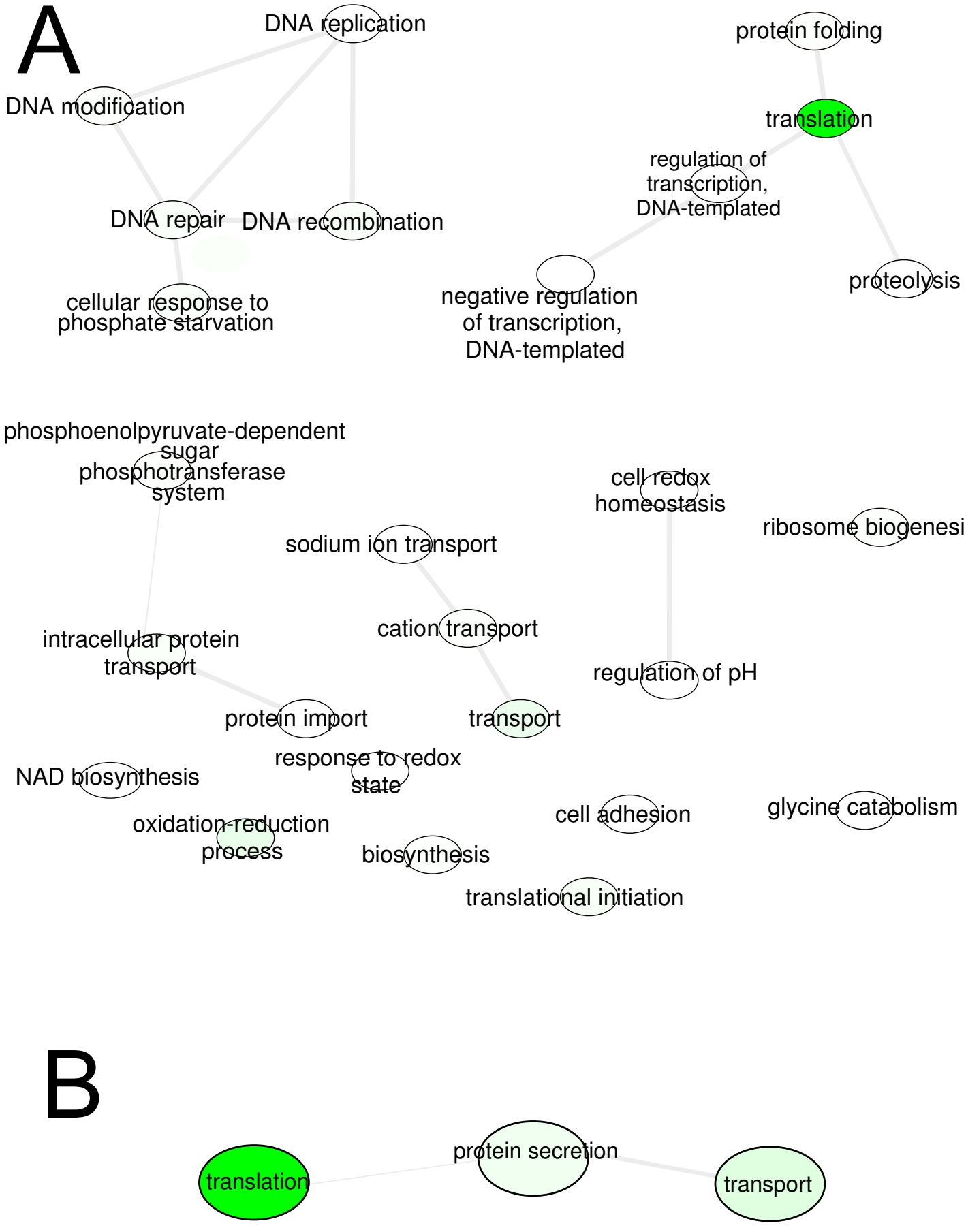


Figure S8. GO terms associated with changes in gene expression profiles in members of the orange complex when comparing baselines of active and inactive sites. GO terms were assigned to differentially expressed genes in progression and summaryzed using REVIGO. a) GO terms associated with up-regulated genes in active sites baselines b) GO terms associated with down-regulated genes in active sites baselines.

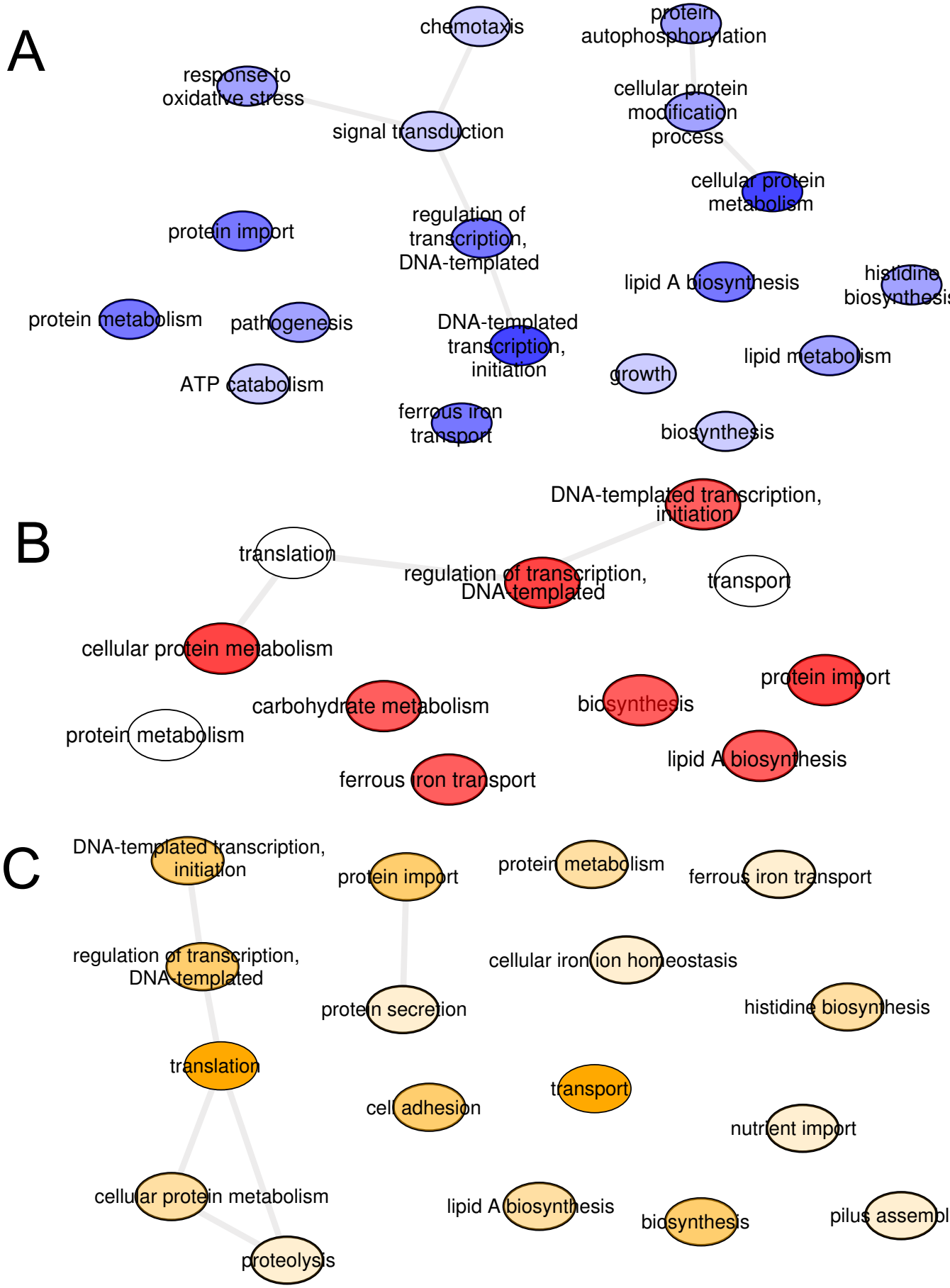


Figure S9. GO terms associated with changes in gene expression of putative virulence factors in the oral community during periodontitis progression. GO terms were assigned to differentially expressed putative virulence factors in progression and summaryzed using REVIGO.
a) GO terms enrichment analysis of virulence factors in the whole community b) GO terms associated with up-regulated virulence factors in the red complex c) GO terms associated with up-regulated virulence factors in the orange complex

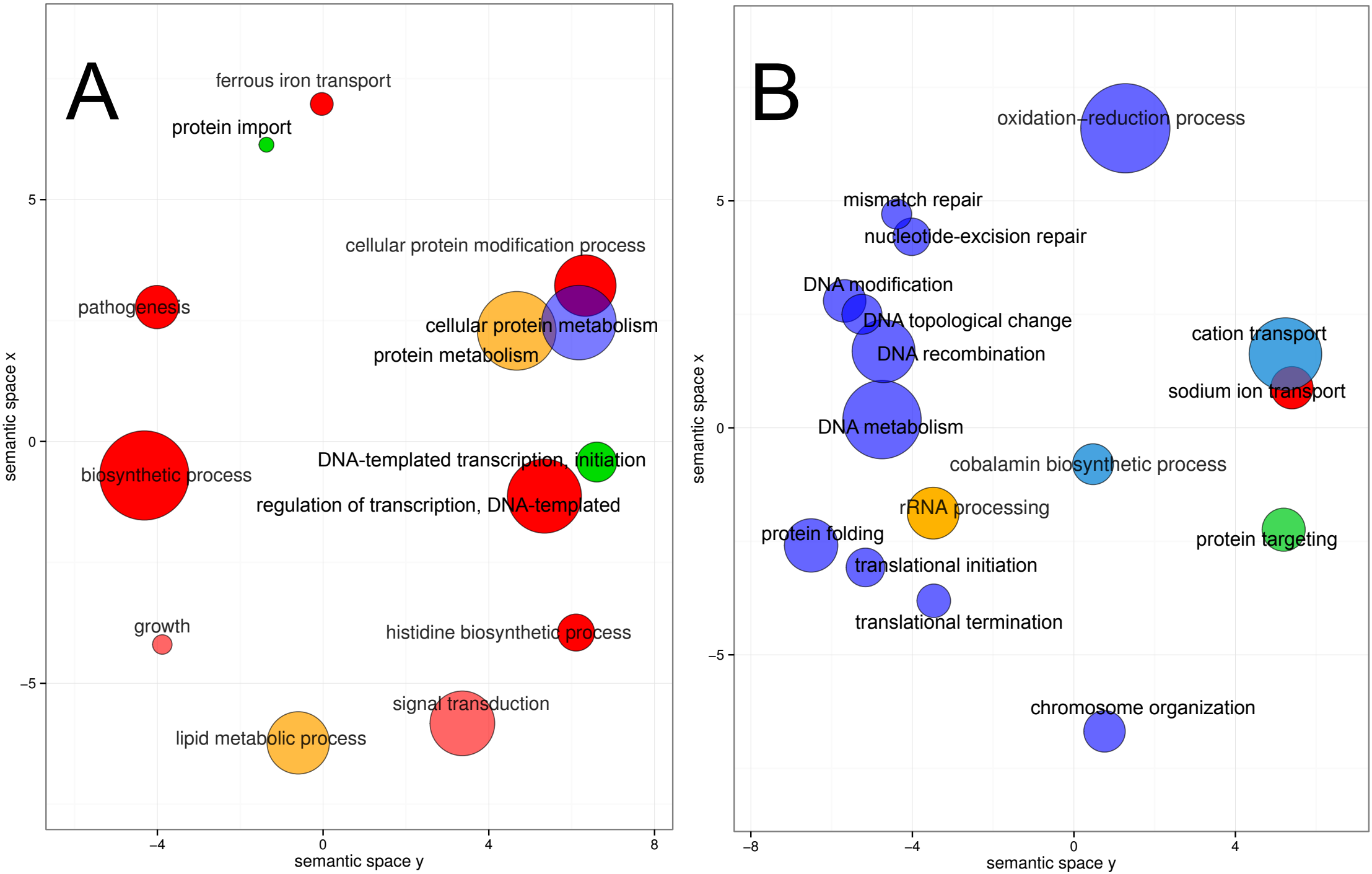


Figure S10. GO terms enrichment analysis of virulence factors comparing baselines. GO terms enrichment was performed using GOseq and summaryzed using REVIGO.
A) GO terms over-represented in progressing sites baselines B) GO terms over-represented in non-progressing sites baselines.

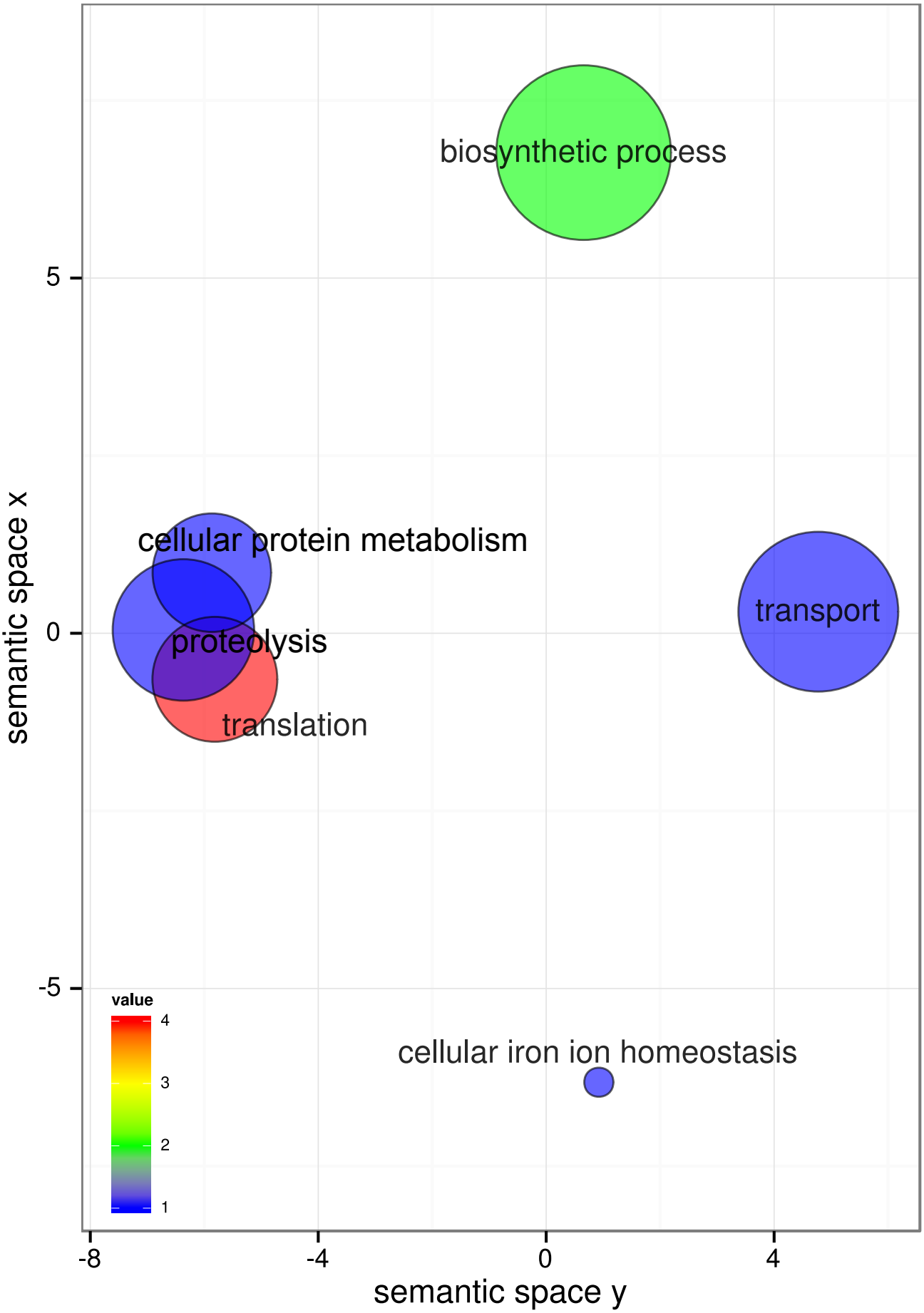


Figure S11. GO terms enrichment analysis of virulence factors in the orange complex comparing baselines. GO terms enrichment was performed using GOseq and summaryzed using REVIGO. GO terms over-represented in baselines of progressing sites.

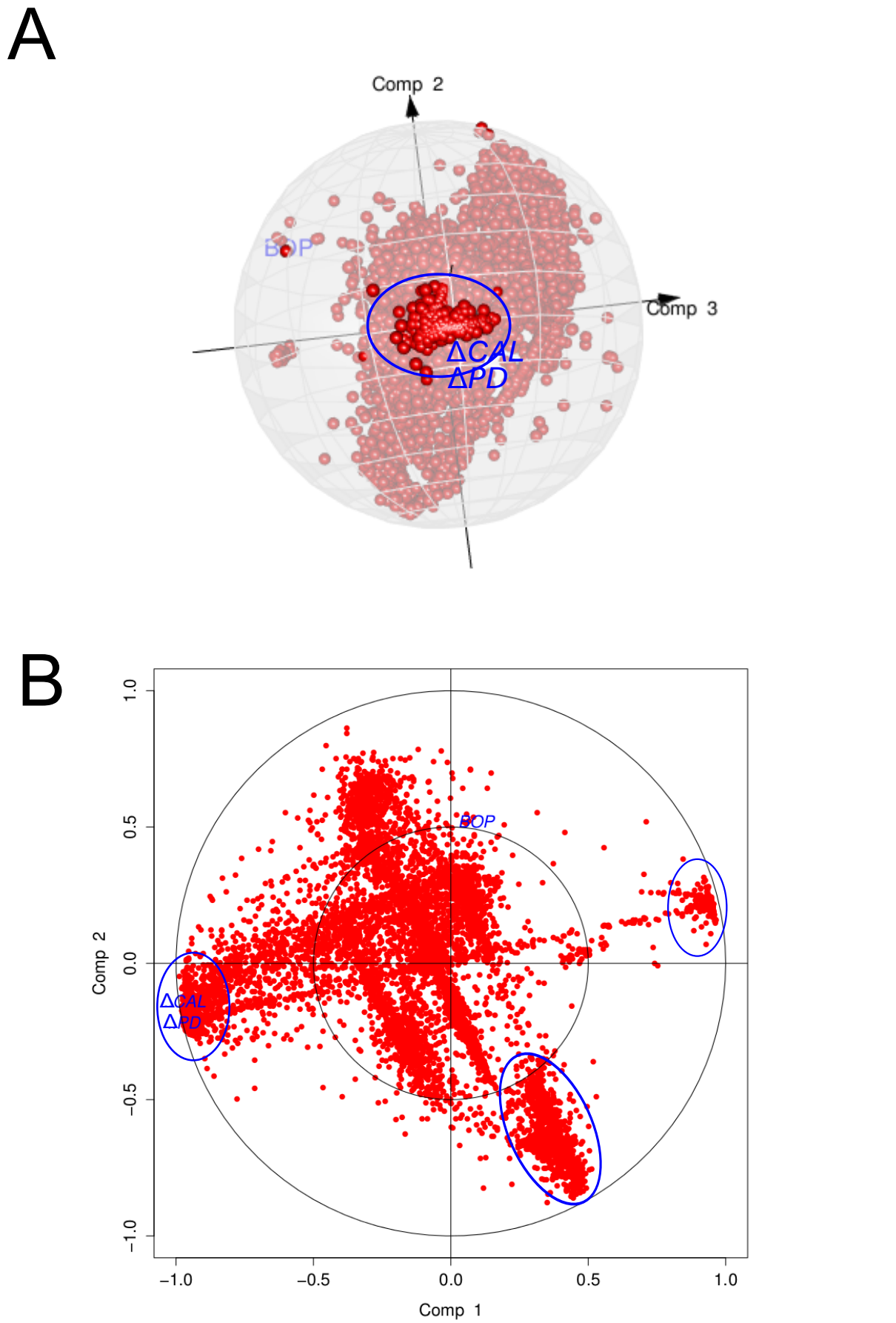


Figure S12. Correlation Circle plots of sPLS analysis. Correlation Circle plots were obtained to assess correlation of the evolution of bleeding on probing (BOP), increase in pocket depth (ΔPD) and increase in clinical attachment level (ΔCAL). A) 3D representation of gene expression associations with evolution of clinical traits (components 1 to 3). B) Gene expression associations with evolution of clinical traits of the 2 first components.

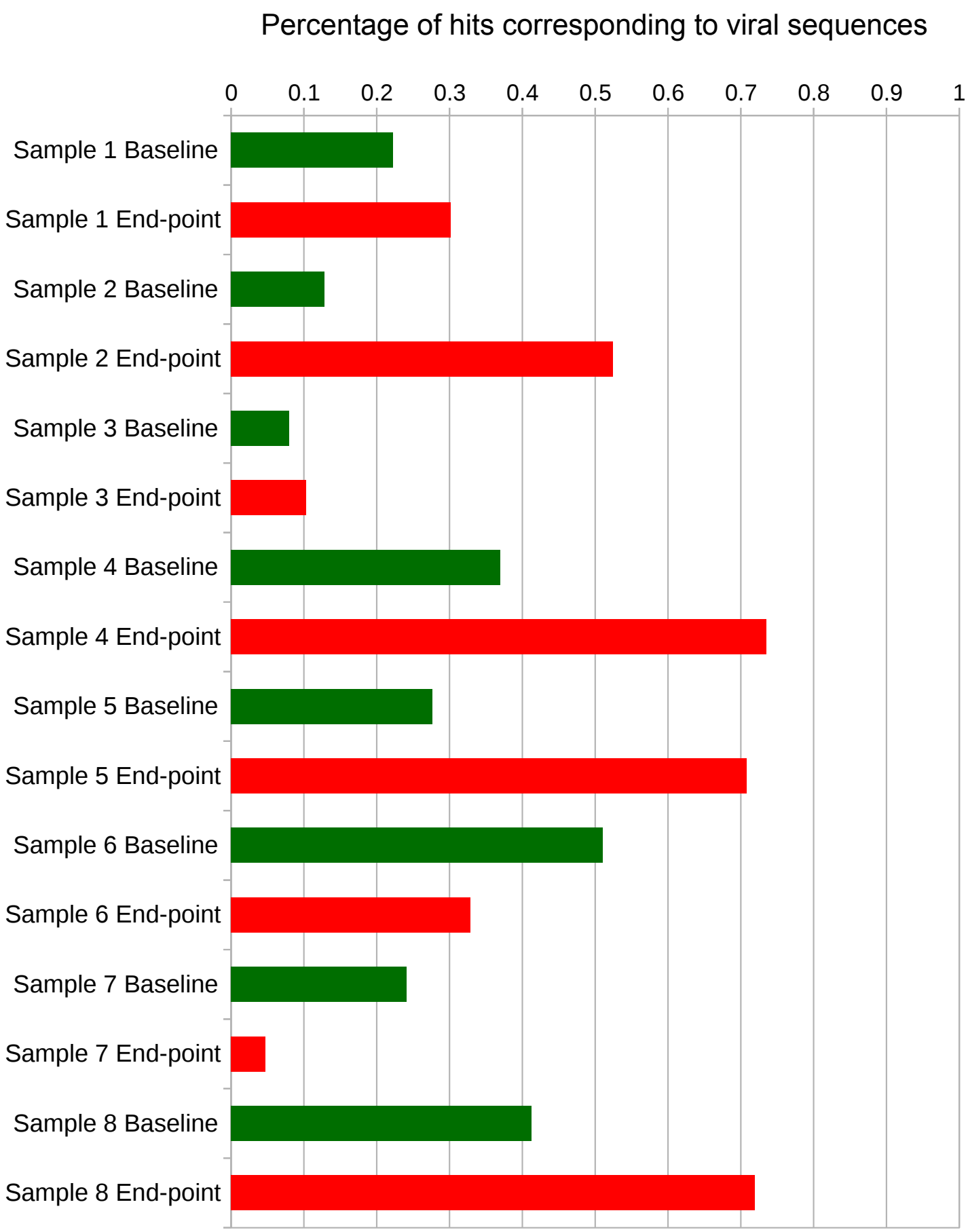
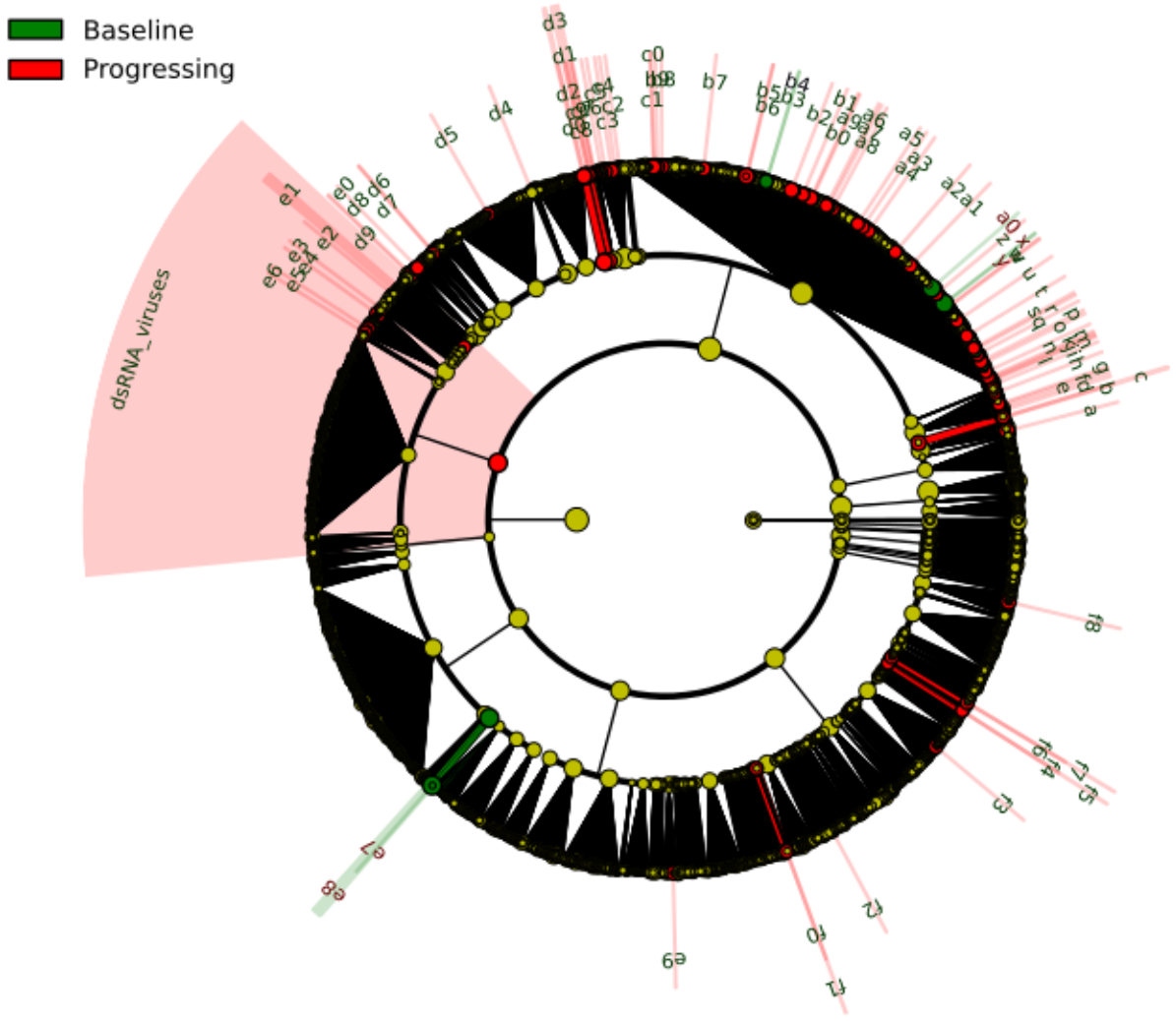


Figure S13. Percentage of hits corresponding to viral sequences. Sequences were aligned against a database containing all viral sequences in NCBI. Bars represent the percentage of hits that corresponded to viral sequences.



- a: Fowl_adenovirus_C
- b: Heliothis_virescens_ascovirus_3e
- c: Ascoviridae
- d: Bombyx_mori_NPV
- e: Chrysodeixis_chalcites_nucleopolyhedrovirus
- f: Epinotia_aporema_granulovirus
- g: Erwinia_phage_ENT90
- h: Rachiplusia_ou_MNPV
- i: Acinetobacter_phage_Acj9
- j: Aeromonas_phage_31
- k: Aeromonas_phage_Aes012
- l: Aeromonas_phage_phiAS4
- m: Agrobacterium_phage_7_7_1
- n: Bacillus_phage_BPS13
- o: Bacillus_phage_Megatron
- p: Bacillus_phage_SPBc2
- q: Bacteriophage_RB32
- r: Burkholderia_phage_AH2
- s: Burkholderia_phage_BcepB1A
- t: Cellulophaga_phage_phi14_2
- u: Cronobacter_phage_ENT39118
- v: Enterobacteria_phage_CC31
- w: Enterobacteria_phage_EPS7
- z: Enterobacteria_phage_RTP
- x: Enterobacteria_phage_HK629
- y: Enterobacteria_phage_HK630
- a0: Enterobacteria_phage_lambda
- b4: Pseudomonas_phage_MP38
- e7: Enterobacteria_phage_phiX174_sensu_lato
- e8: Microviridae
- a1: Haemophilus_phage_HP2
- a2: Lactococcus_phage_P092
- a3: Mycobacterium_phage_Artemis2UCLA
- a4: Mycobacterium_phage_Bruin
- a5: Mycobacterium_phage_Che8
- a6: Mycobacterium_phage_MichelleMyBell
- a7: Mycobacterium_phage_Obama12
- a8: Mycobacterium_phage_PLot
- a9: Mycobacterium_phage_Spud
- b0: Mycobacterium_phage_Thibault
- b1: Myobacteriophage_Papyrus
- b2: Prochlorococcus_phage_P_SSM2
- b3: Pseudomonas_phage_MP29
- b5: Ralstonia_phage_RSB3_DNA
- b6: Ralstonia_phage_RSL1
- b7: Staphylococcus_phage_EW
- b8: Synechococcus_phage_S_IOM18
- b9: Synechococcus_phage_S_SSM5
- c0: Thermus_phage_P23_45
- c1: Thermus_phage_TMA
- e9: Mouse_astrovirus_M_52_USA_2008
- f0: Fathead_minnow_picornavirus_strain_09_283
- f1: Fathead_minnow_picornavirus
- f2: Bat_hepevirus
- f3: Papaya_ringspot_virus
- f4: Seneca_valley_virus
- f5: Seneca_valley_virus
- f6: Solenopsis_invicta_virus_3
- f7: Solenopsis_invicta_virus_1
- f8: Pepper_mild_mottle_virus
- c2: Callitrichine_herpesvirus_3_strain_CJ0149
- c3: Equid_herpesvirus_1
- c4: Gallid_herpesvirus_2
- c5: Meleagrid_herpesvirus_1
- c6: Ranid_herpesvirus_1_strain_McKinnell
- c7: Invertebrate_iridovirus_25_complete
- c8: Lymphocystis_disease_virus___isolate_China
- c9: Acidianus_filamentous_virus_9
- d0: Sulfolobus_islandicus_rod_shaped_virus_1
- d1: Ligamenvirales
- d2: Helicoverpa_zea_nudivirus_2
- d3: Nudiviridae
- d4: Phaeocystis_globosa_virus_strain_16T
- d5: Glypta_fumiferanae_ichnovirus_segment_C10
- d6: Pigeonpox_virus_isolate_FeP2
- d7: Rabbit_fibroma_virus
- d8: Vibrio_phage_henriette_12B8
- d9: Halovirus_PH1
- e0: Micromonas_pusilla_virus_12T
- e1: Chrysoviridae
- e2: Pseudomonas_phage_phi15
- e3: Sclerotinia_sclerotiorum_hypovirus_2_isolate_5472
- e4: Dill_cryptic_virus_1_isolate_IPP_hortorum_segment_1
- e5: Rhizoctonia_solani_dsRNA_virus_2_segment_2
- e6: Vicia_cryptic_virus_RNA2

Figure S14. Statistical differences in viral composition of transcript libraries. Hit counts were analyzed using LEfSe with default parameters, to identify significant differences at species level between the microbial communities compared. Comparison progressing site baselines to end point of the same sites.