

## Additional file 1

Description for all matrixes: The part of the matrix from the upper right corner shows differences in the number of deletions between sequence pairs; the part of the matrix from the lower left corner shows differences in the number of nucleotides between the sequence pairs. The green color indicates species pairs that differ in the lowest number of nucleotides or deletions, the red color the pairs with the highest number of nucleotide or deletion differences between sequences.



Table S2

A matrix showing the nucleotides and deletions differences between pairs of 16S rRNA gene sequences for *Enterococcus* species.

	1	2	3	4	5	6	7	8	9
E._avium	1	0	0	3	3	3	3	3	0
E._casseliflavus	2	22	0	3	3	3	3	3	0
E._cecorum	3	59	53	3	3	3	3	3	0
E._durans	4	22	34	71	0	0	0	0	3
E._faecium	5	24	35	74	7	0	0	0	3
E._faecalis	6	66	62	90	59	55	0	0	3
E._hirae	7	20	34	71	5	7	56	0	3
E._porcinus	8	27	35	75	16	14	63	13	3
E._raffinosis	9	5	20	59	23	25	67	21	26



Table S4

A matrix showing the nucleotides and deletions differences between pairs of *sodA* gene sequences for *Enterococcus* species.

	1	2	3	4	5	6	7	8	9
E._avium	1	0	0	0	0	0	0	0	0
E._casseliflavus	2	96	0	0	0	0	0	0	0
E._cecorum	3	119	114	0	0	0	0	0	0
E._durans	4	112	98	121	0	0	0	0	0
E._faecium	5	99	77	106	95	0	0	0	0
E._faecalis	6	97	75	106	94	5	0	0	0
E._hiraе	7	96	80	105	62	78	80	0	0
E._porcinus	8	118	105	113	70	99	100	60	0
E._raffinosis	9	52	84	124	100	94	93	90	114



Table S6

A matrix showing the nucleotides and deletions differences between pairs of *tuf* gene sequences for *Enterococcus* species.

	1	2	3	4	5	6	7	8	9
E._avium	1	0	0	0	0	0	0	0	0
E._casseliflavus	2	70	0	0	0	0	0	0	0
E._cecorum	3	71	76	0	0	0	0	0	0
E._durans	4	80	75	72	0	0	0	0	0
E._faecium	5	86	71	75	28	0	0	0	0
E._faecalis	6	69	73	75	89	93	0	0	0
E._hirae	7	80	73	70	10	26	87	0	0
E._porcinus	8	107	99	99	57	72	114	57	0
E._raffinosis	9	19	73	72	81	83	74	82	104





Table S8

A matrix showing the nucleotides and deletions differences between pairs of *rpoB* gene sequences for *Enterococcus* species.

	1	2	3	4	5	6	7	8	9
E._avium	1	0	0	0	0	0	0	0	0
E._casseliflavus	2	118	0	0	0	0	0	0	0
E._cecorum	3	113	110	0	0	0	0	0	0
E._durans	4	108	98	98	0	0	0	0	0
E._faecium	5	114	99	101	62	0	0	0	0
E._faecalis	6	103	94	94	99	92	0	0	0
E._hirae	7	116	115	94	69	71	104	0	0
E._porcinus	8	114	113	94	83	77	83	71	0
E._raffinosis	9	64	102	113	103	118	92	117	109



Table S10

A matrix showing the nucleotides and deletions differences between pairs of 16S-23S rRNA region sequences for *Enterococcus* species.

		1	2	3	4	5	6	7	8	9
E._avium	1		55	77	148	160	47	149	135	29
E._casseliflavus	2	238		62	115	125	28	116	102	32
E._cecorum	3	333	294		159	167	58	156	142	58
E._durans	4	332	280	416		70	115	55	35	125
E._faecium	5	336	279	416	204		123	17	37	137
E._faecalis	6	289	261	354	333	307		110	96	24
E._hirae	7	349	295	431	104	180	348		22	126
E._porcinus	8	341	271	418	114	193	336	103		112
E._raffinosis	9	62	199	307	294	295	252	309	302	

Table S11

The length of 16S-23S rRNA region, 16S rRNA gene, intergenic spacer region and 23S rRNA region for all *Streptococcus* and *Enterococcus* species.

Species	16S-23S rRNA region length (bp)	16S rRNA gene length (bp)	Intergenic spacer region length (bp)	23S rRNA gene length (bp)
<i>Streptococcus acidominimus</i>	4486	1523	468	2495
<i>Streptococcus adjacens</i>	4251	1525	218	2508
<i>Streptococcus anginosus</i>	4411	1531	387	2493
<i>Streptococcus australis</i>	4254	1521	240	2493
<i>Streptococcus canis</i>	4301	1524	283	2494
<i>Streptococcus constellatus</i>	4340	1531	315	2494
<i>Streptococcus cremoris</i>	4317	1522	300	2495
<i>Streptococcus criceti</i>	4649	1532	625	2492
<i>Streptococcus cristatus</i>	4260	1522	248	2490
<i>Streptococcus difficilis</i>	4295	1524	277	2494
<i>Streptococcus downei</i>	4564	1533	538	2493
<i>Streptococcus durans</i>	4366	1535	326	2505
<i>Streptococcus dysgalactiae</i>	4301	1524	283	2494
<i>Streptococcus equi</i>	4429	1524	414	2491
<i>Streptococcus equinus</i>	4732	1531	708	2493
<i>Streptococcus gallolyticus</i>	4284	1524	270	2490
<i>Streptococcus gordonii</i>	4267	1532	243	2492
<i>Streptococcus infantarius</i>	4284	1524	269	2491
<i>Streptococcus infantis</i>	4258	1523	242	2493
<i>Streptococcus intermedius</i>	4291	1531	267	2493
<i>Streptococcus mitis</i>	4258	1521	244	2493
<i>Streptococcus mutans</i>	4408	1532	384	2492
<i>Streptococcus oligofermentans</i>	4263	1524	249	2490
<i>Streptococcus oralis</i>	4256	1522	242	2492
<i>Streptococcus ovis</i>	4295	1533	265	2497
<i>Streptococcus parasanguinis</i>	4261	1523	245	2493
<i>Streptococcus pasteurianus</i>	4284	1524	268	2492
<i>Streptococcus pluranimalium</i>	4304	1534	276	2494
<i>Streptococcus pneumoniae</i>	4258	1521	244	2493
<i>Streptococcus porcinus</i>	4558	1524	538	2496
<i>Streptococcus pseudopneumoniae</i>	4260	1521	244	2495
<i>Streptococcus pseudoporcinus</i>	4498	1524	478	2496
<i>Streptococcus pyogenes</i>	4436	1524	418	2494
<i>Streptococcus saccharolyticus</i>	4263	1533	224	2506
<i>Streptococcus salivarius</i>	4283	1523	269	2491
<i>Streptococcus sanguinis</i>	4275	1523	260	2492
<i>Streptococcus sinensis</i>	4270	1532	248	2490
<i>Streptococcus sobrinus</i>	4424	1532	403	2489

Species	16S-23S rRNA region length (bp)	16S rRNA gene length (bp)	Intergenic spacer region length (bp)	23S rRNA gene length (bp)
<i>Streptococcus suis</i>	4420	1524	401	2495
<i>Streptococcus tigurinus</i>	4259	1524	242	2493
<i>Streptococcus uberis</i>	4352	1524	336	2492
<i>Streptococcus urinalis</i>	4292	1524	273	2495
<i>Enterococcus avium</i>	4273	1540	231	2502
<i>Enterococcus casseliflavus</i>	4266	1541	220	2505
<i>Enterococcus cecorum</i>	4224	1540	182	2502
<i>Enterococcus durans</i>	4313	1541	268	2504
<i>Enterococcus faecium</i>	4381	1541	335	2505
<i>Enterococcus faecalis</i>	4262	1542	217	2503
<i>Enterococcus hirae</i>	4366	1541	321	2504
<i>Enterococcus porcinus</i>	4346	1541	301	2504
<i>Enterococcus raffinosus</i>	4258	1540	215	2503

Table S12

The intraspecies polymorphism of 16S-23S rRNA region sequence within *Streptococcus* and *Enterococcus* genera<sup>a</sup>.

Species	The lowest intraspecies nucleotide difference compared to GenBank sequences	The highest intraspecies nucleotide difference compared to GenBank sequences	Amount of complete genome sequences/genome assembly/ 16S-23S rRNA region (GenBank)
<i>Streptococcus acidominimus</i>	-	-	0/1
<i>Streptococcus anginosus</i>	0 (0%)	49 (1.11%)	6/1
<i>Streptococcus australis</i>	111 (2.6%)	111 (2.6%)	0/1
<i>Streptococcus canis</i>	3 (0.07%)	3 (0.07%)	0/1
<i>Streptococcus constellatus</i>	0 (0%)	18 (0.41%)	3/0
<i>Streptococcus cremoris</i>	0 (0%)	18 (0.42%)	14/0
<i>Streptococcus criceti</i>	0 (0%)	1 (0.02%)	0/0/3
<i>Streptococcus durans</i>	0 (0%)	73 (2.74%)	3/0
<i>Streptococcus dysgalactiae</i>	0 (0%)	39 (0.91%)	6/7
<i>Streptococcus equi</i>	0 (0%)	67 (1.51%)	6/9
<i>Streptococcus equinus</i>	-	-	0/4
<i>Streptococcus gallolyticus</i>	0 (0%)	7 (0.16%)	5/1
<i>Streptococcus gordonii</i>	0 (0%)	21 (0.49%)	5/3
<i>Streptococcus infantarius</i>	0 (0%)	40 (0.93%)	2/0
<i>Streptococcus intermedius</i>	0 (0%)	20 (0.47%)	6/1
<i>Streptococcus mitis</i>	0 (0%)	37 (0.87%)	5/0
<i>Streptococcus mutans</i>	0 (0%)	20 (0.45%)	9/2/3
<i>Streptococcus oligofermentans</i>	0 (0%)	4 (0.09%)	1/2
<i>Streptococcus oralis</i>	0 (0%)	55 (1.29%)	4/1
<i>Streptococcus parasanguinis</i>	0 (0%)	40 (0.94%)	2/0
<i>Streptococcus pasteurianus</i>	0 (0%)	2 (0.05%)	2/1
<i>Streptococcus pluranimalium</i>	0 (0%)	16 (0.37%)	2/0
<i>Streptococcus pneumoniae</i>	0 (0%)	496 (11.65%)	52/83
<i>Streptococcus porcinus</i>	0 (0%)	0 (0%)	0/1
<i>Streptococcus pseudopneumoniae</i>	16 (0.4%)	16 (0.4%)	1/0
<i>Streptococcus pseudoporcinus</i>	0 (0%)	0 (0%)	0/1
<i>Streptococcus pyogenes</i>	0 (0%)	98 (2.2%)	136/65
<i>Streptococcus salivarius</i>	0 (0%)	33 (0.77%)	9/2
<i>Streptococcus sanguinis</i>	7 (0.16%)	31 (0.72%)	1/5
<i>Streptococcus sobrinus</i>	0 (0%)	15 (0.34%)	4/2
<i>Streptococcus suis</i>	0 (0%)	58 (1.31%)	43/2
<i>Streptococcus tigurinus</i>	38 (0.9%)	38 (0.9%)	1/0
<i>Streptococcus uberis</i>	0 (0%)	5 (0.11%)	2/2
<i>Streptococcus urinalis</i>	0 (0%)	0 (0%)	0/1
<i>Enterococcus casseliflavus</i>	1 (0.02%)	4 (0.09%)	2/0
<i>Enterococcus cecorum</i>	1 (0.02%)	28 (0.66%)	6/1

<b>Species</b>	<b>The lowest intraspecies nucleotide difference compared to GenBank sequences</b>	<b>The highest intraspecies nucleotide difference compared to GenBank sequences</b>	<b>Amount of complete genome sequences/genome assembly/ 16S-23S rRNA region (GenBank)</b>
<i>Enterococcus durans</i>	0 (0%)	42 (0.97%)	3/0
<i>Enterococcus faecium</i>	0 (0%)	102 (2.33%)	70/58
<i>Enterococcus faecalis</i>	0 (0%)	17 (0.4%)	31/1
<i>Enterococcus hirae</i>	0 (0%)	71 (2.67%)	3/1

<sup>a</sup> For species not included in a Table (*Streptococcus adjacens*, *Streptococcus cristatus*, *Streptococcus difficilis*, *Streptococcus downei*, *Streptococcus infantis*, *Streptococcus ovis*, *Streptococcus saccharolyticus*, *Streptococcus sinensis*, *Enterococcus avium*, *Enterococcus porcinus*, *Enterococcus raffinosus*), there are no reference genomes available.

Figure S1

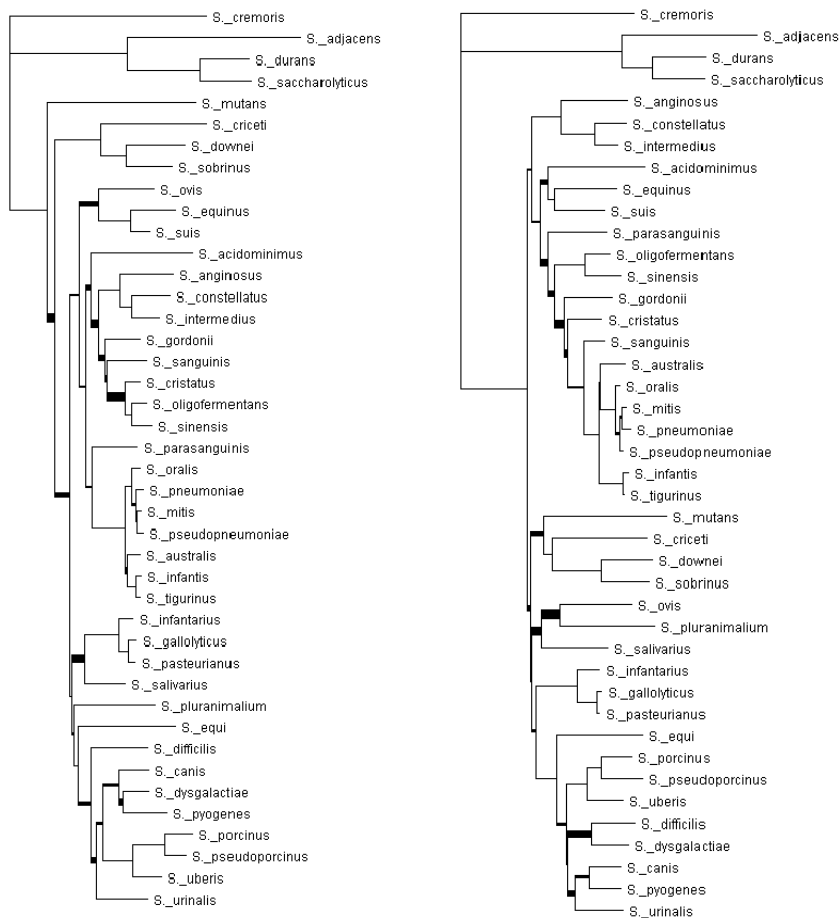
The comparison between phylogenetic trees based on 16S-23S rRNA region and 16S rRNA, *rpoB*, *sodA* and *tuf* genes for both *Streptococcus* and *Enterococcus* species.

The thicker branches reflect a low topological score and result from different positions on the compared trees.

### *Streptococcus*

#### 16S rRNA gene vs 16S-23S rRNA region

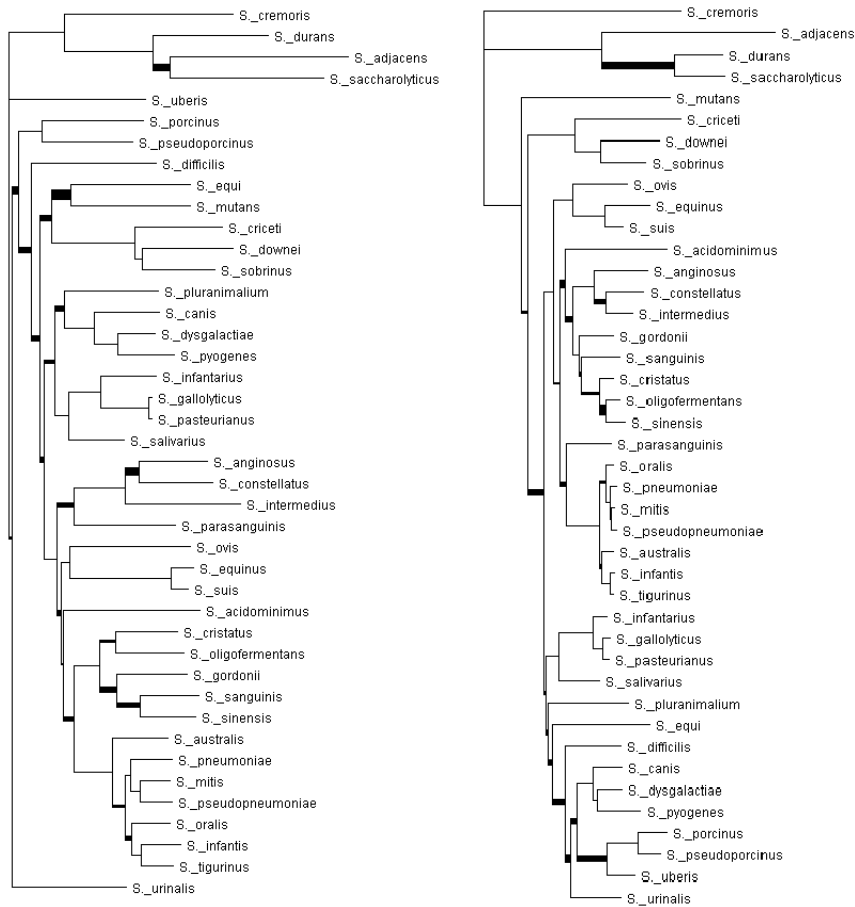
Overall topological score = 72.4 %





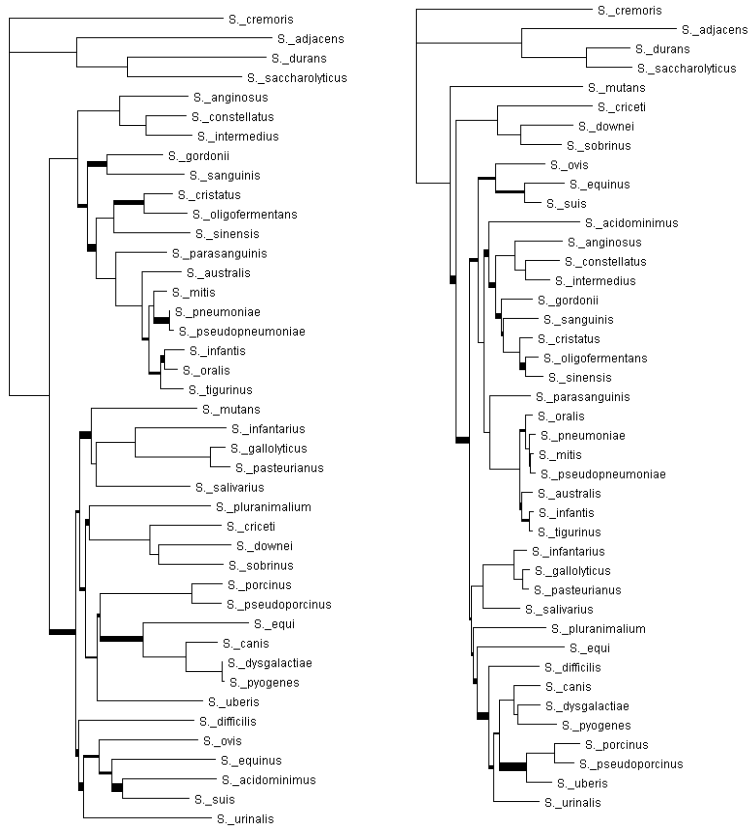
*rpoB* gene vs 16S-23S rRNA region

Overall topological score = 71.6 %



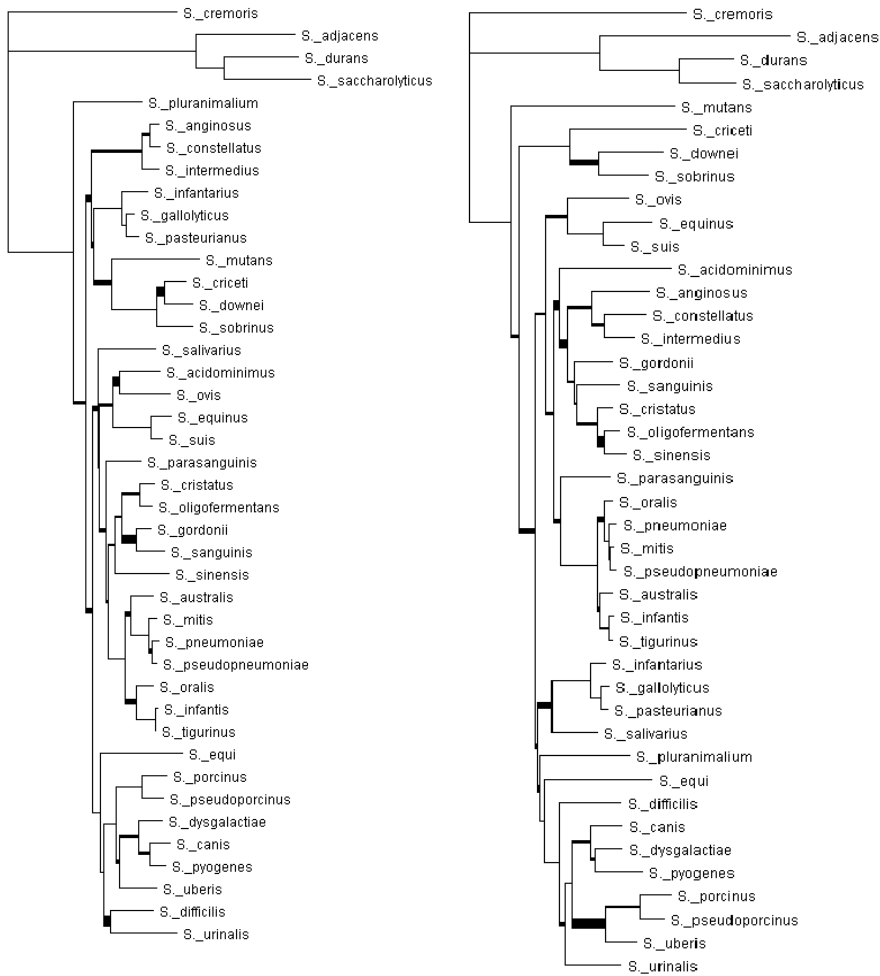
*sodA* gene vs 16S-23S rRNA region

Overall topological score = 70.4 %



*tuf* gene vs 16S-23S rRNA region

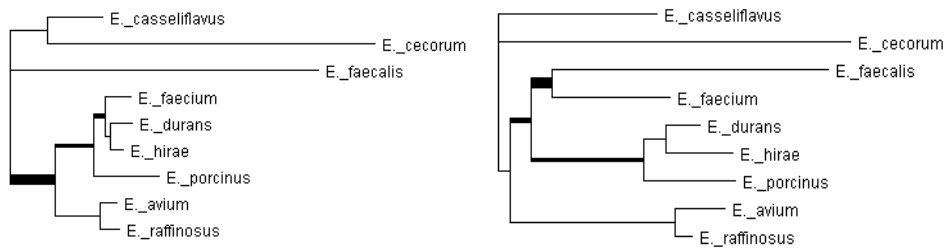
Overall topological score = 70 %



## *Enterococcus*

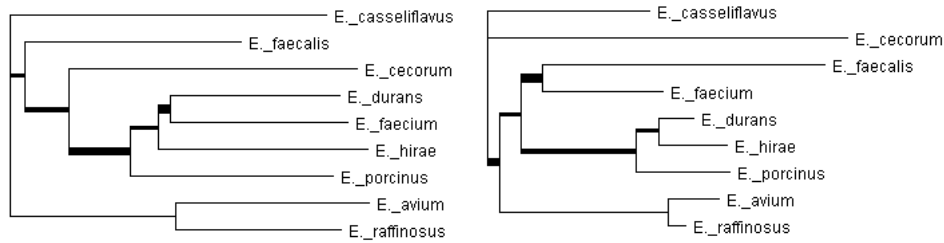
### 16S rRNA gene vs 16S-23S rRNA region

Overall topological score = 76.7 %



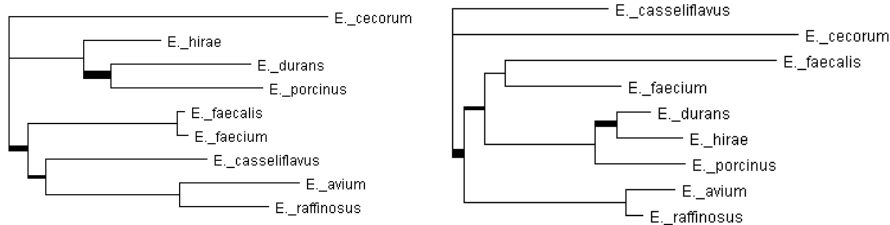
### *rpoB* gene vs 16S-23S rRNA region

Overall topological score = 62.5 %



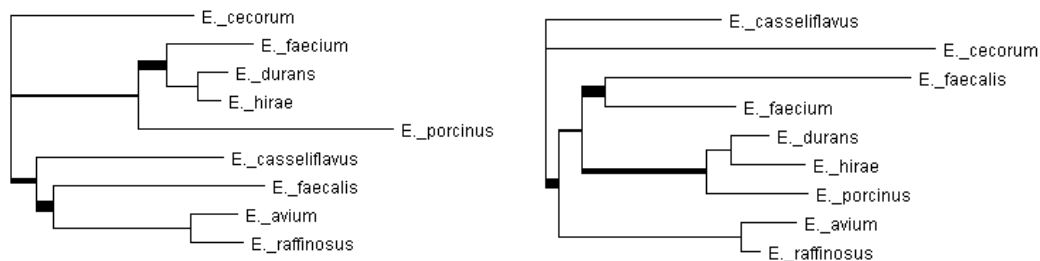
### *sodA* gene vs 16S-23S rRNA region

Overall topological score = 75.6 %



### *tuf* gene vs 16S-23S rRNA region

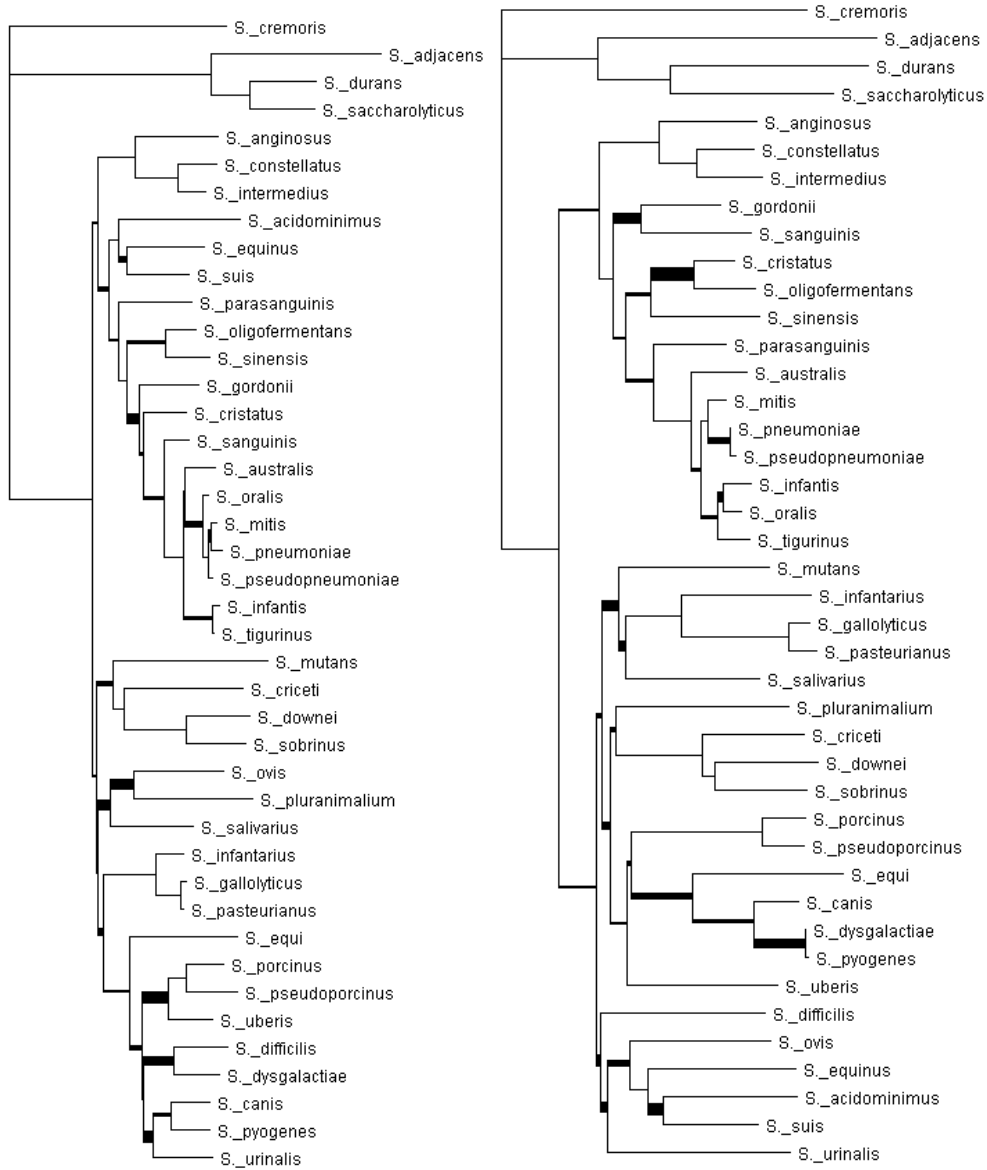
Overall topological score = 66.4 %



*Streptococcus*

16S rRNA gene vs *sodA* gene

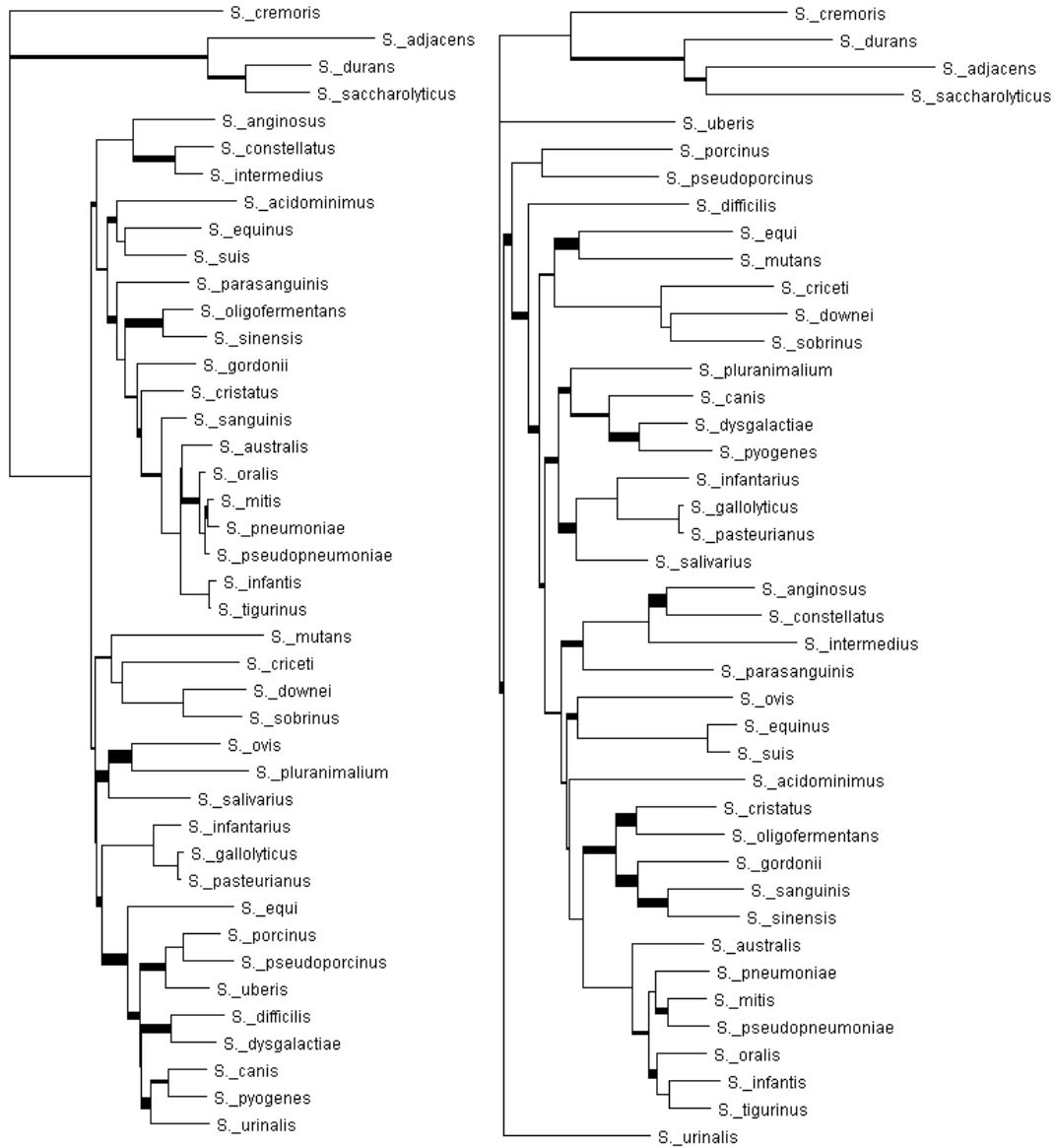
Overall topological score = 66.5 %



*Streptococcus*

16S rRNA gene vs *rpoB* gene

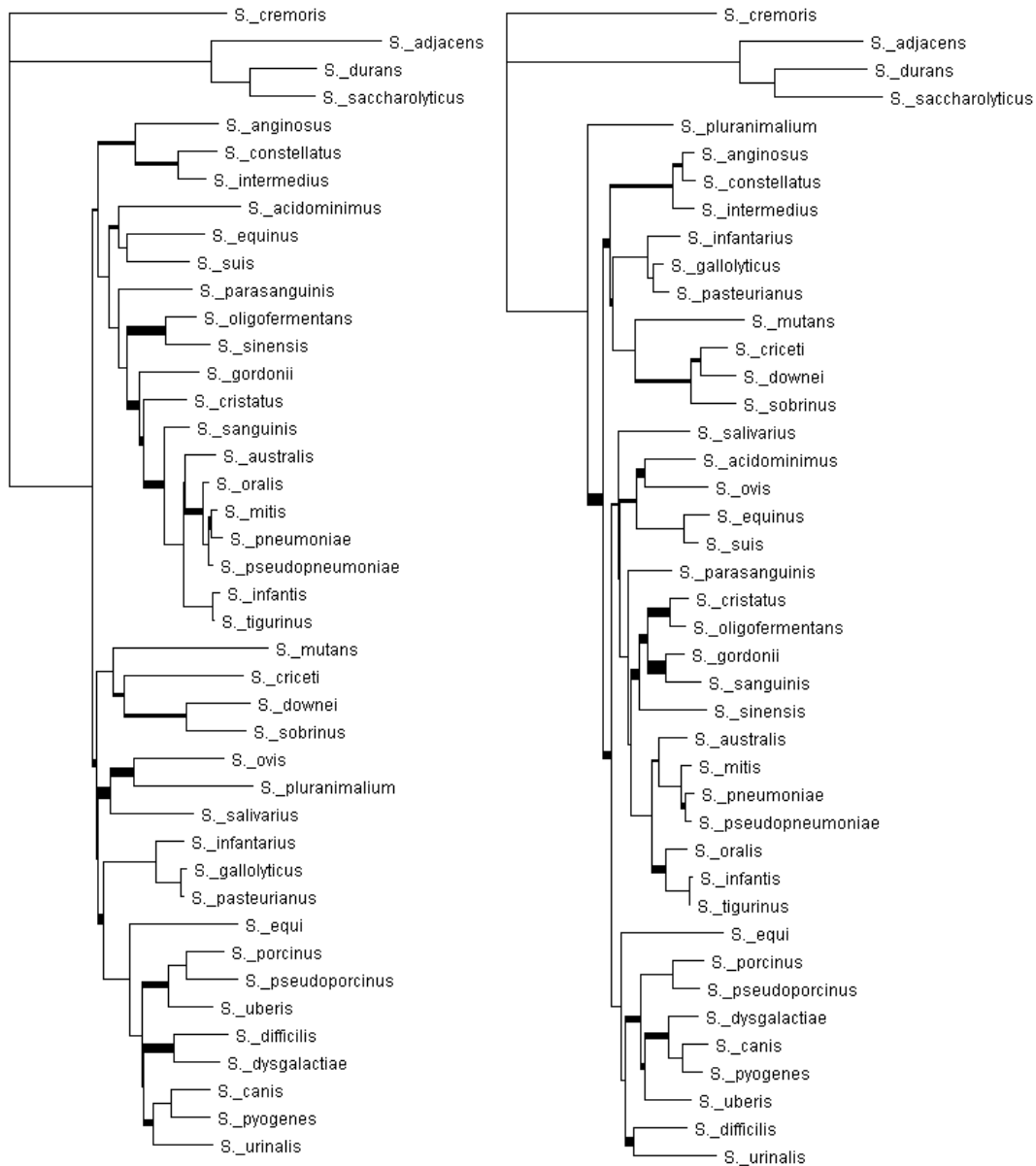
Overall topological score = 61.7 %



# *Streptococcus*

16S rRNA gene vs *tuf* gene

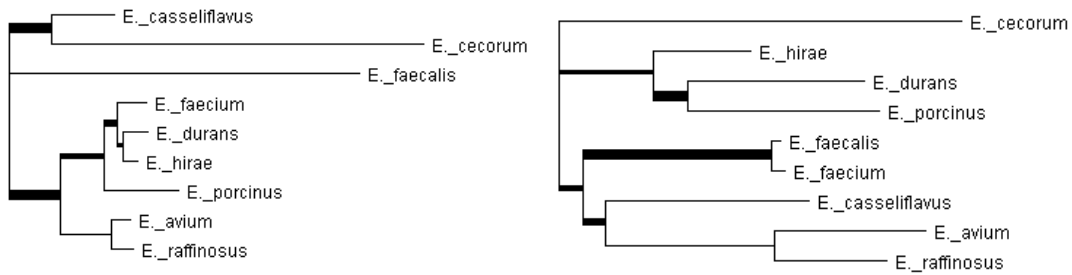
Overall topological score = 69.1 %



*Enterococcus*

16S rRNA gene vs *sodA* gene

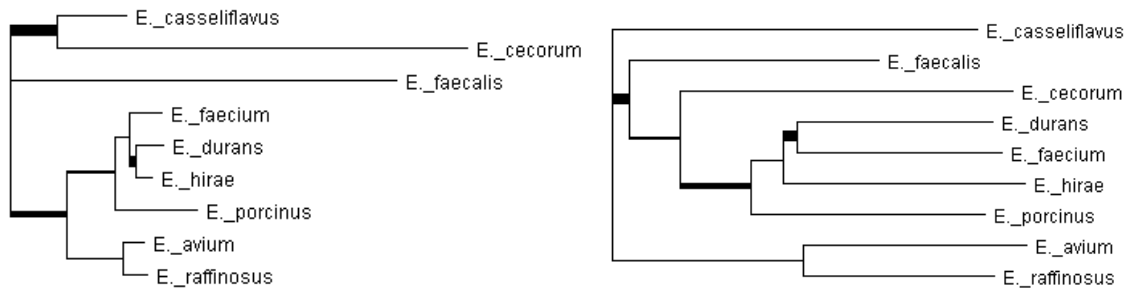
Overall topological score = 56.4 %



*Enterococcus*

16S rRNA gene vs *rpoB* gene

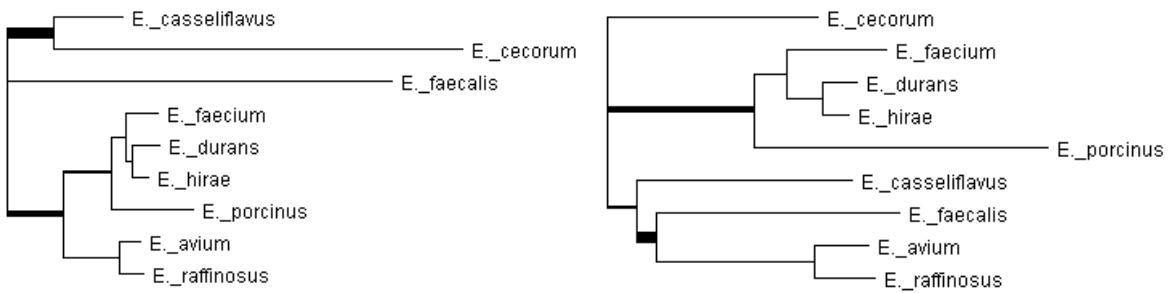
Overall topological score = 67.0 %



*Enterococcus*

16S rRNA gene vs *tuf* gene

Overall topological score = 78.9 %

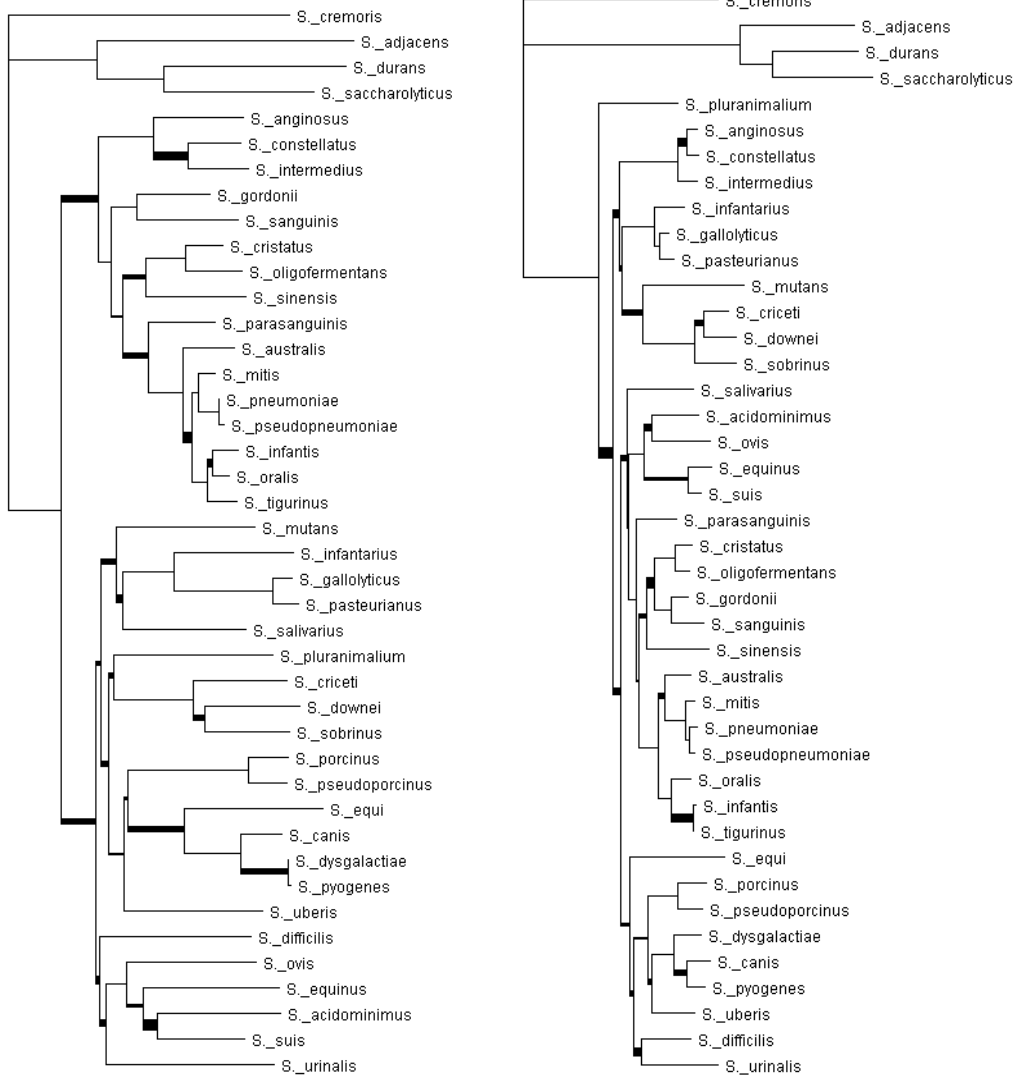




*Streptococcus*

*tuf* gene vs *sodA* gene

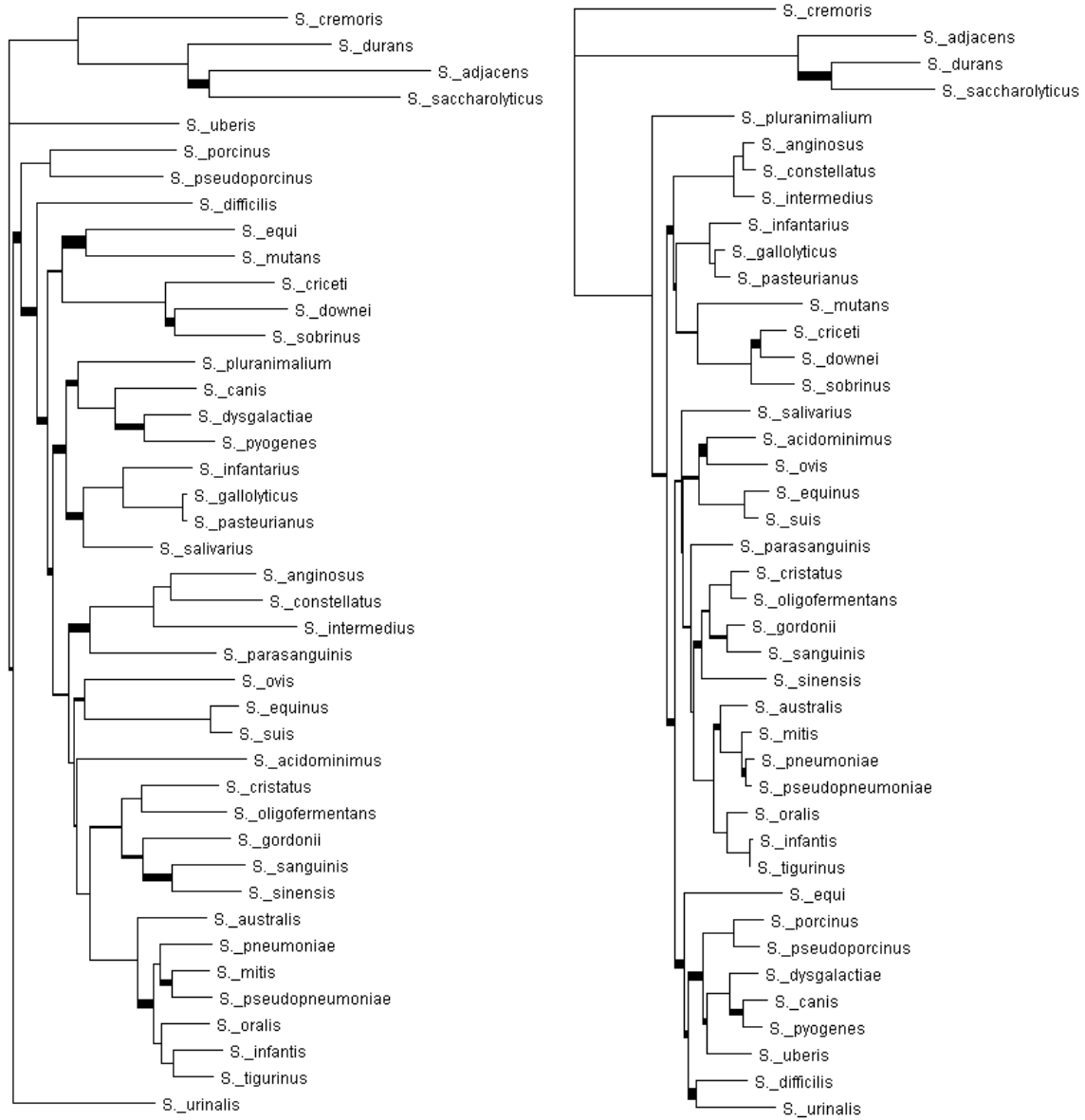
Overall topological score = 70.4 %



# *Streptococcus*

## *rpoB* gene vs *tuf* gene

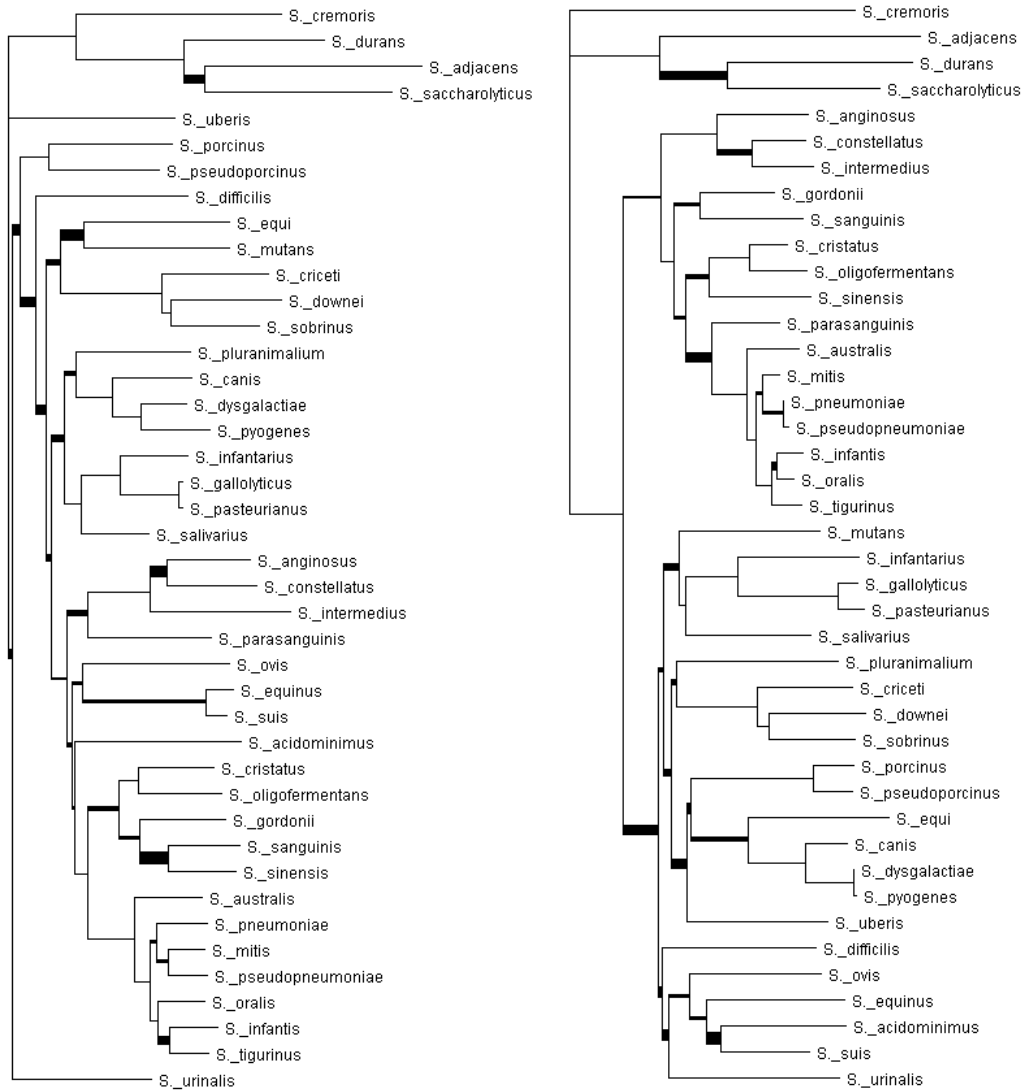
Overall topological score = 70.3 %



# *Streptococcus*

*rpoB* gene vs *sodA* gene

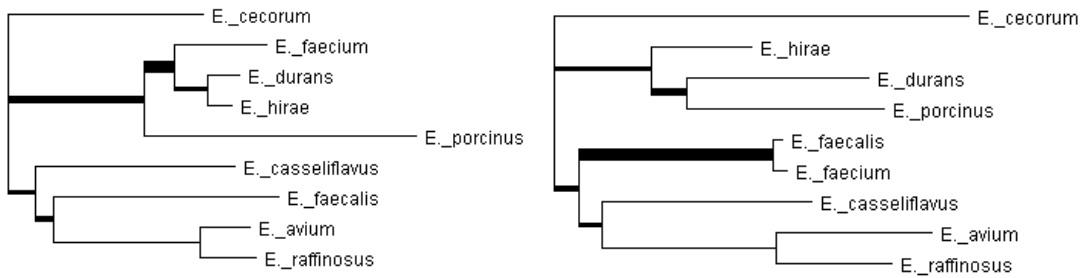
Overall topological score = 67.3 %



*Enterococcus*

*tuf* gene vs *sodA* gene

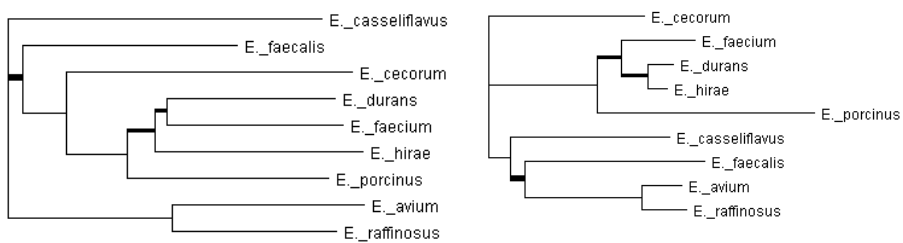
Overall topological score = 62.8 %



*Enterococcus*

*rpoB* gene vs *tuf* gene

Overall topological score = 80.6 %



*Enterococcus*

*rpoB* gene vs *sodA* gene

Overall topological score = 68.9 %

