

## **Supplementary information**

### **Plastid-localized amino acid biosynthetic pathways of Plantae are predominantly composed of non-cyanobacterial enzymes**

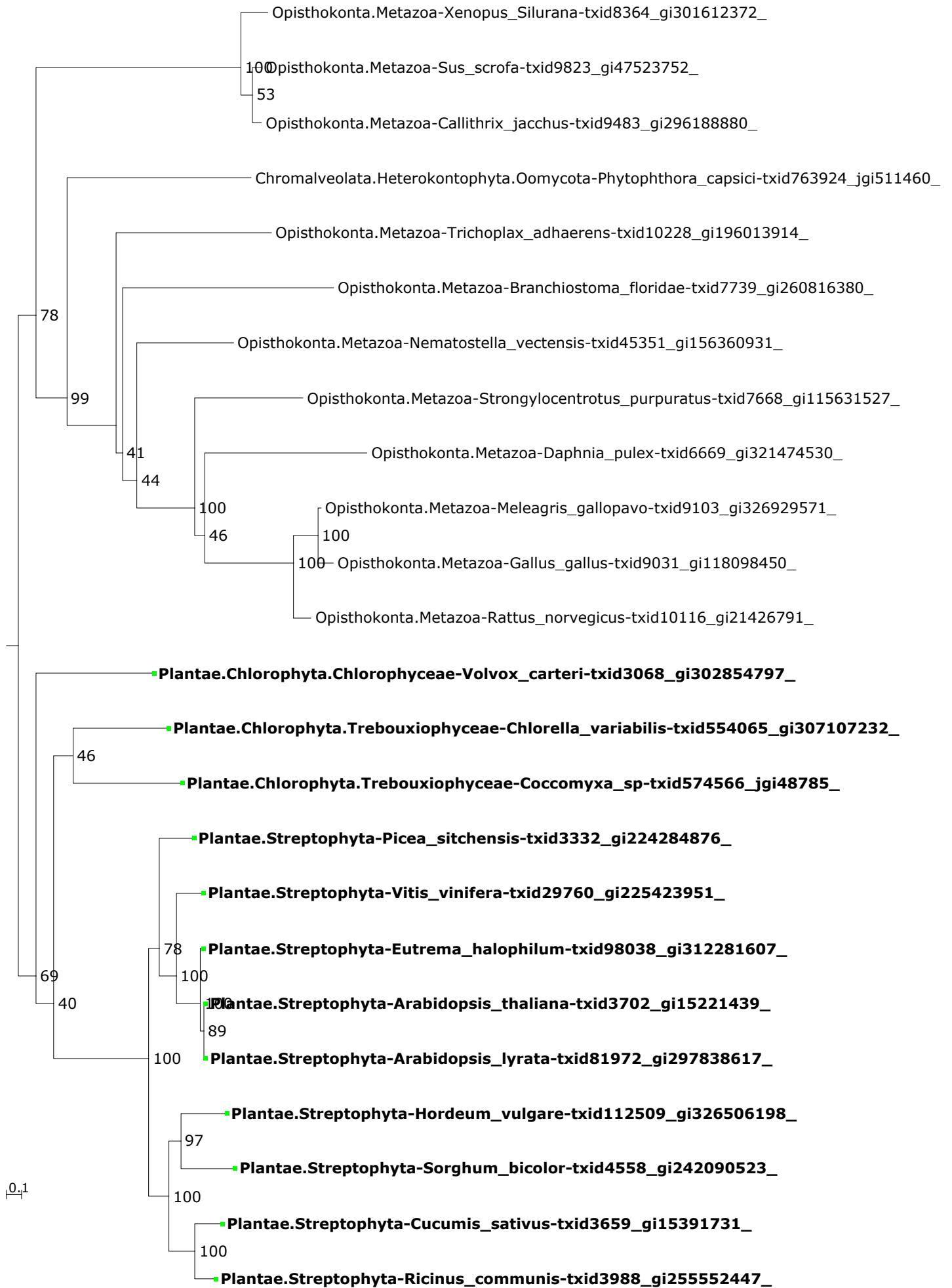
Adrian Reyes-Prieto and Ahmed Moustafa

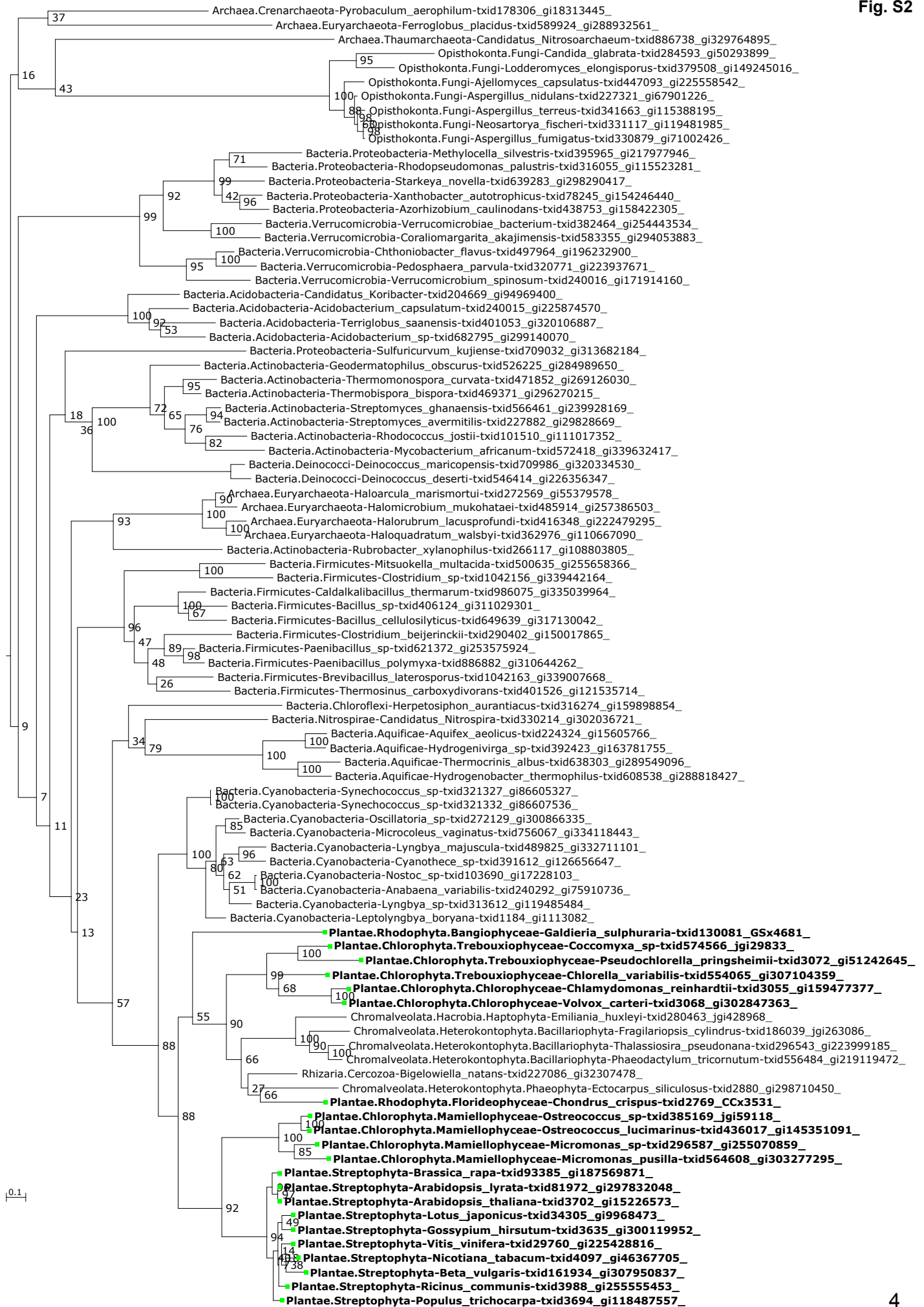
#### **Contents**

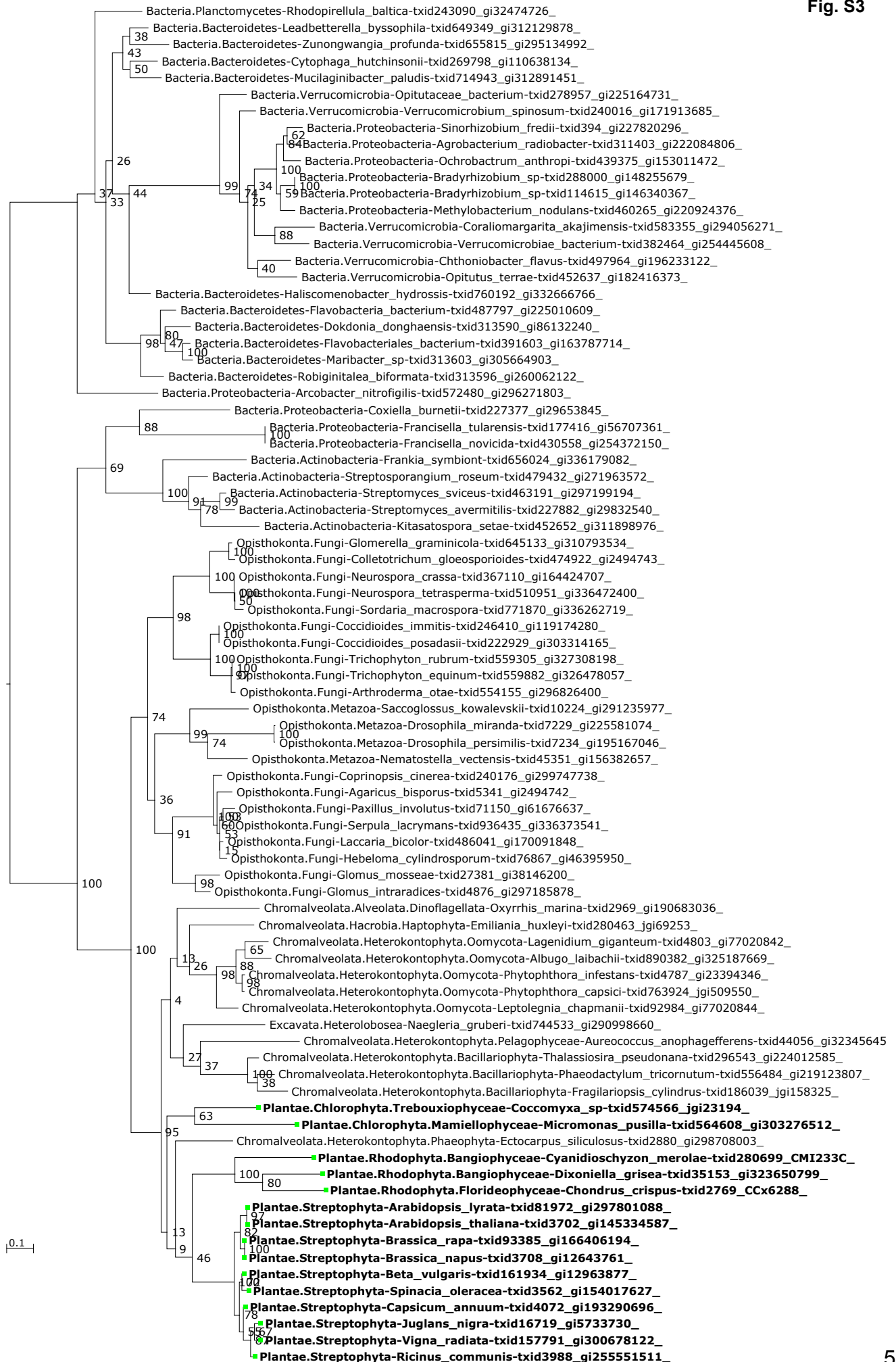
	<b>page</b>
Supplementary Figures S1-S156	2
Supplementary Tables S1-S2	159

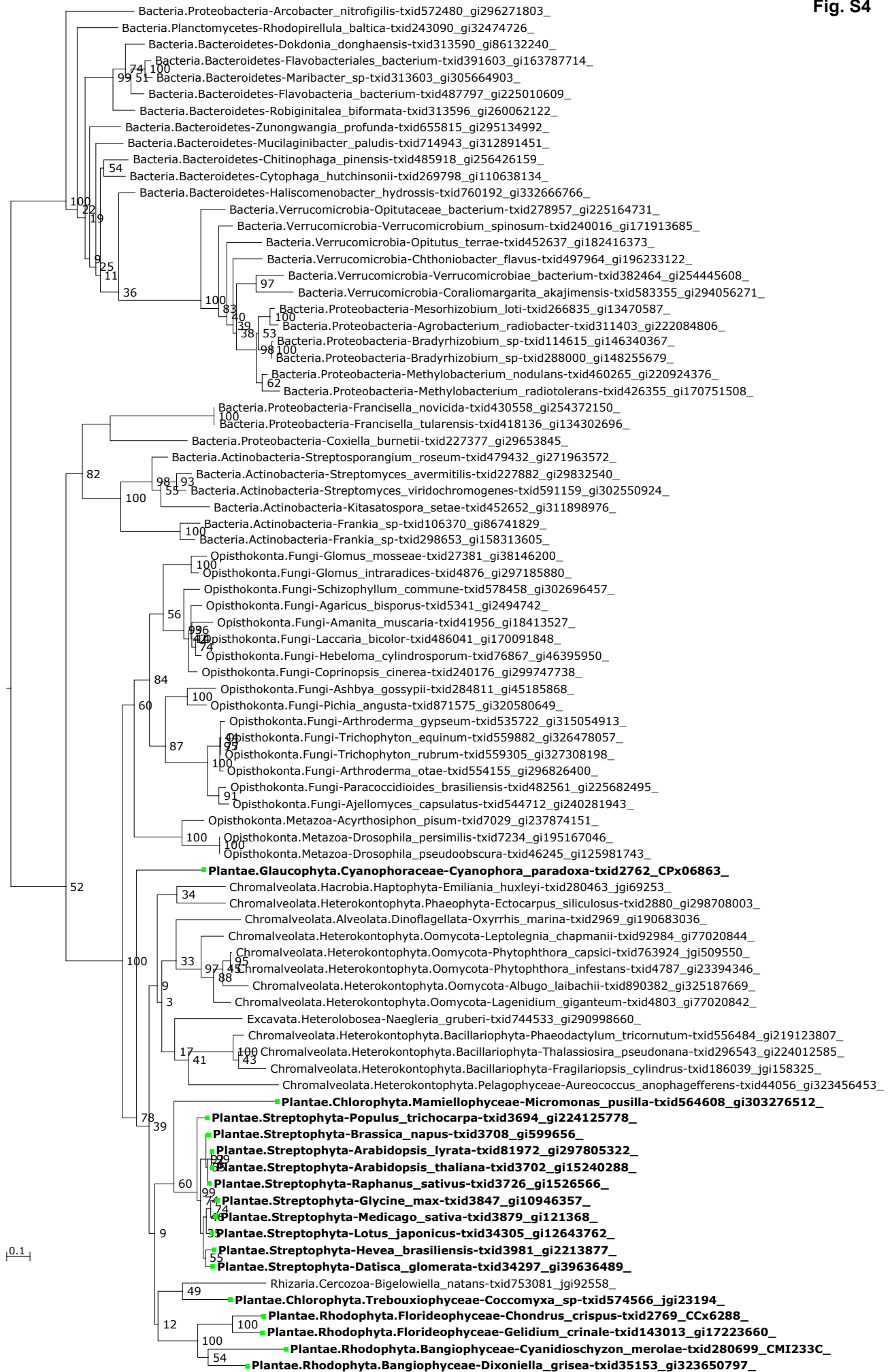
## Figures caption

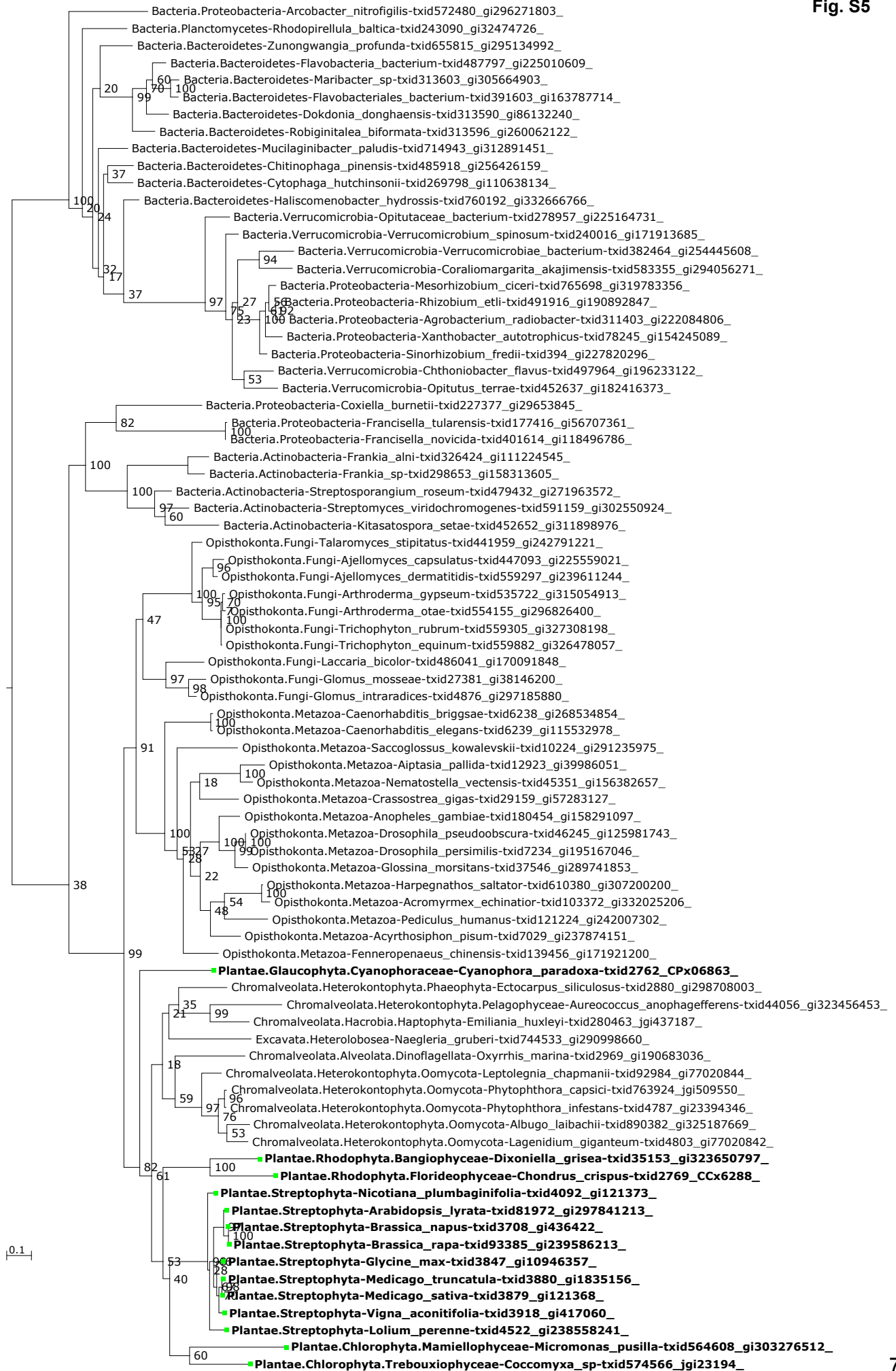
**Supplemental figures S1 to S156.** Maximum likelihood phylogenetic trees of the *Arabidopsis thaliana* enzymes involved in amino acid biosynthesis. Page numbers correspond to detailed descriptions in the main text and information presented in Table S1 (Supplementary Information). The numbers near (right side) nodes indicate RaxML bootstrap support values. Branch lengths are proportional to the number of substitutions per site (see the scale bars)

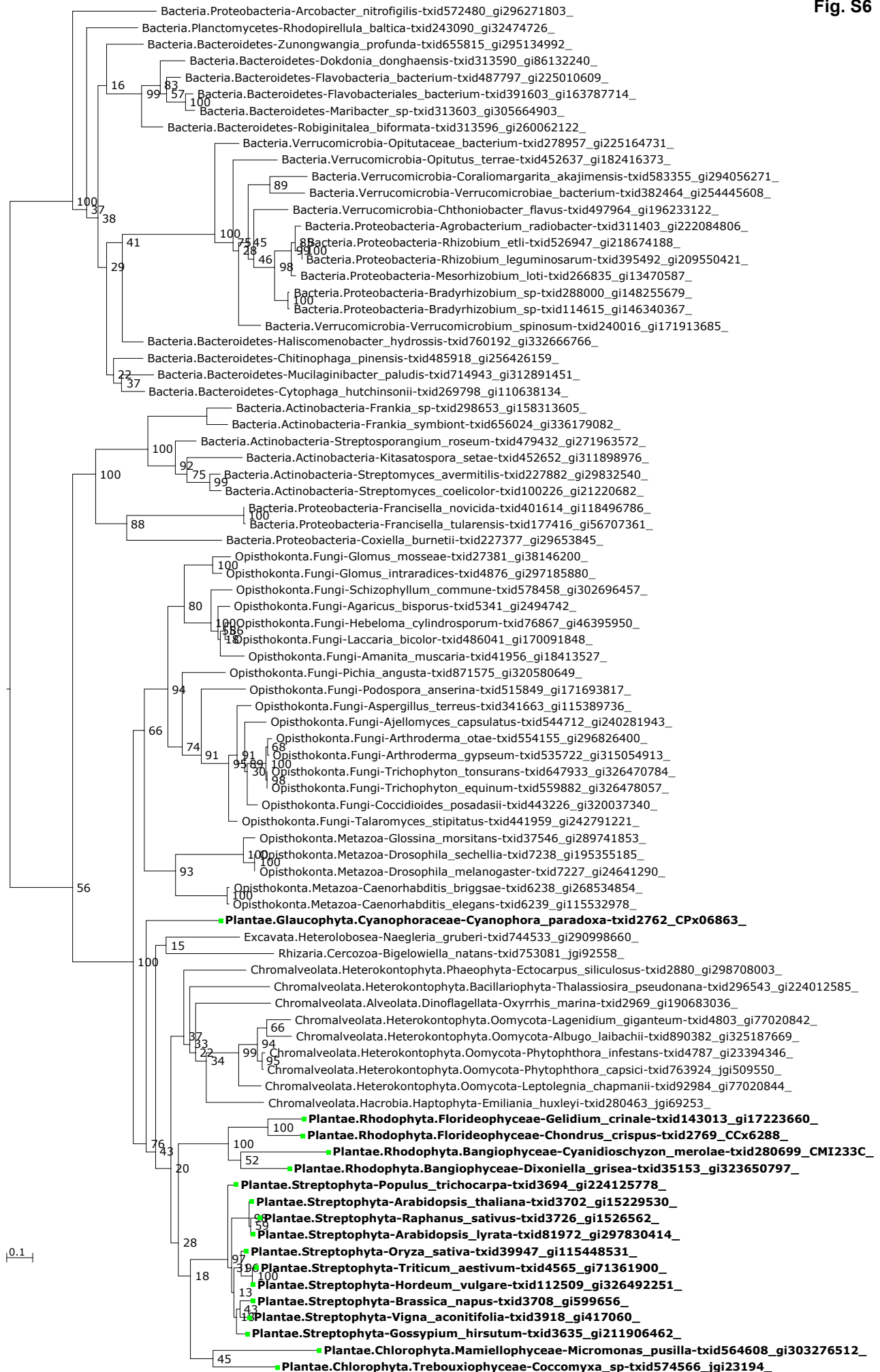




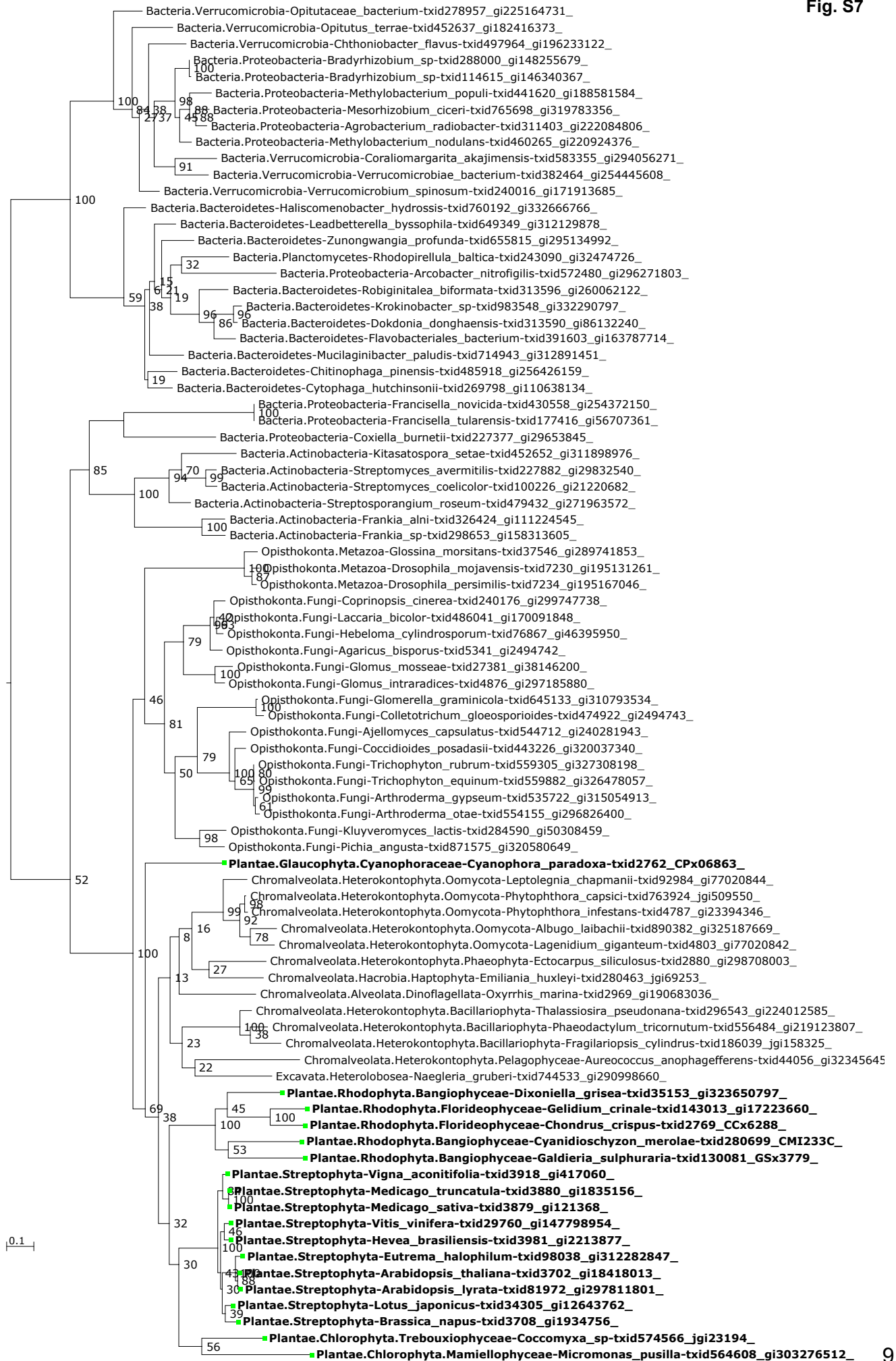


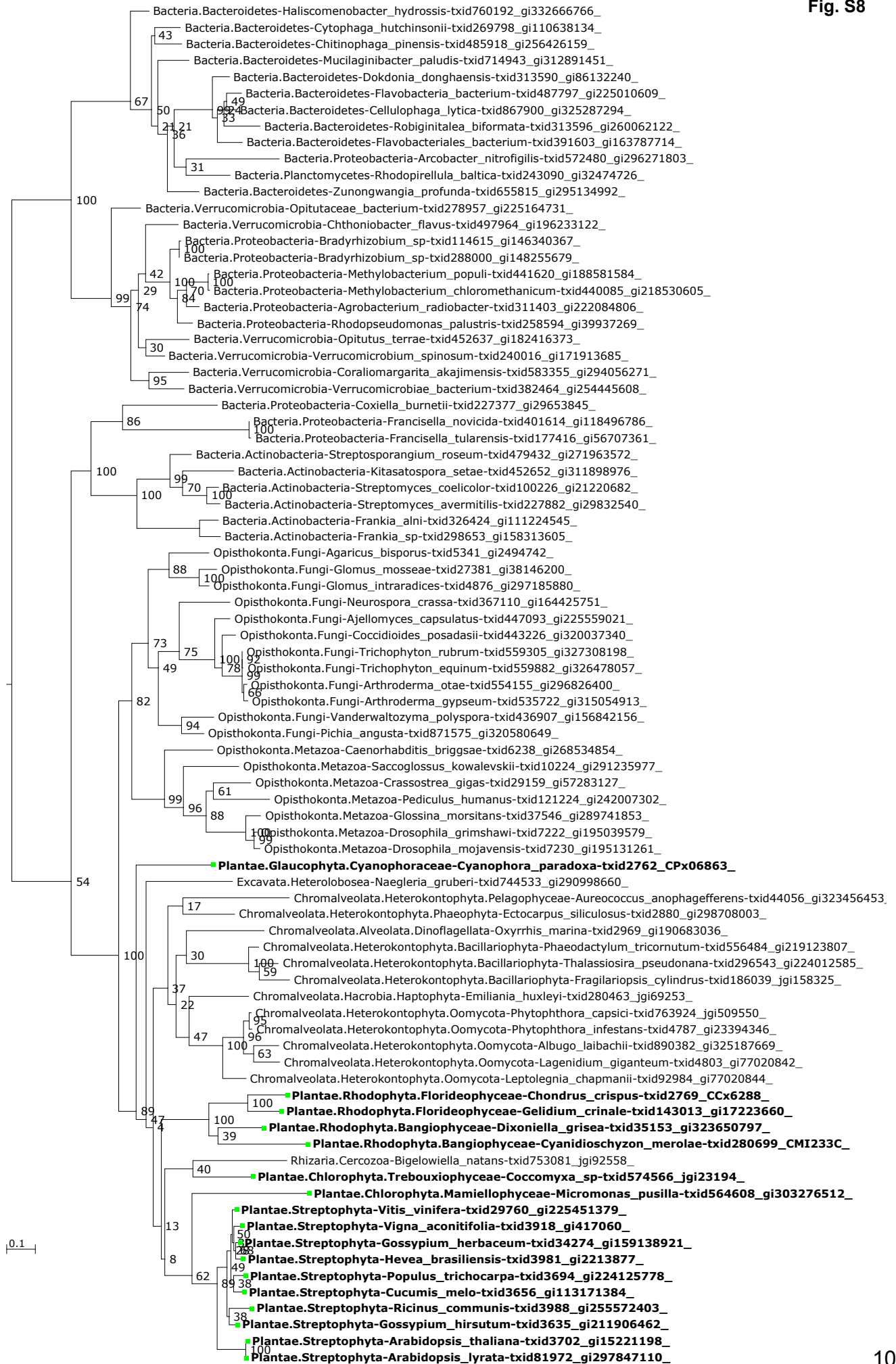


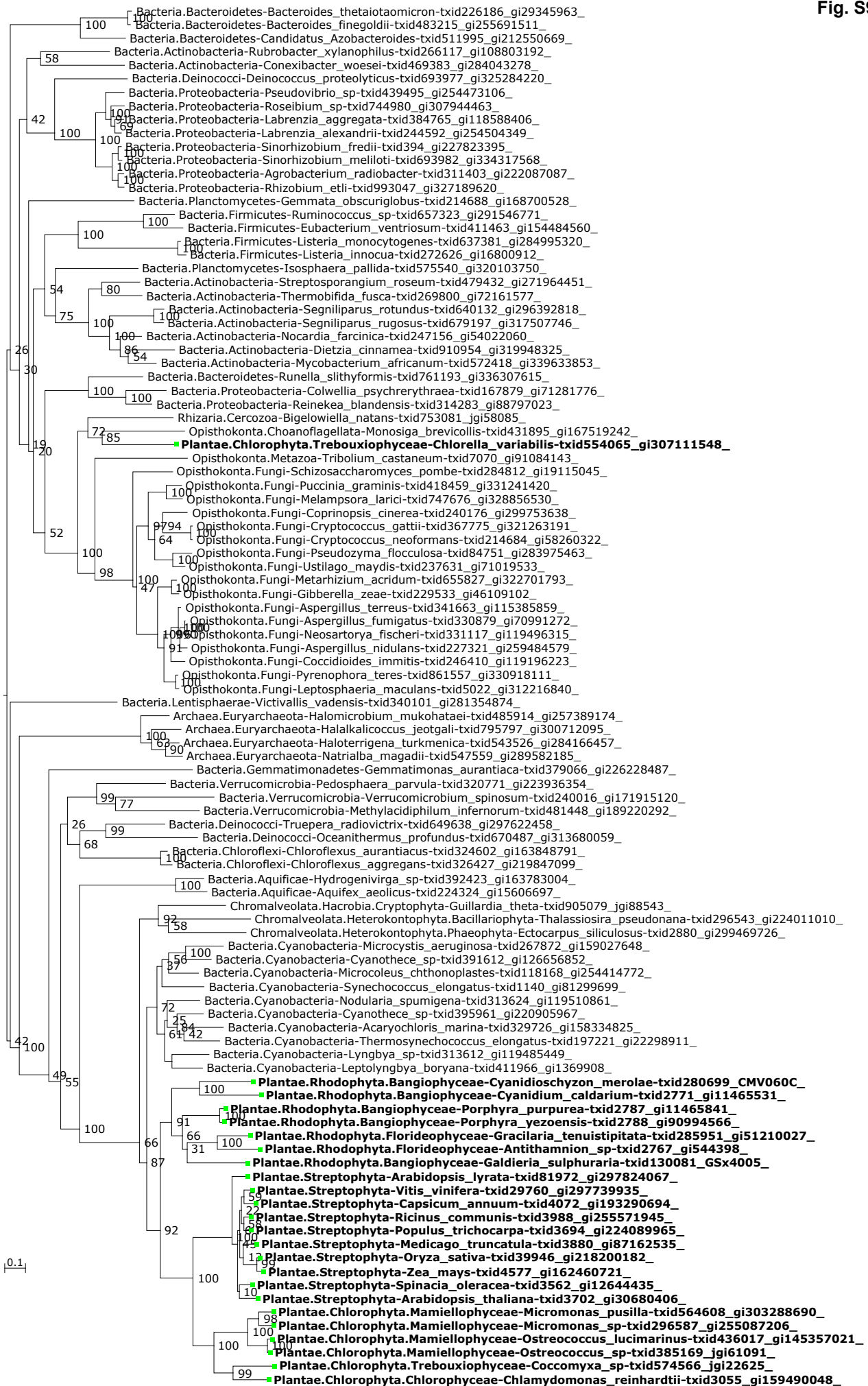


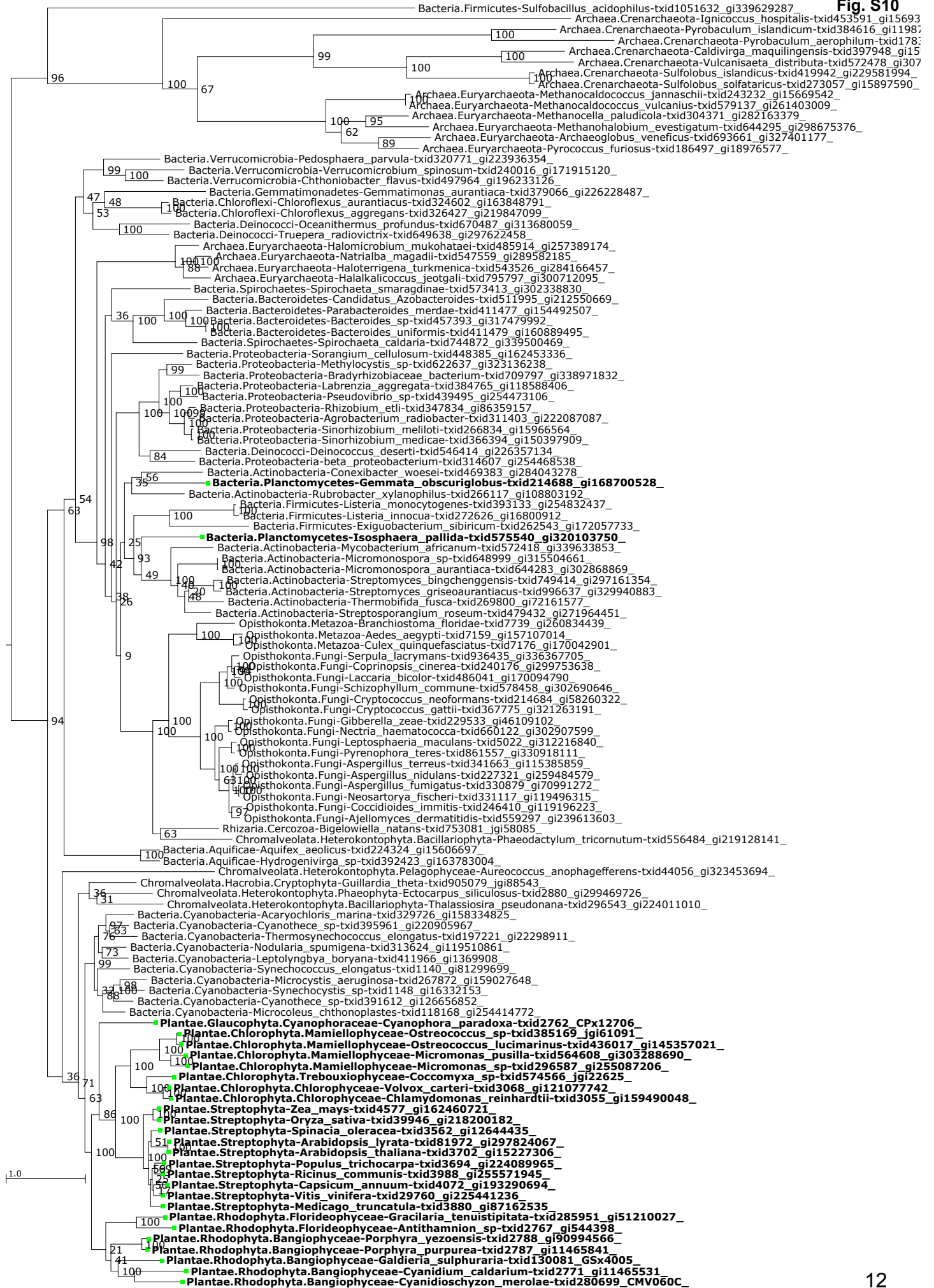


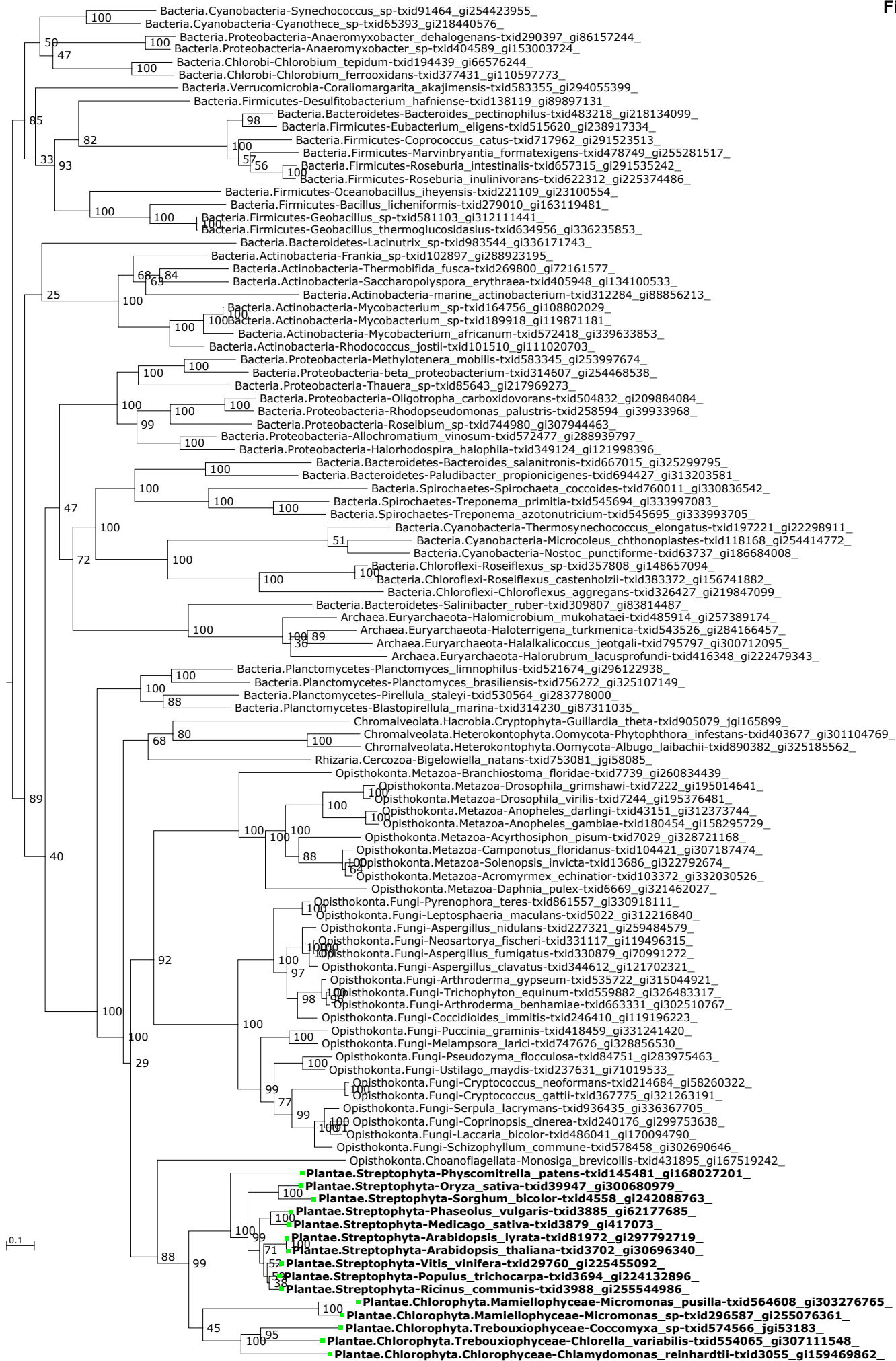


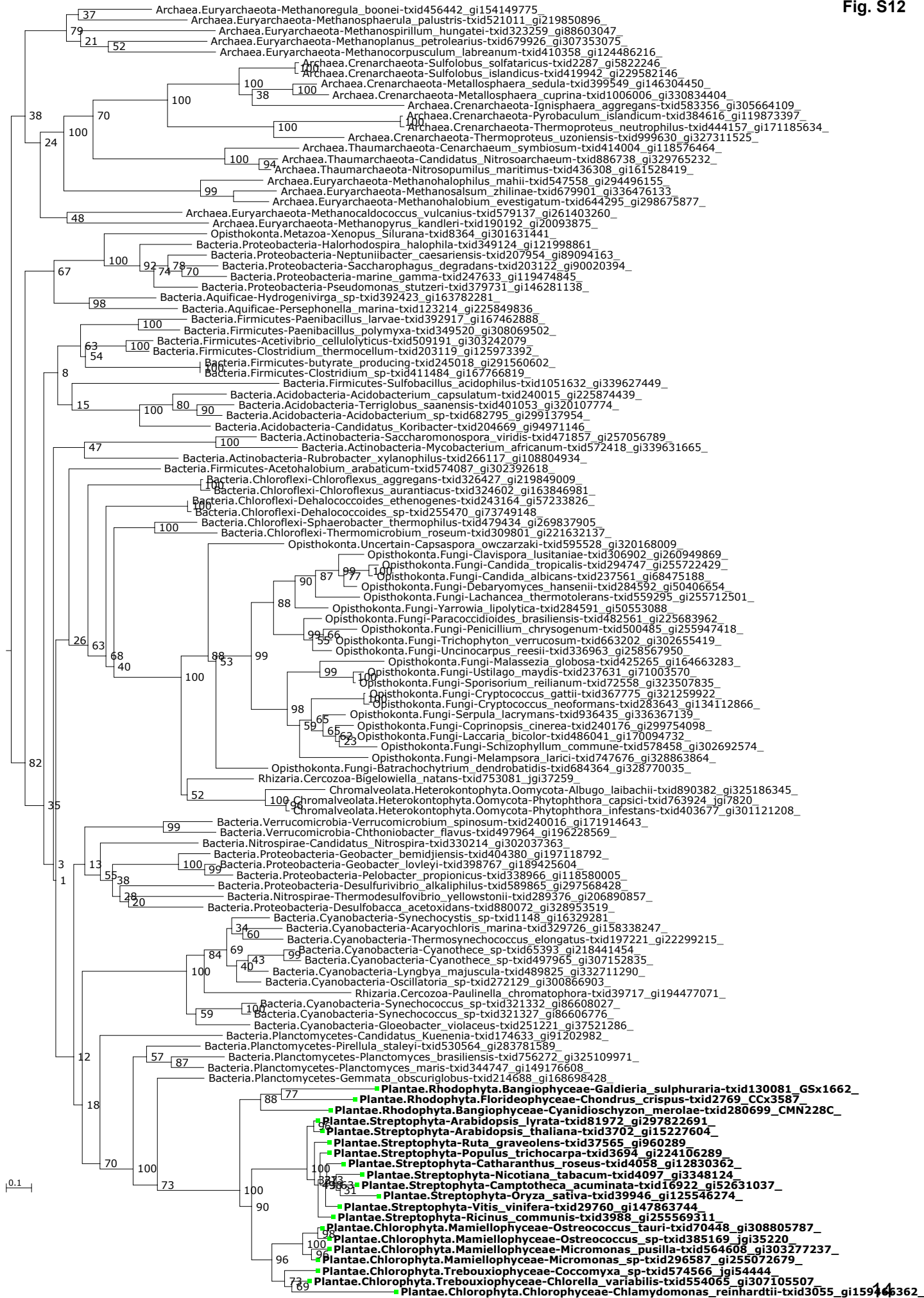


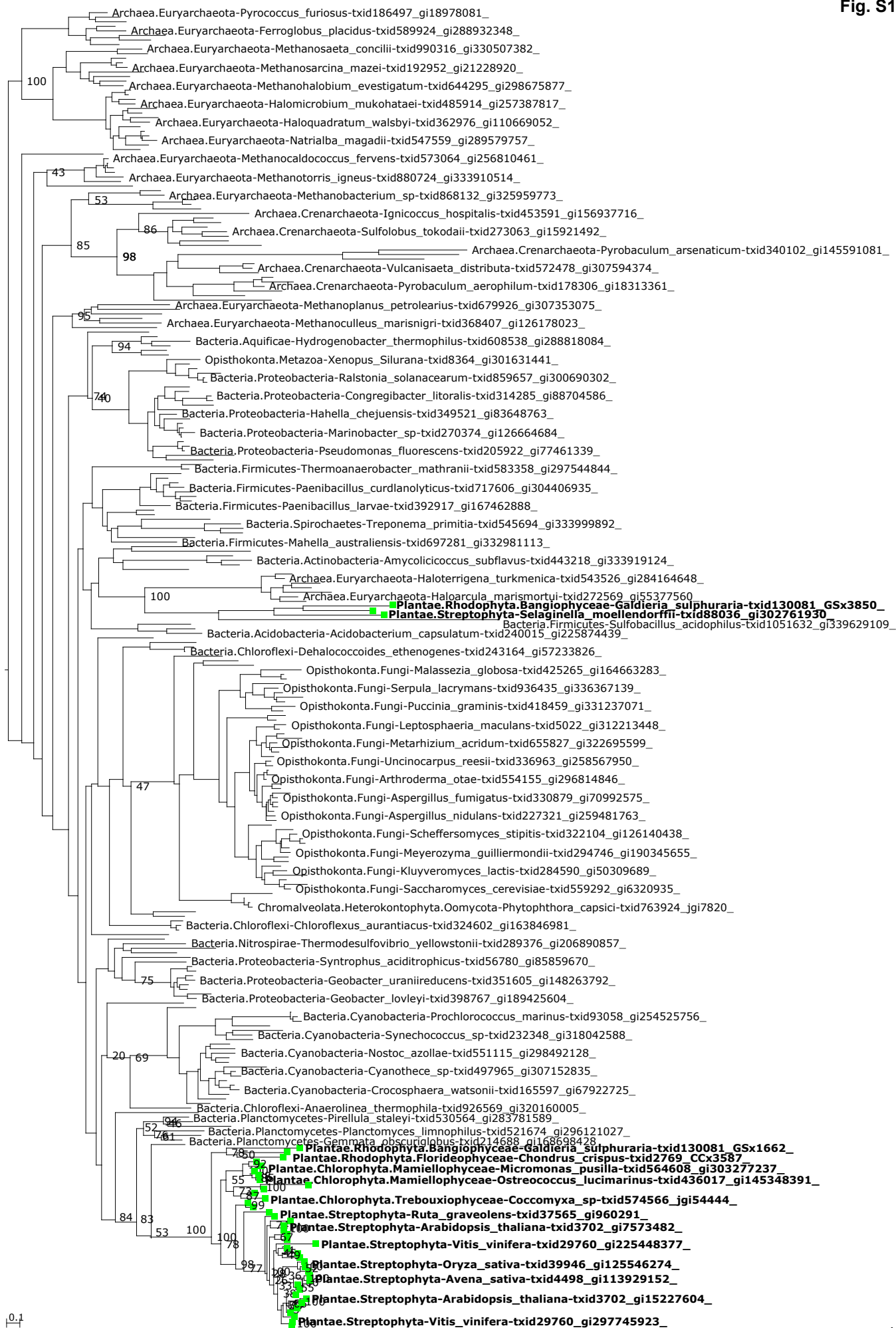


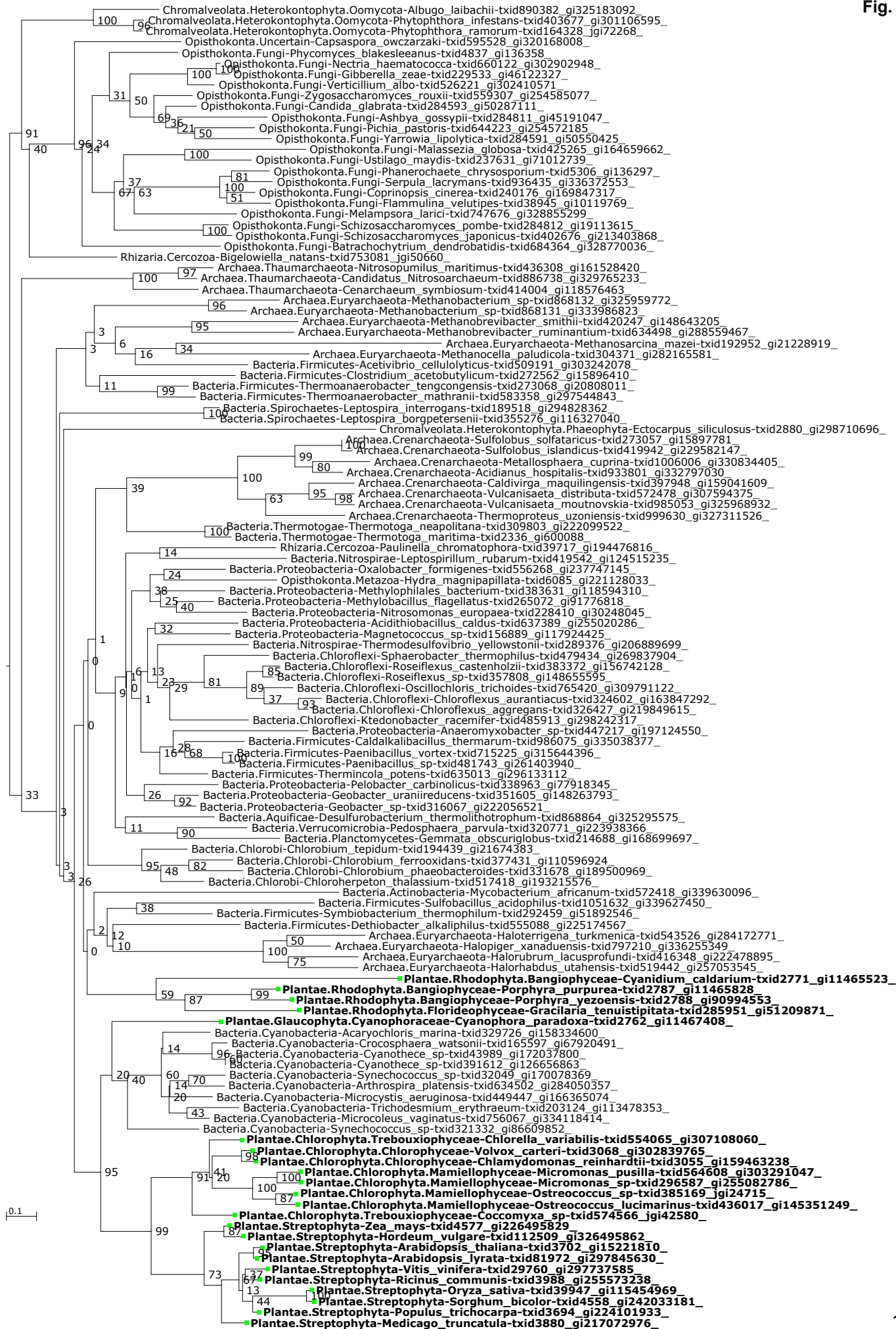












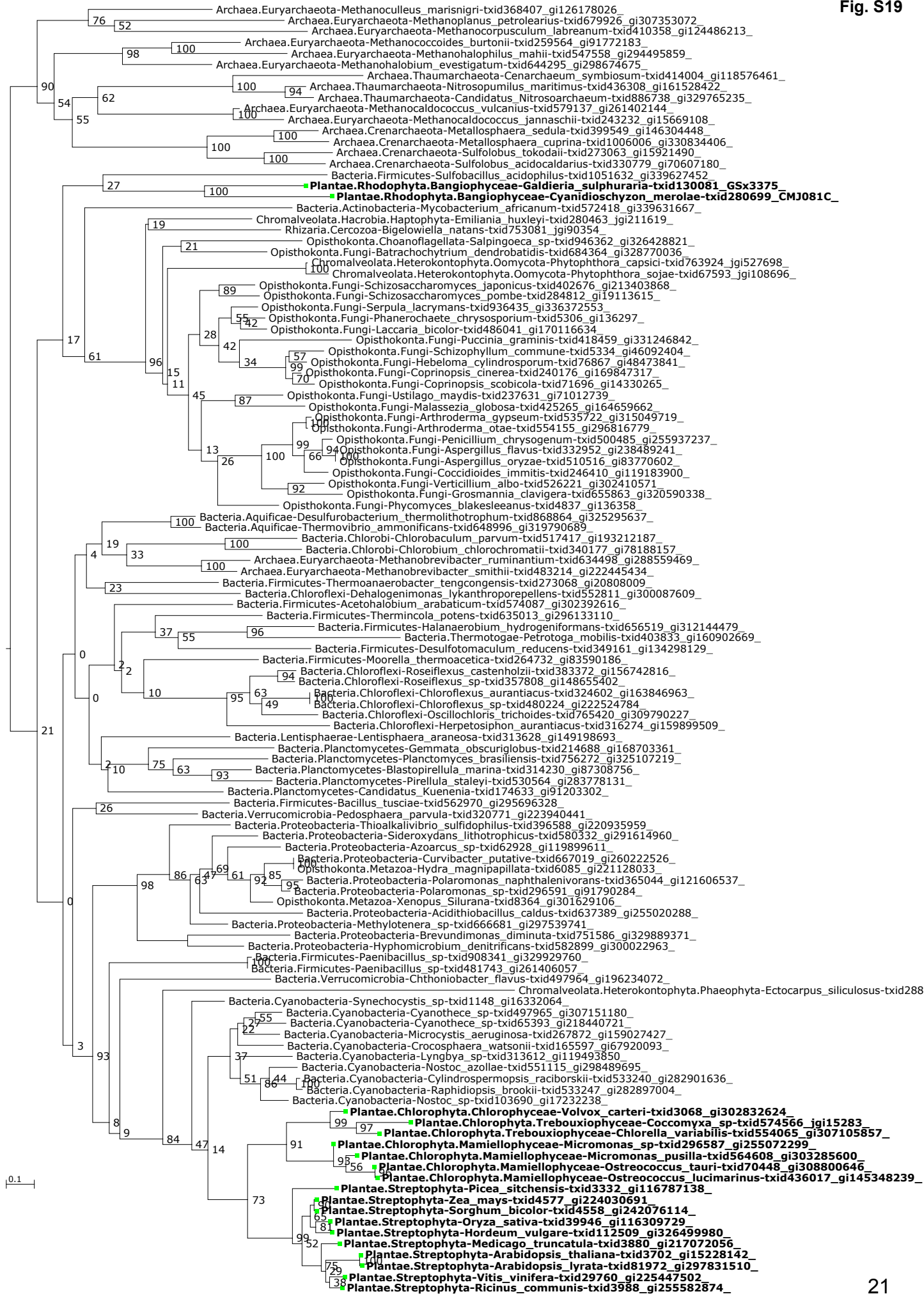


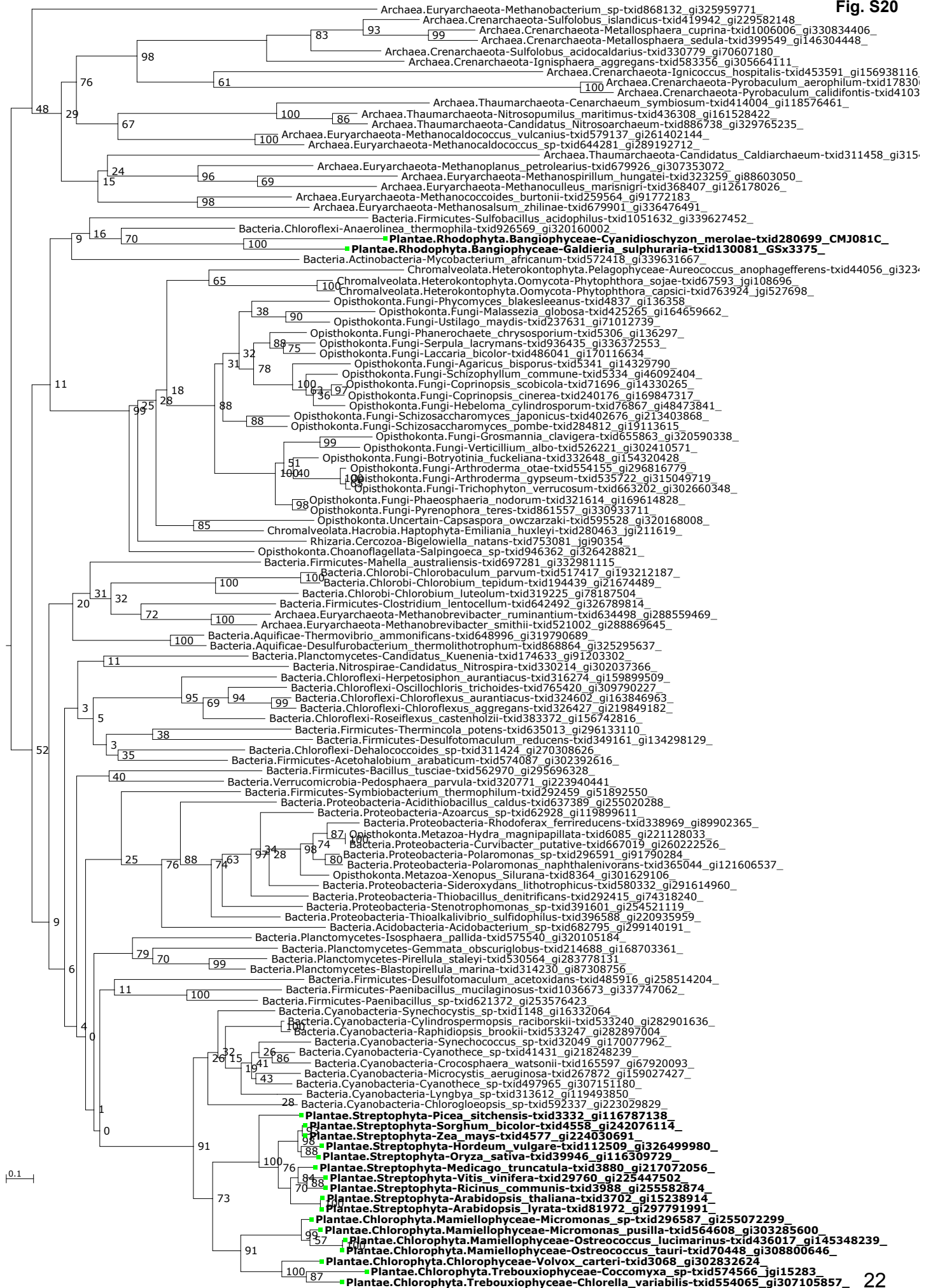










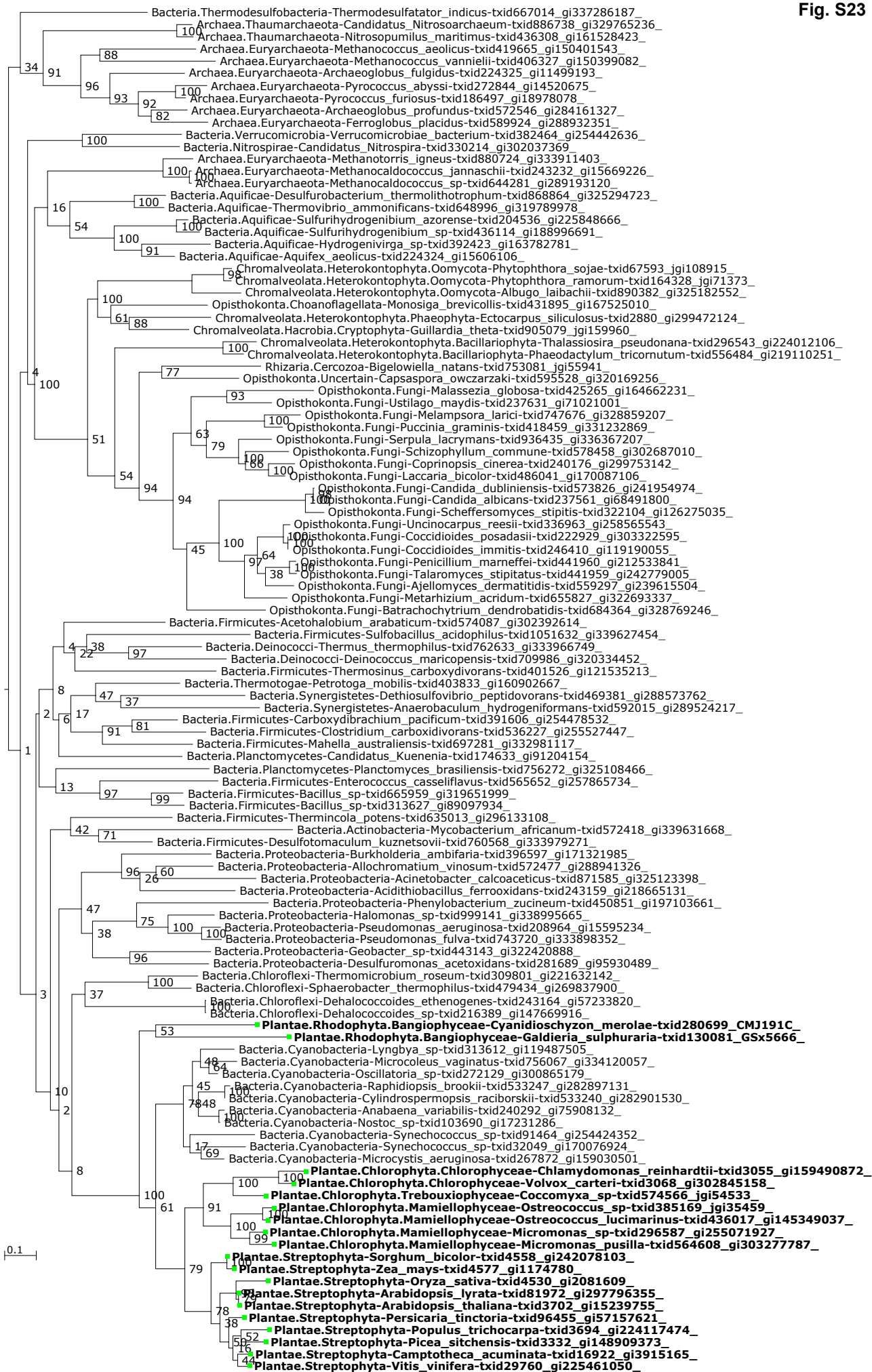


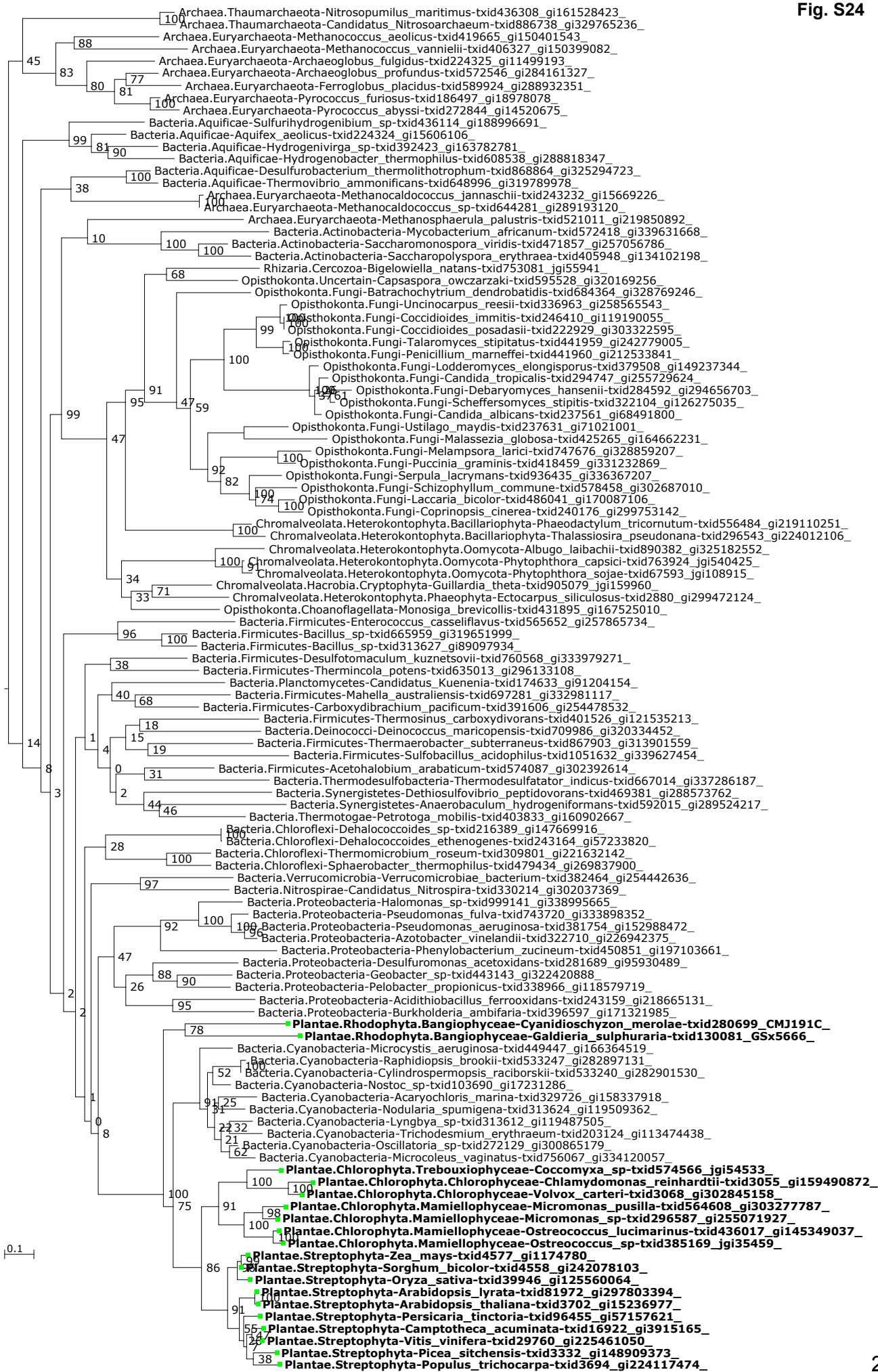
0.1

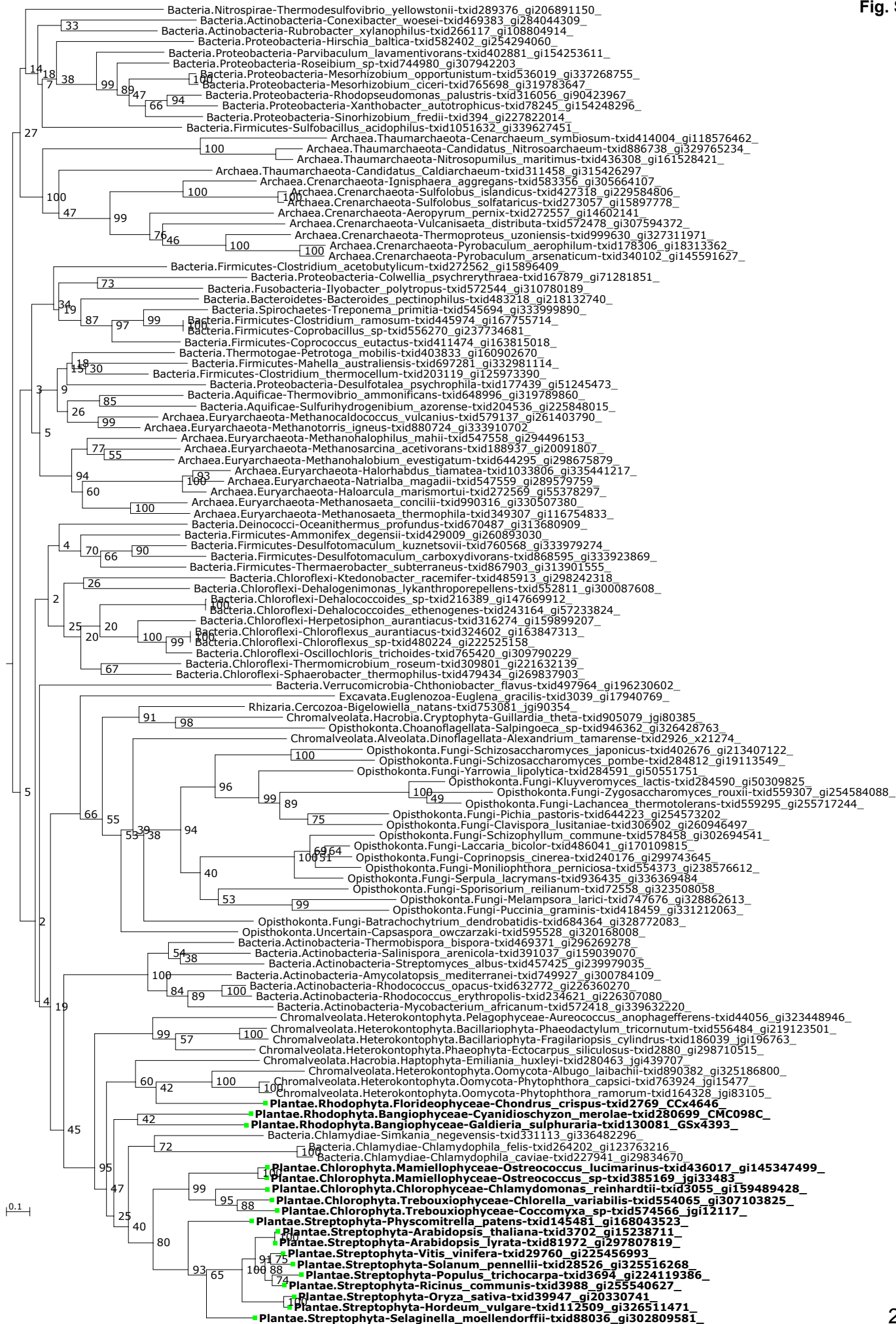


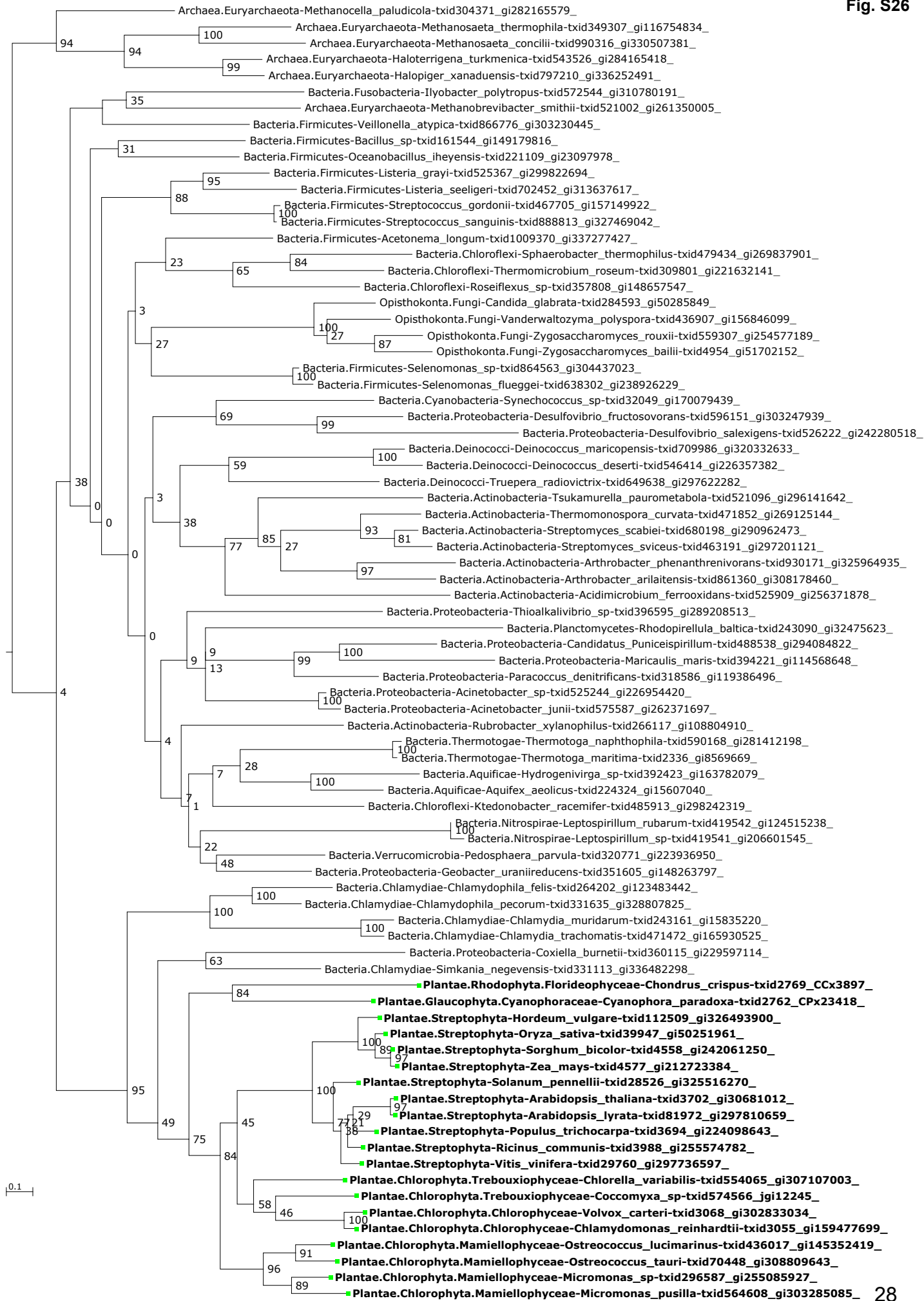




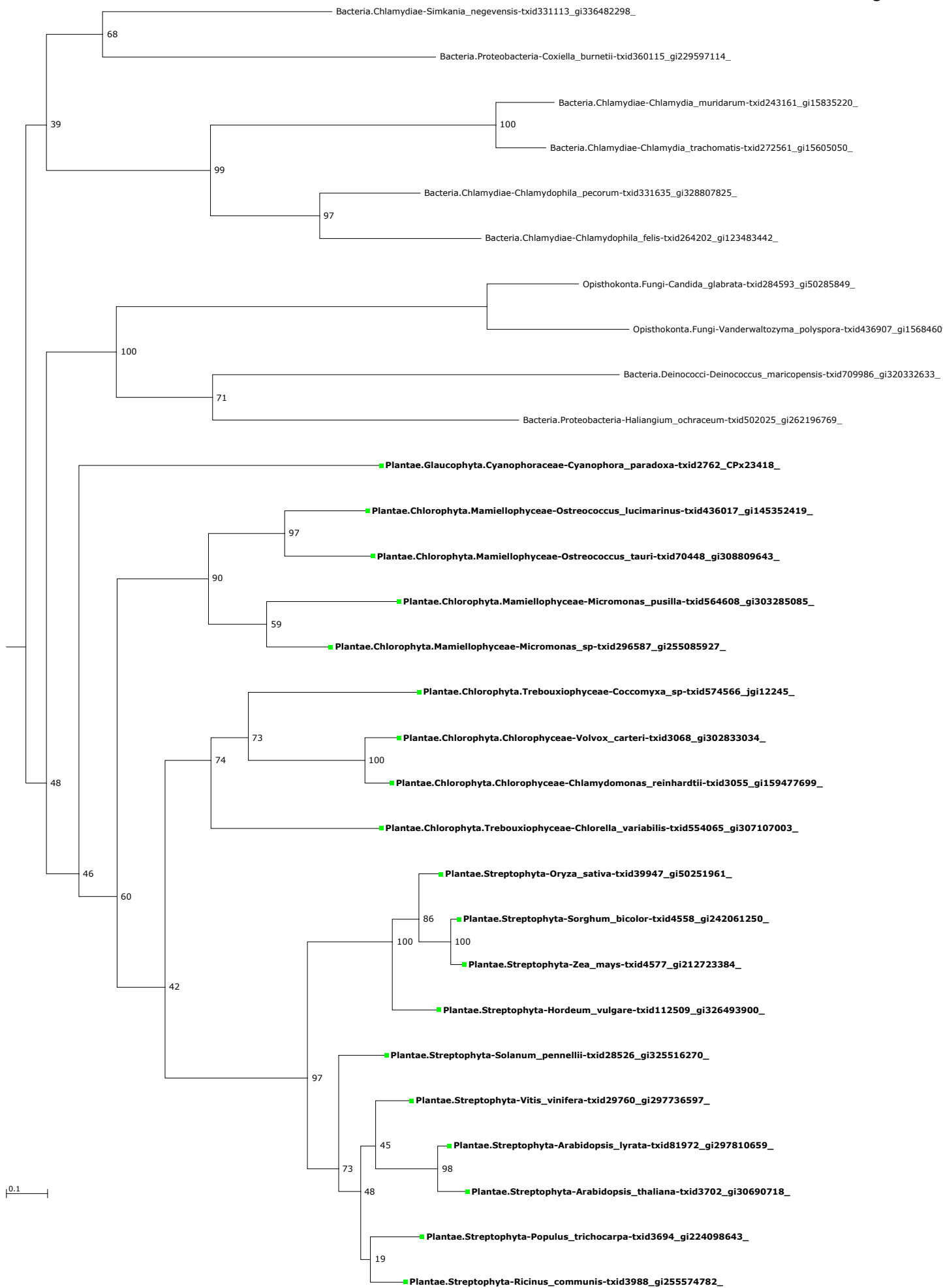


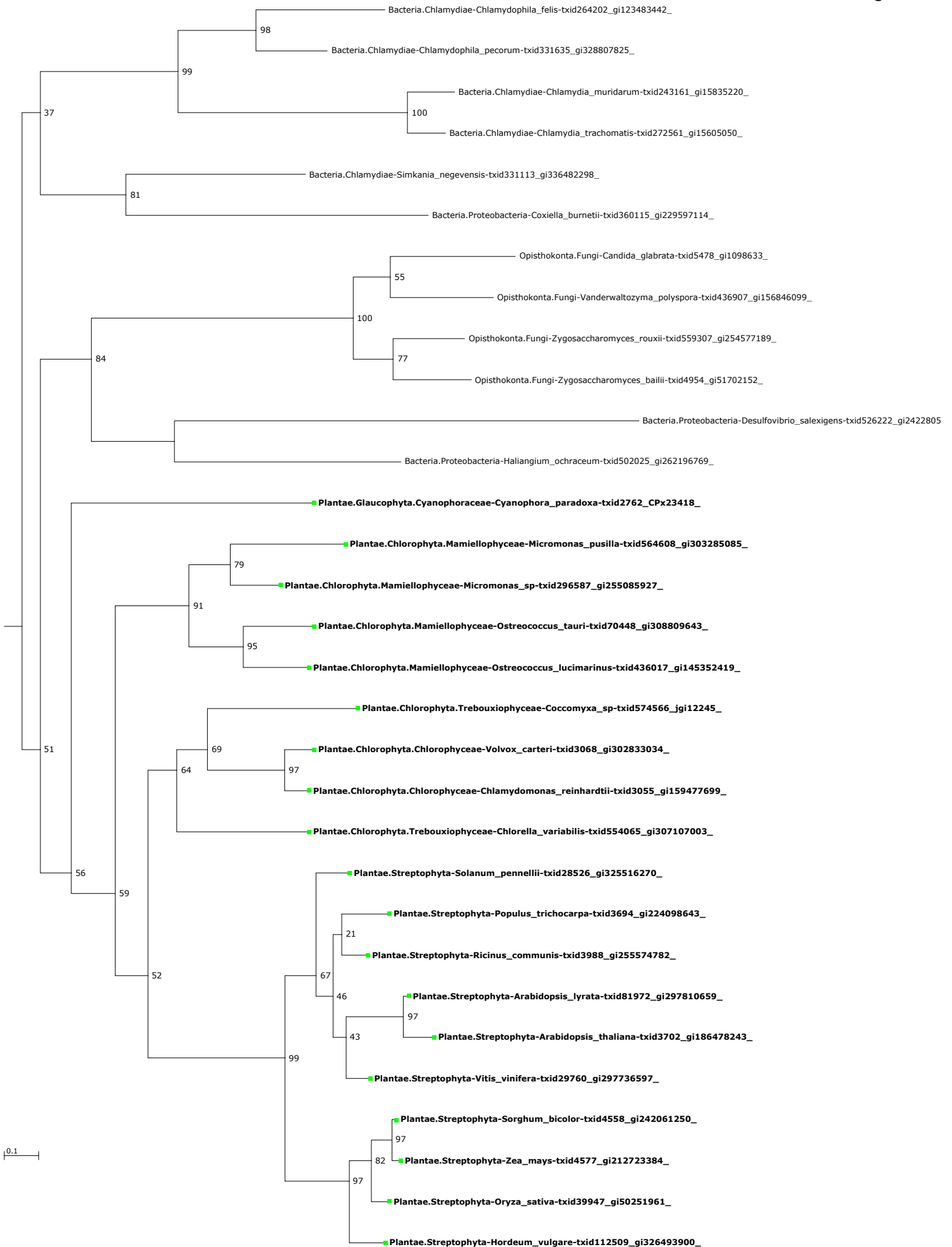


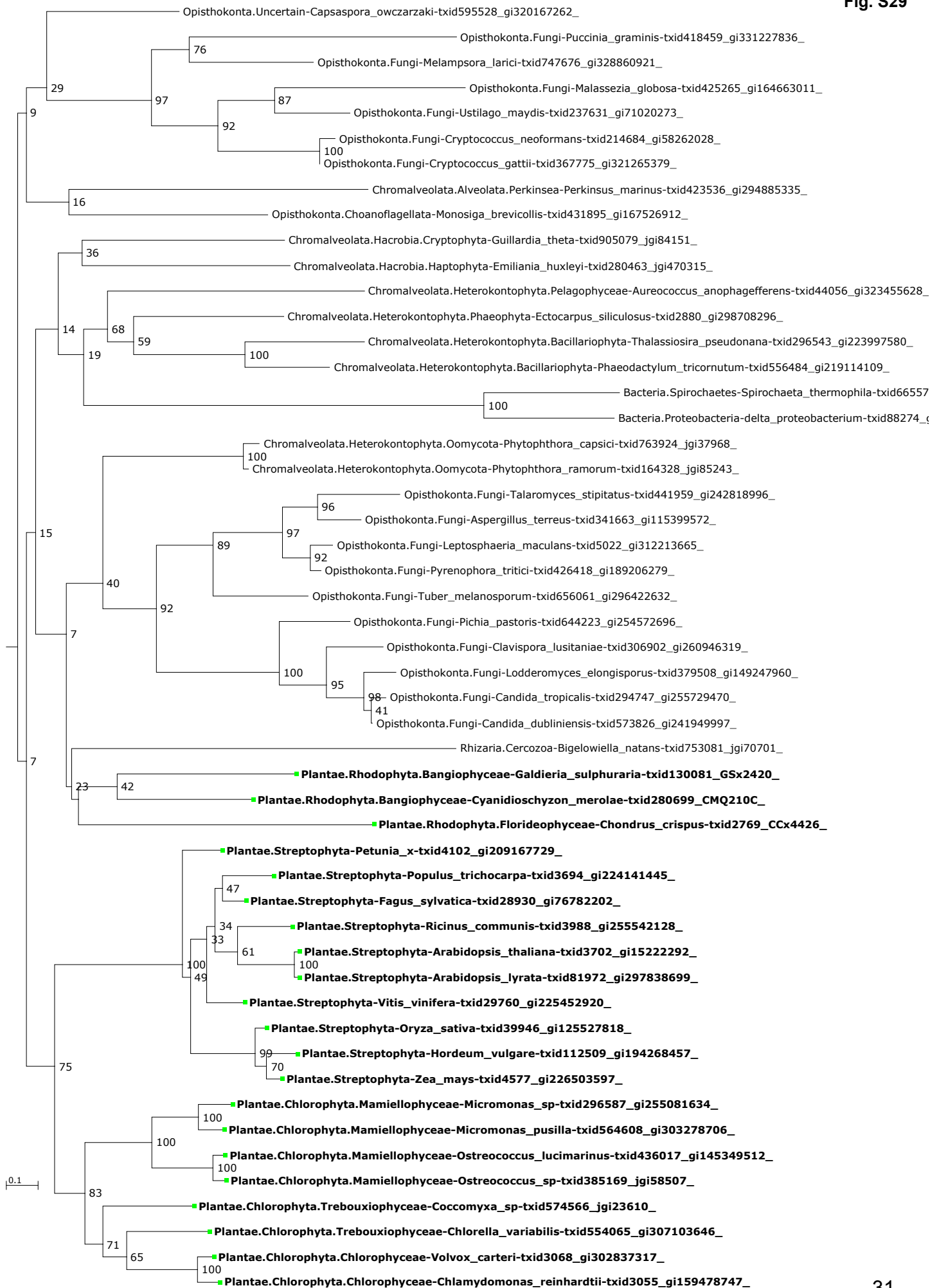


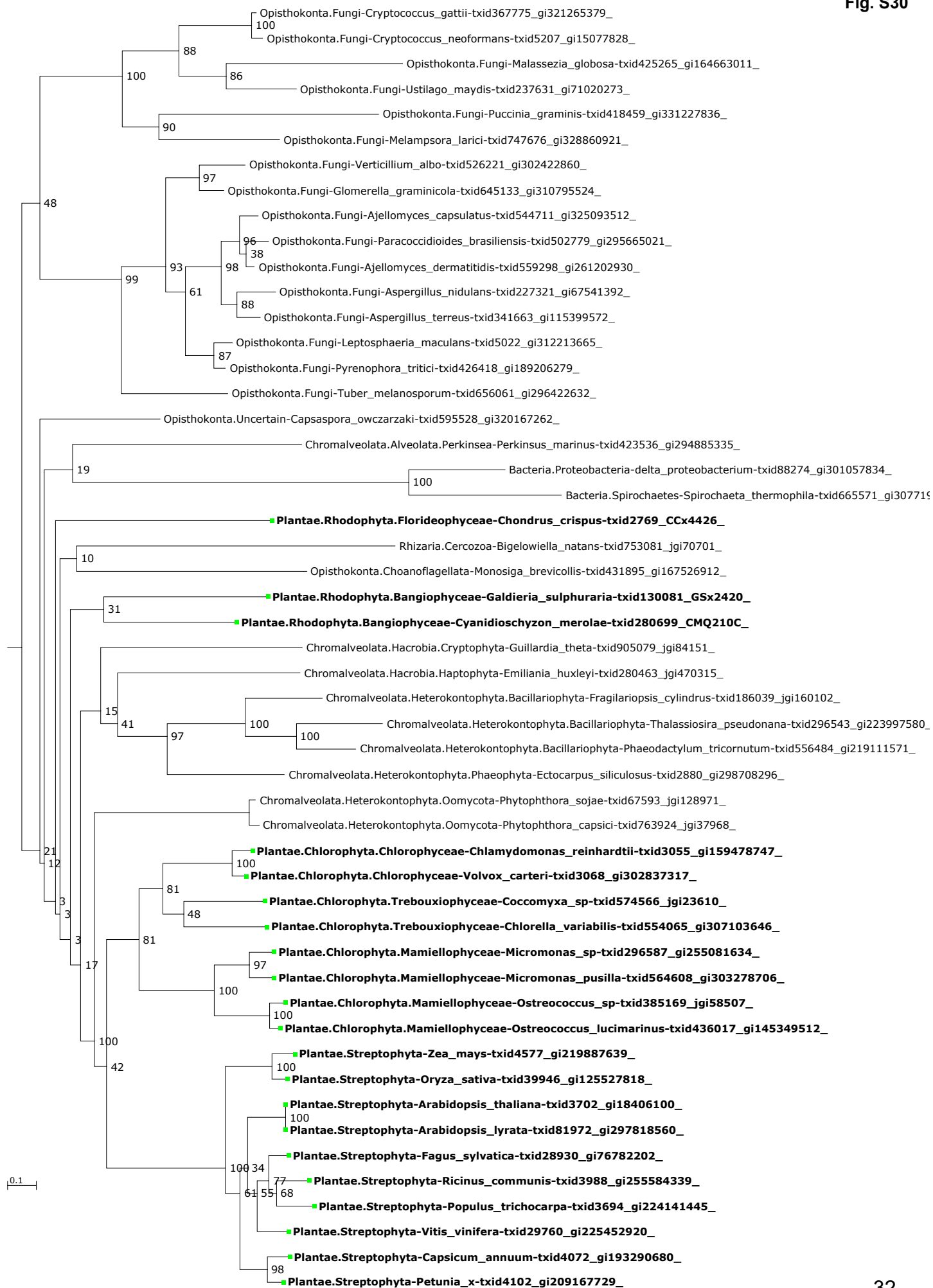


0.1

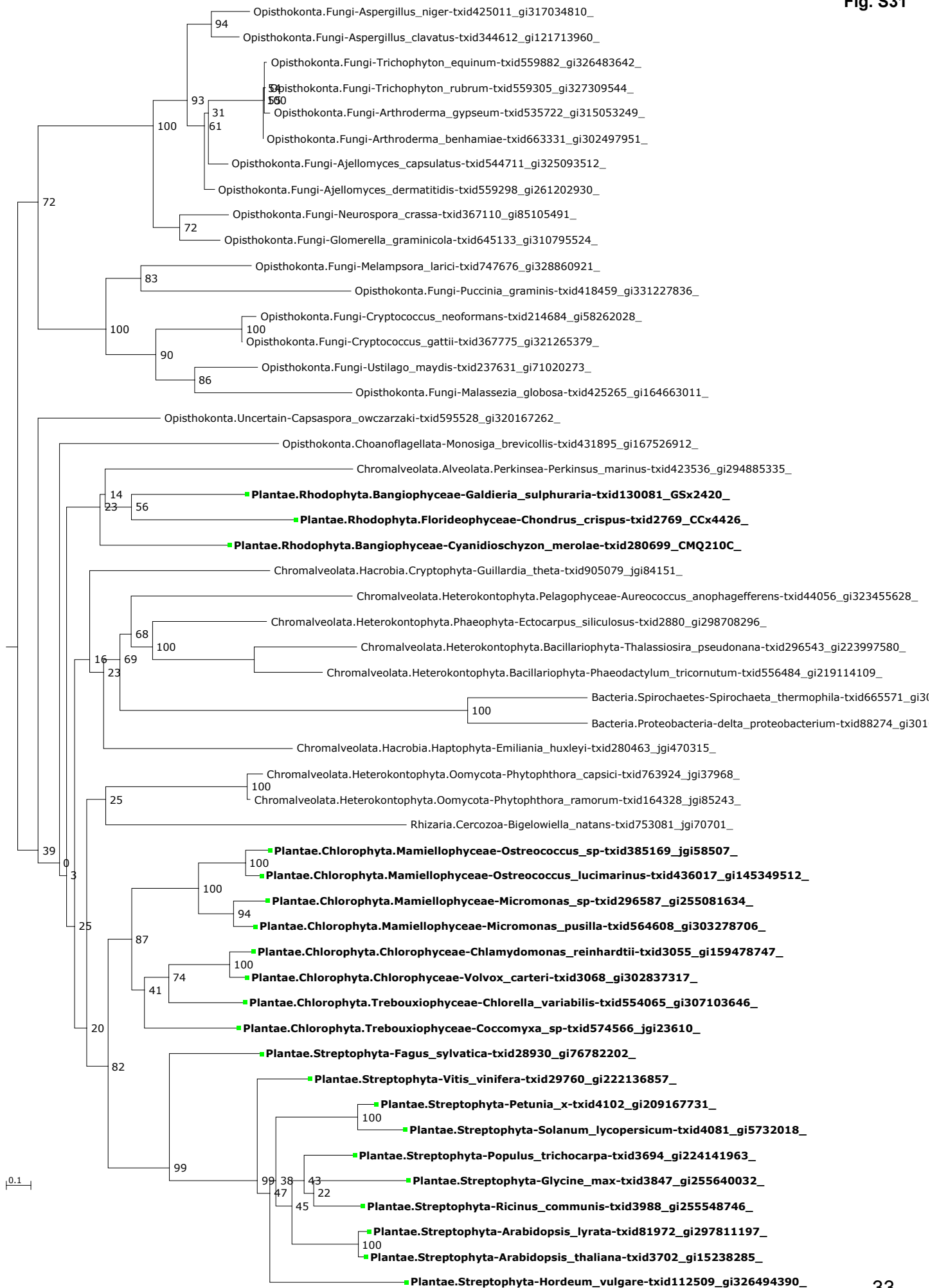






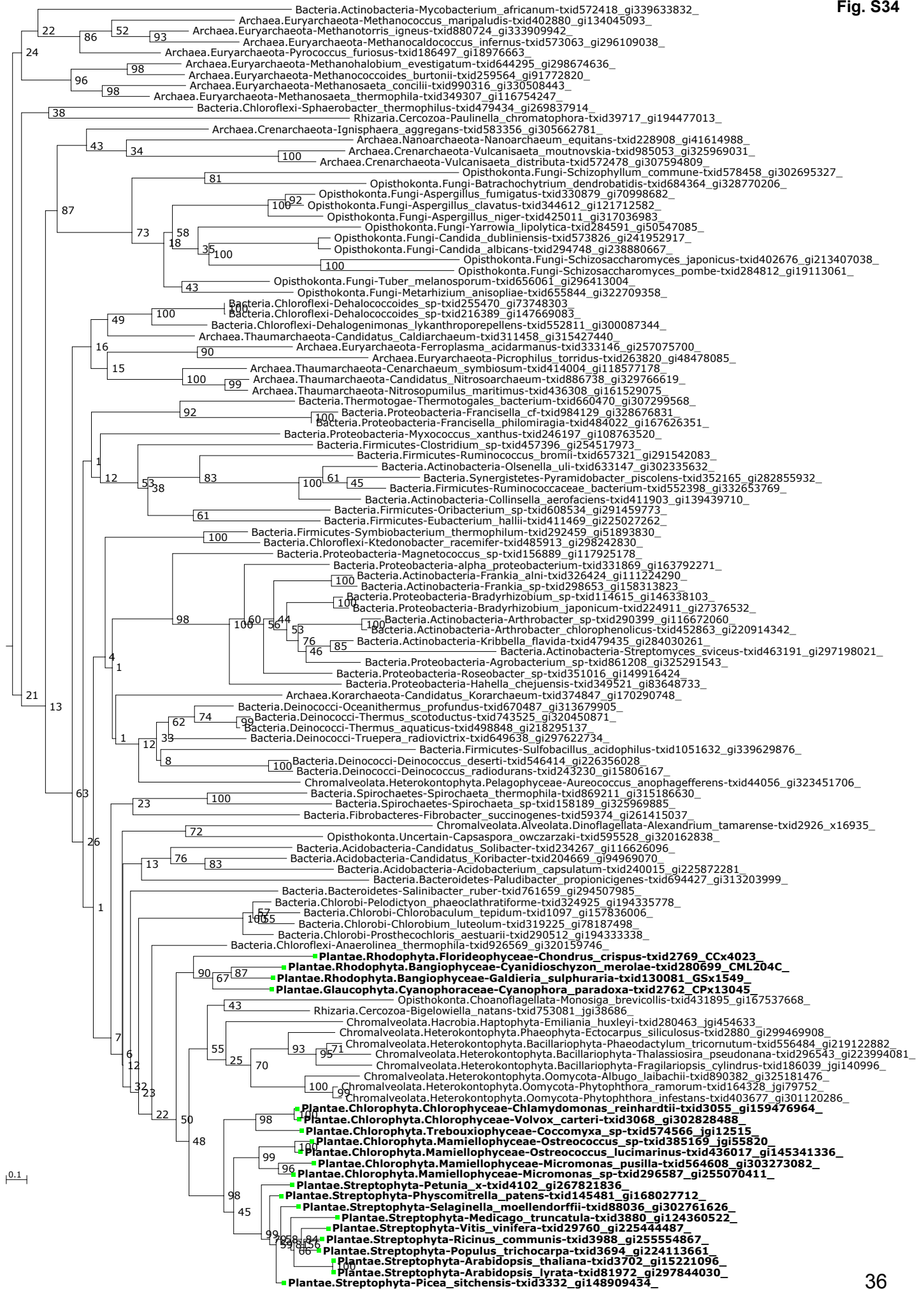


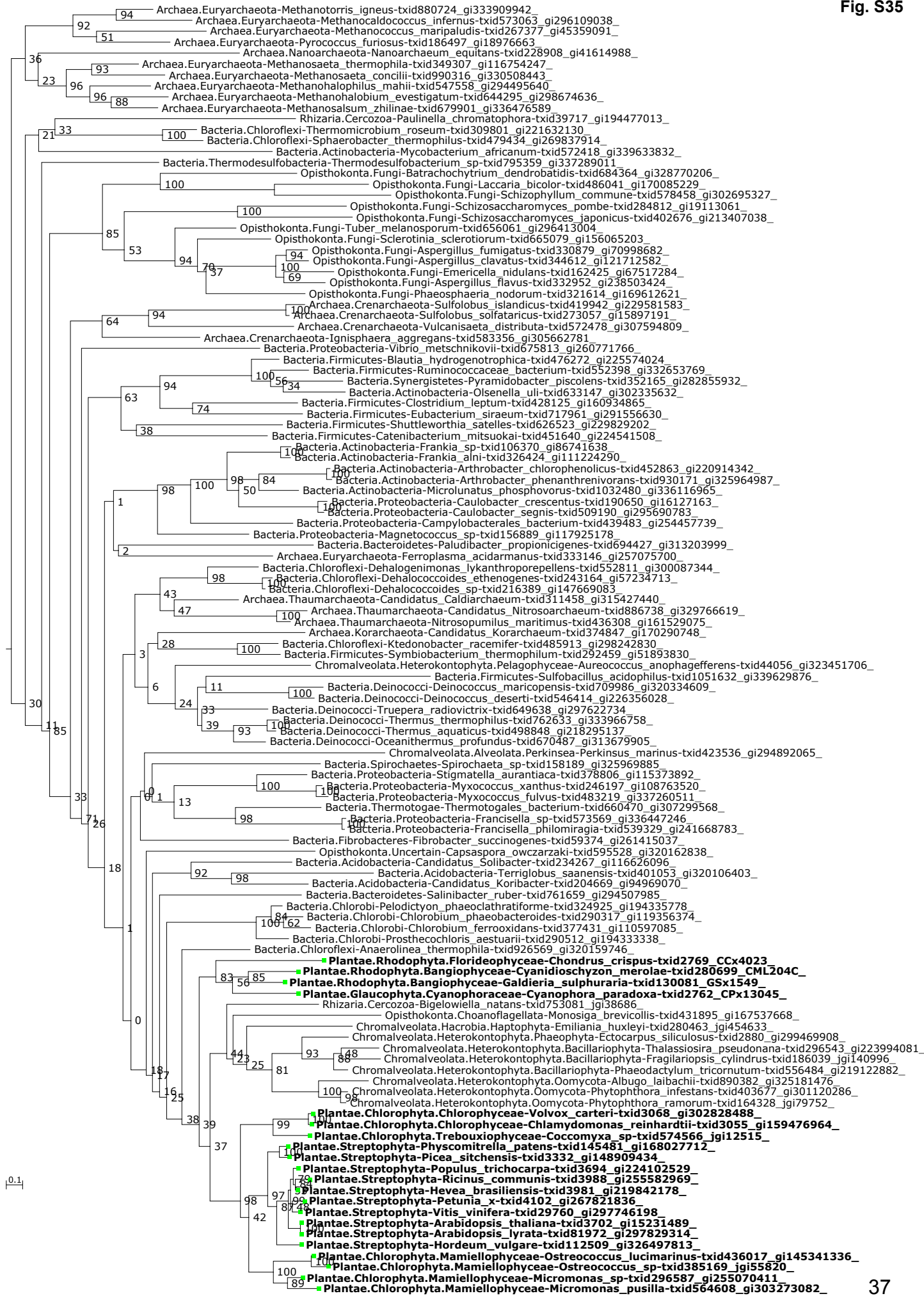


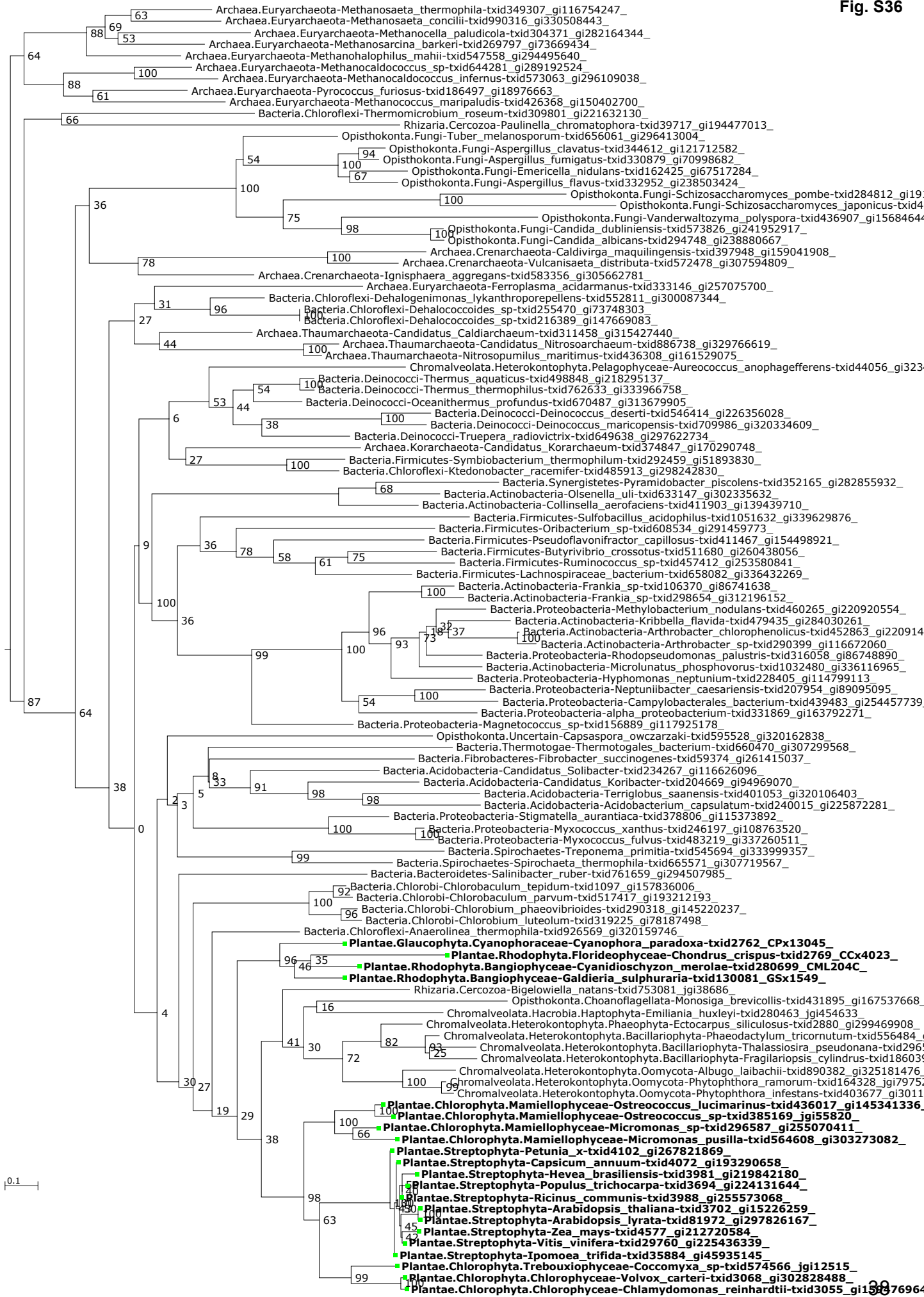


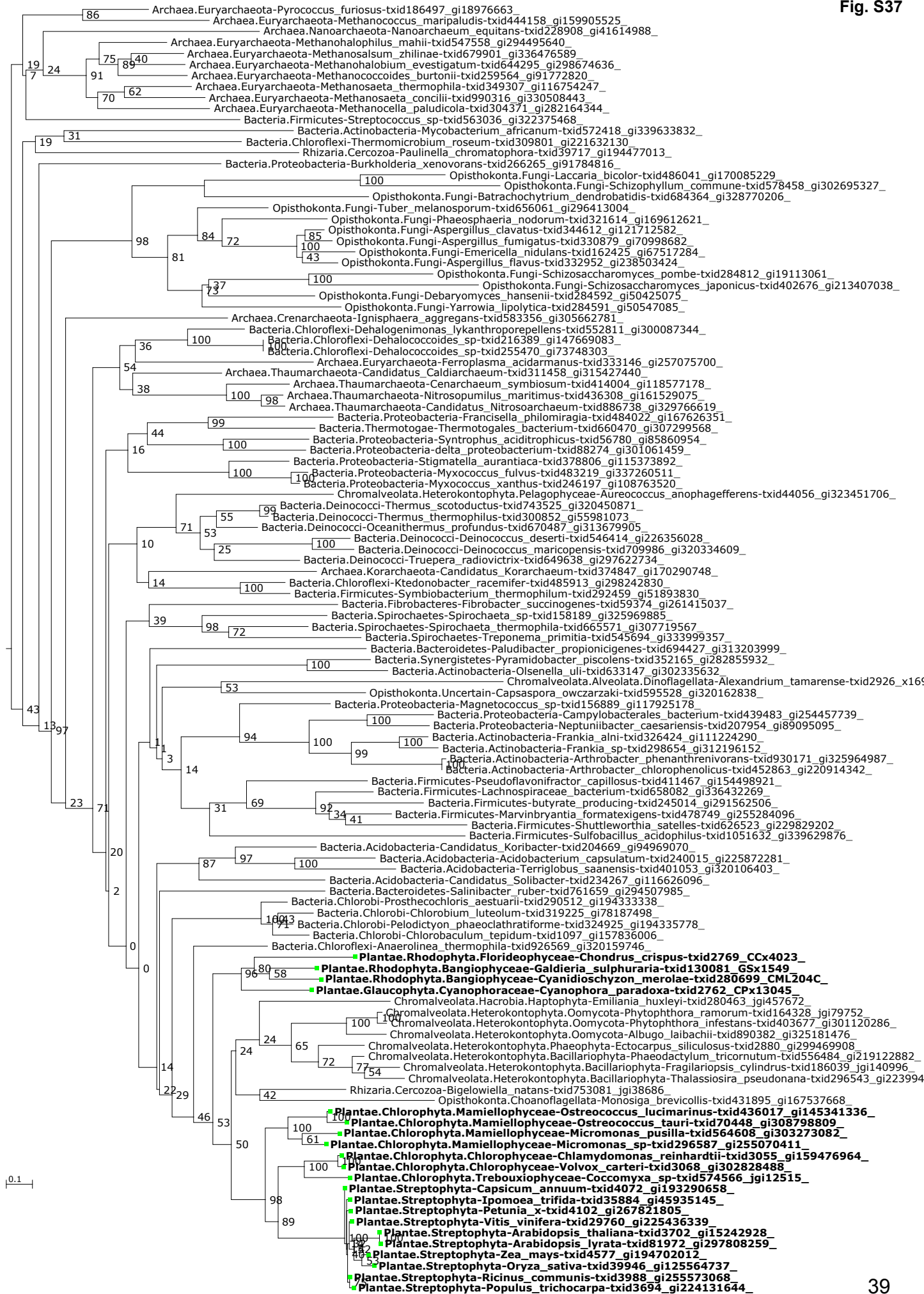


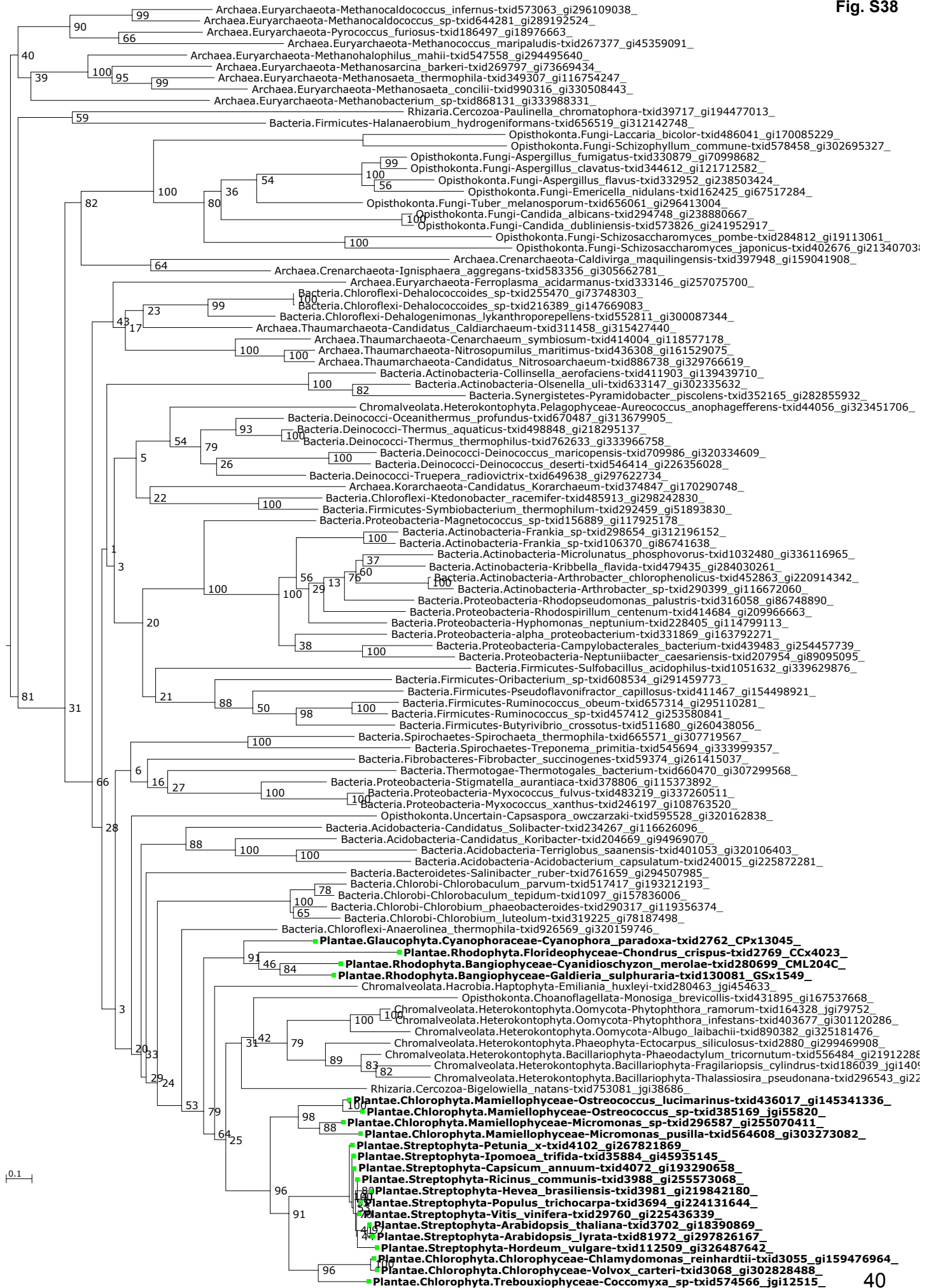




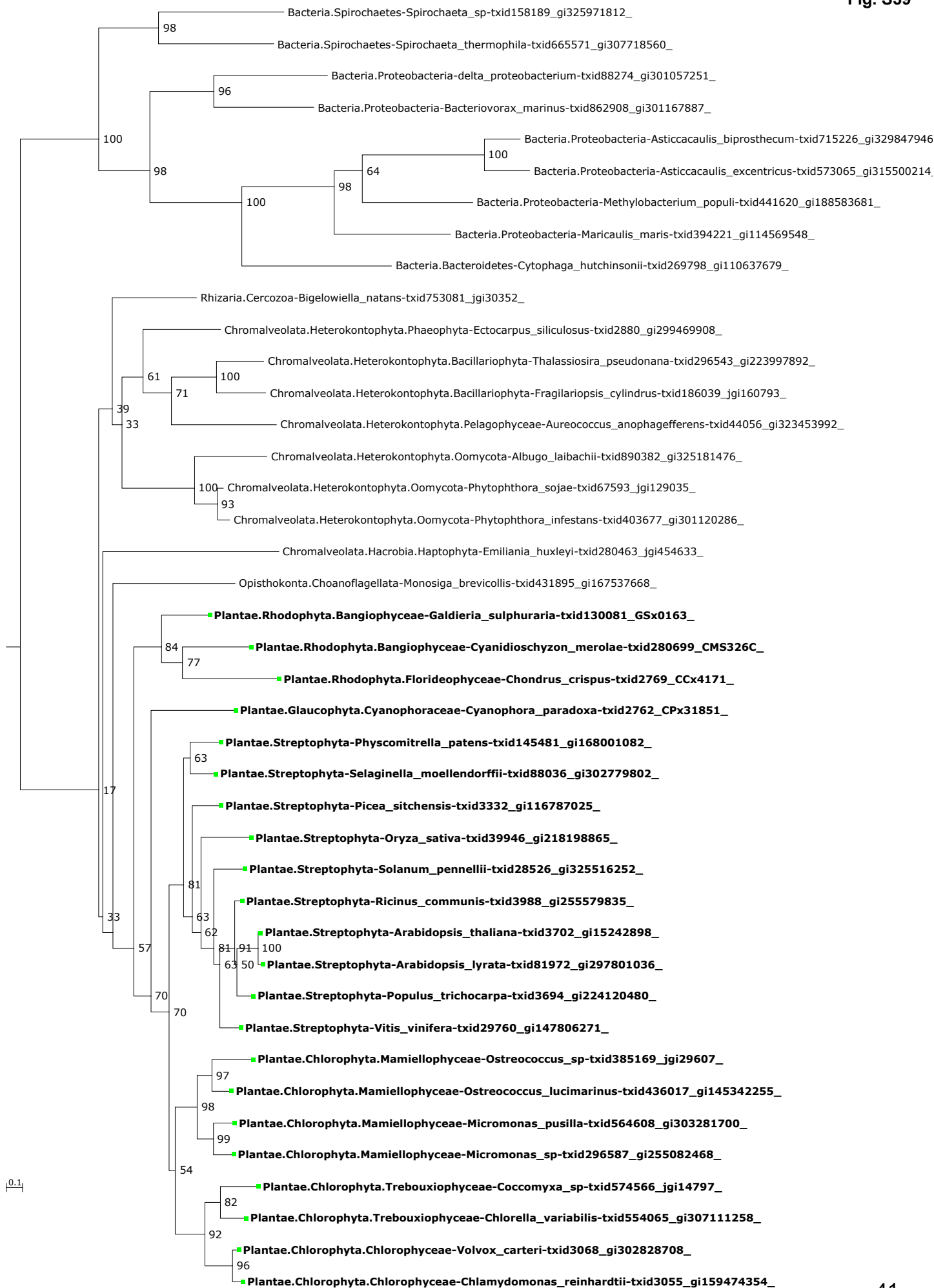




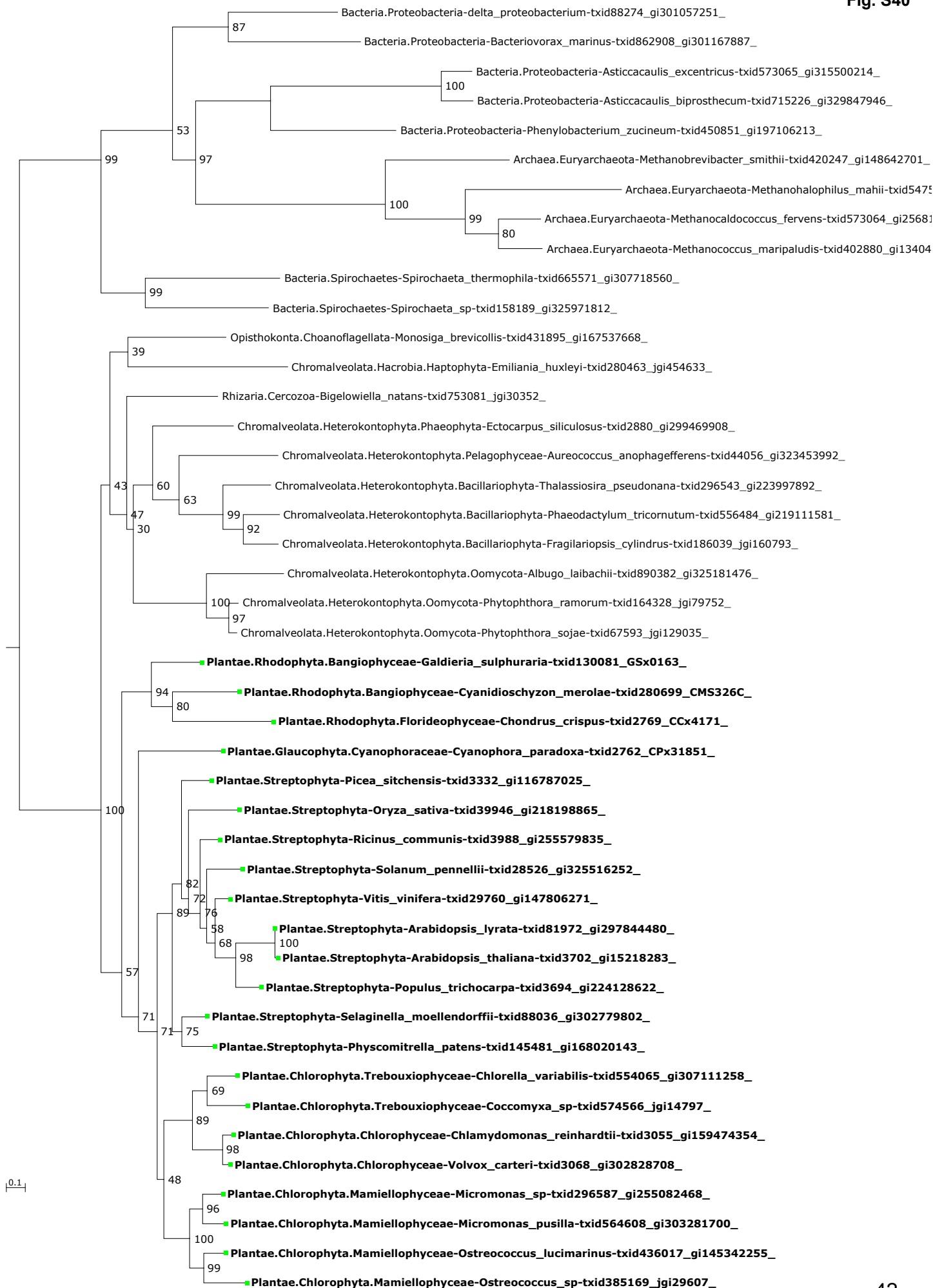


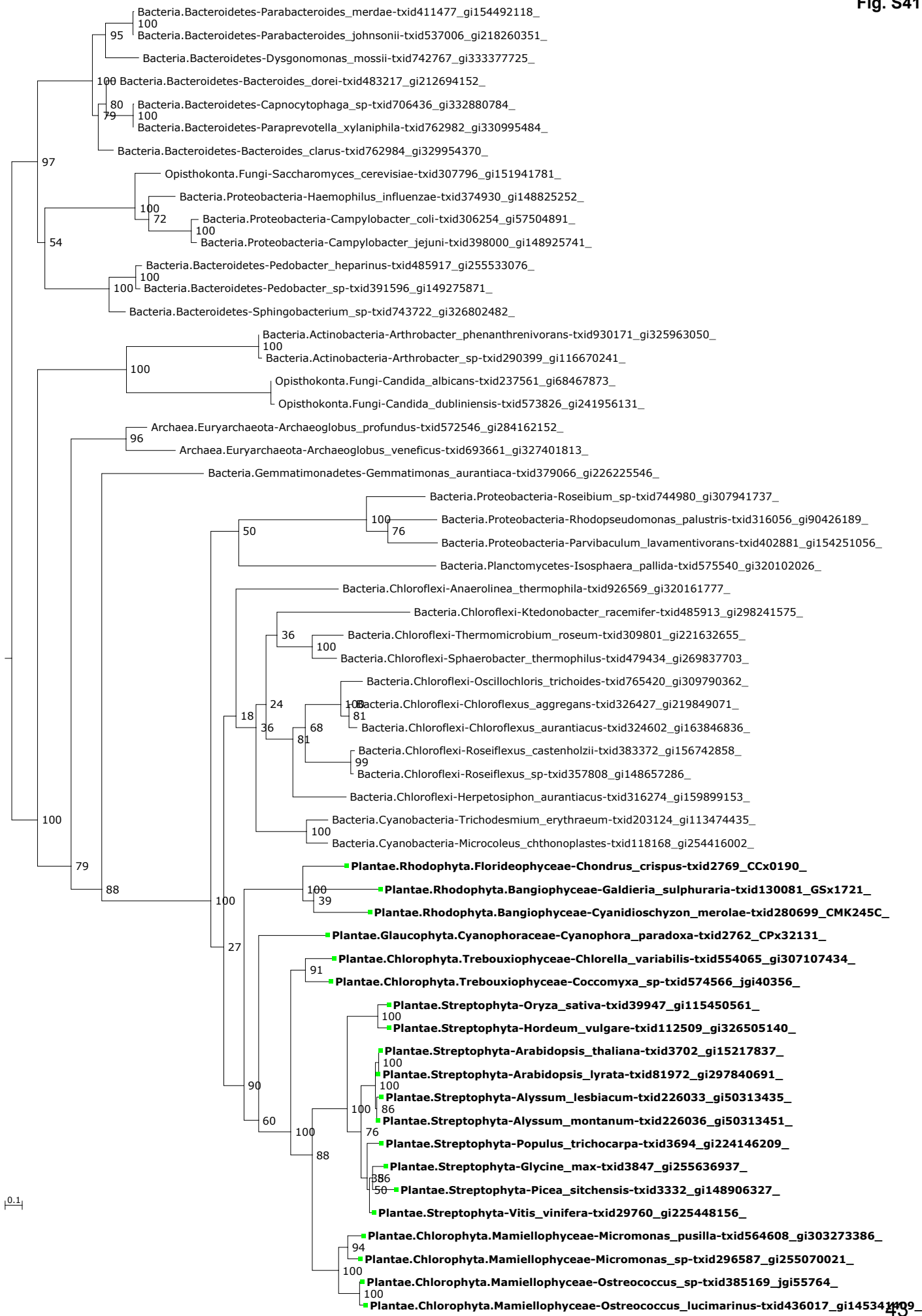




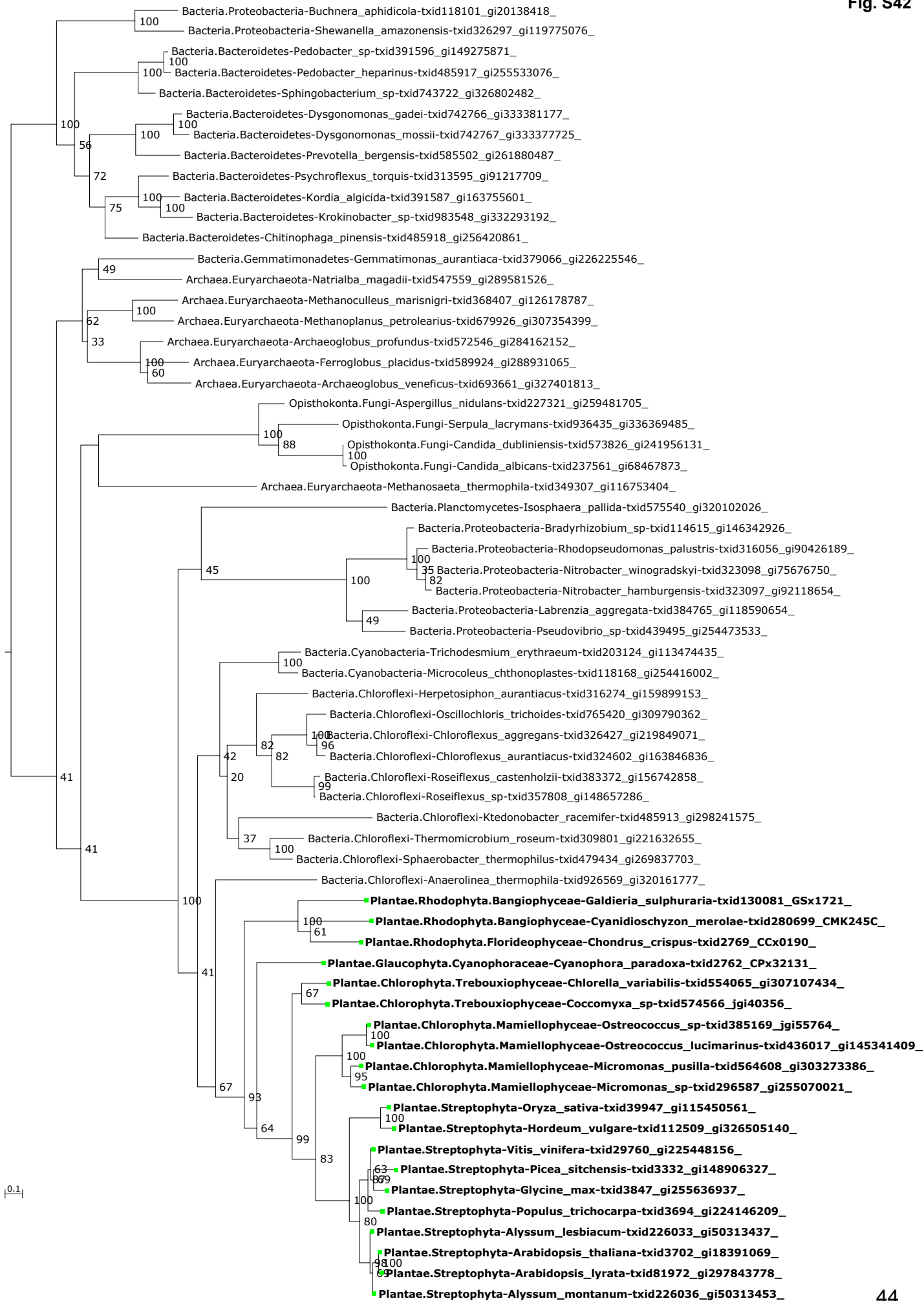


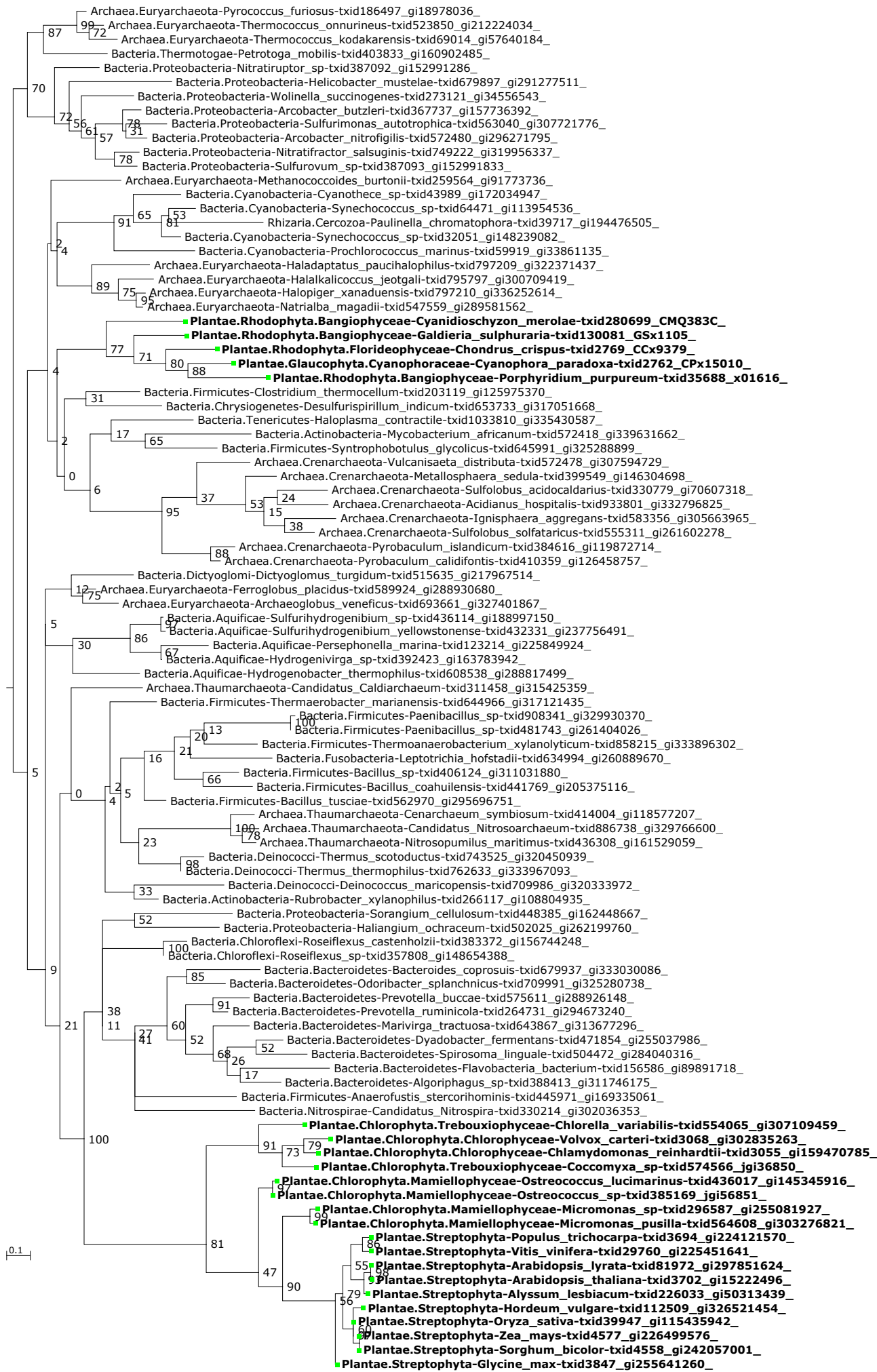
0.1

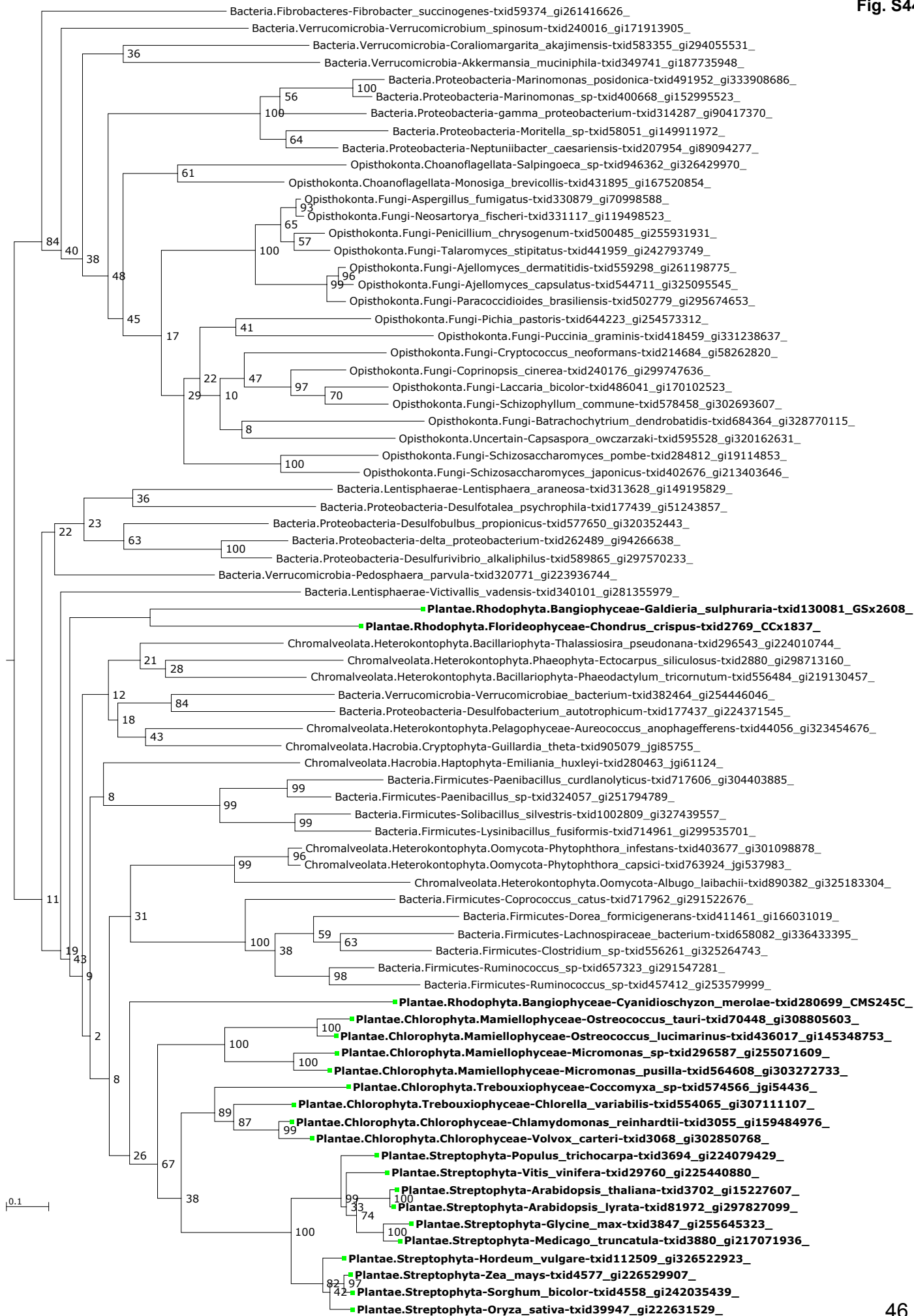


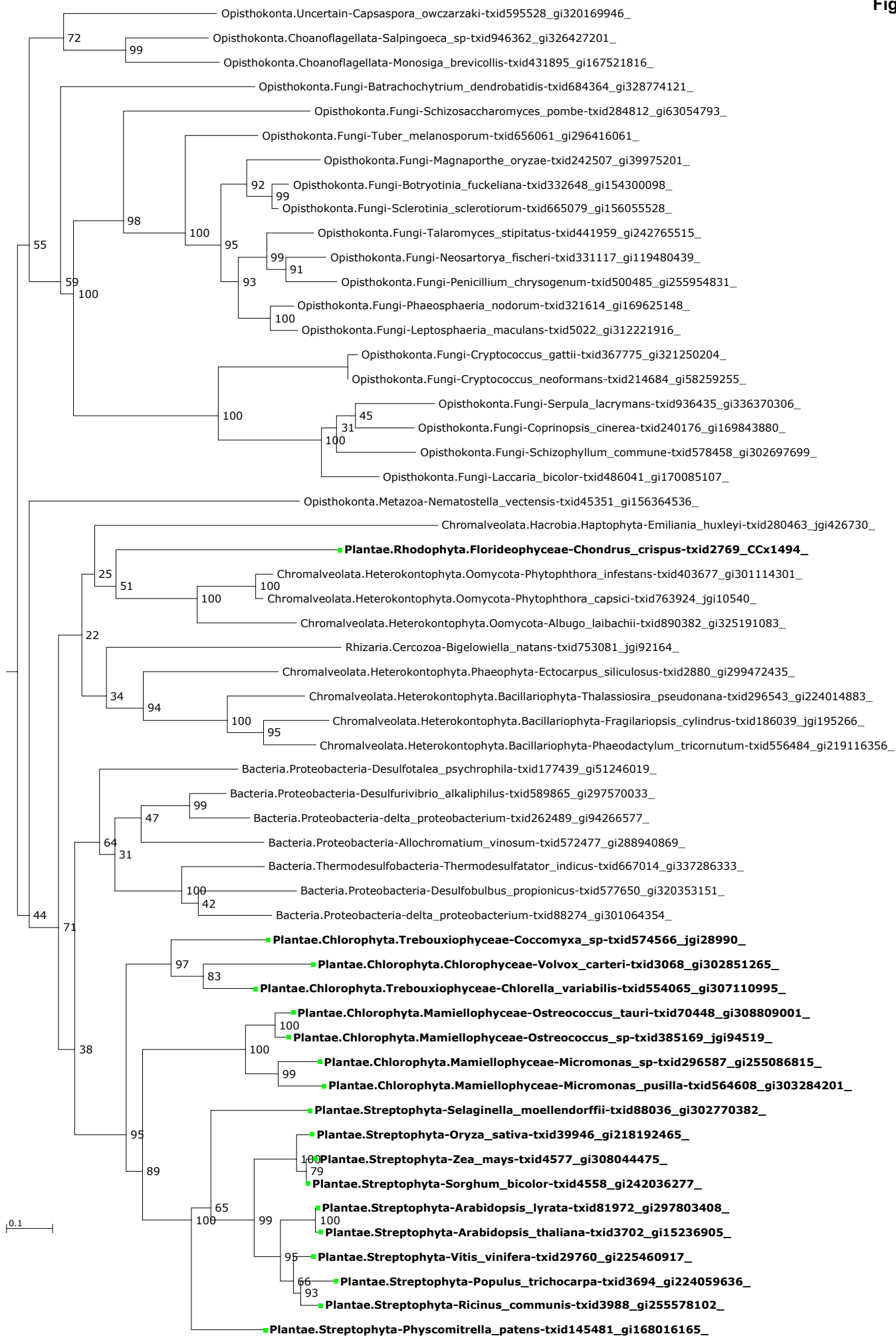


0.1





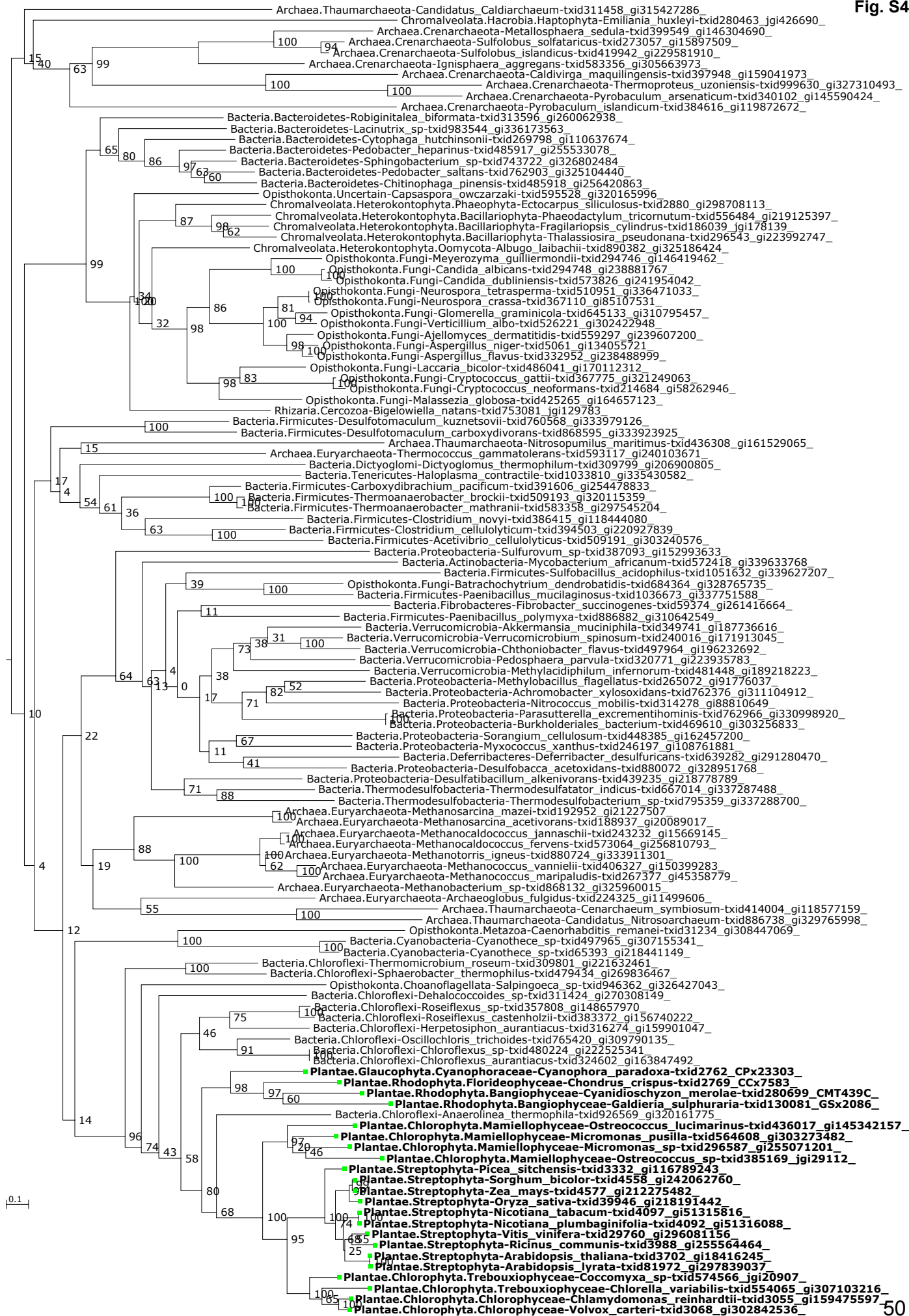


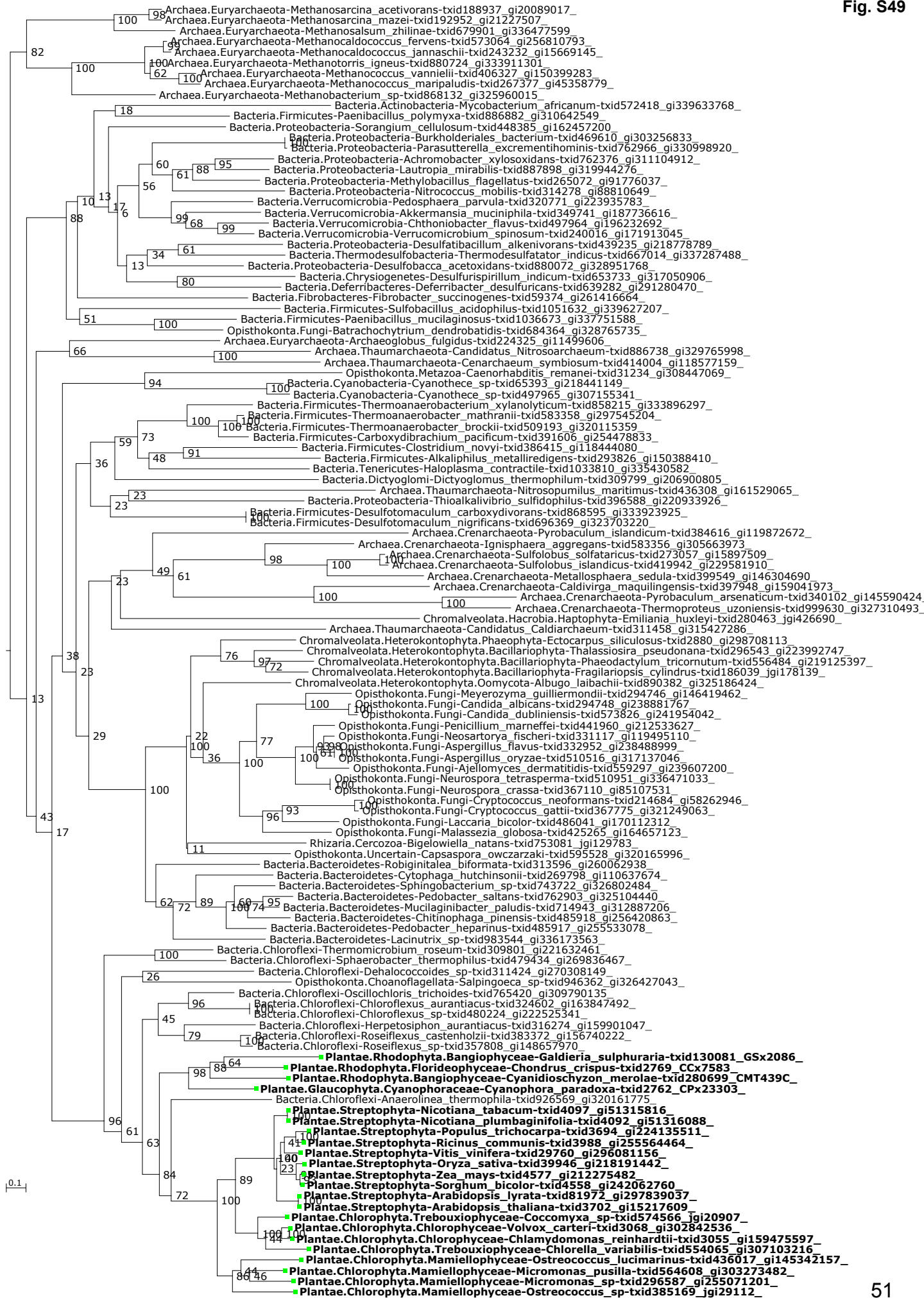




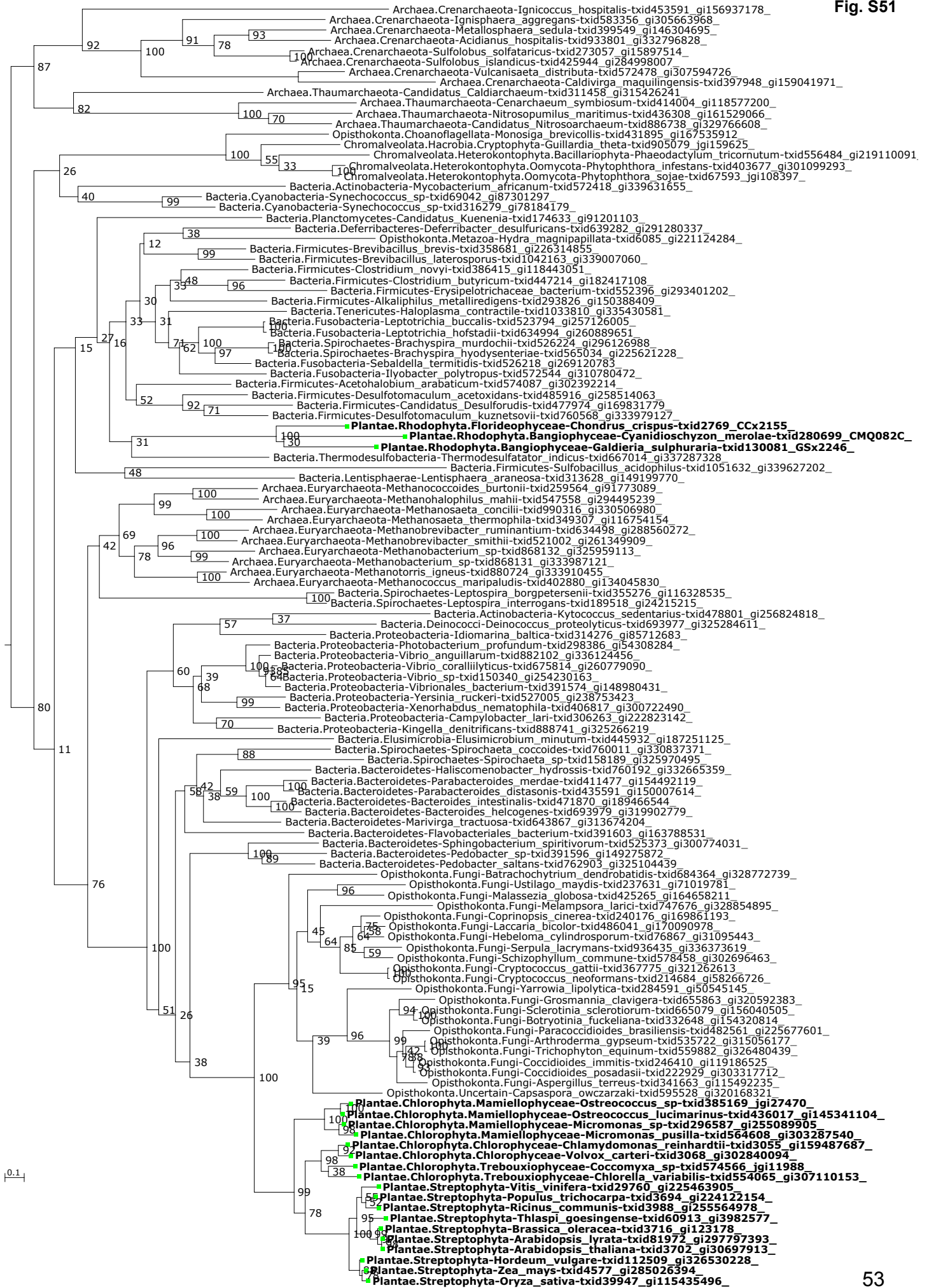


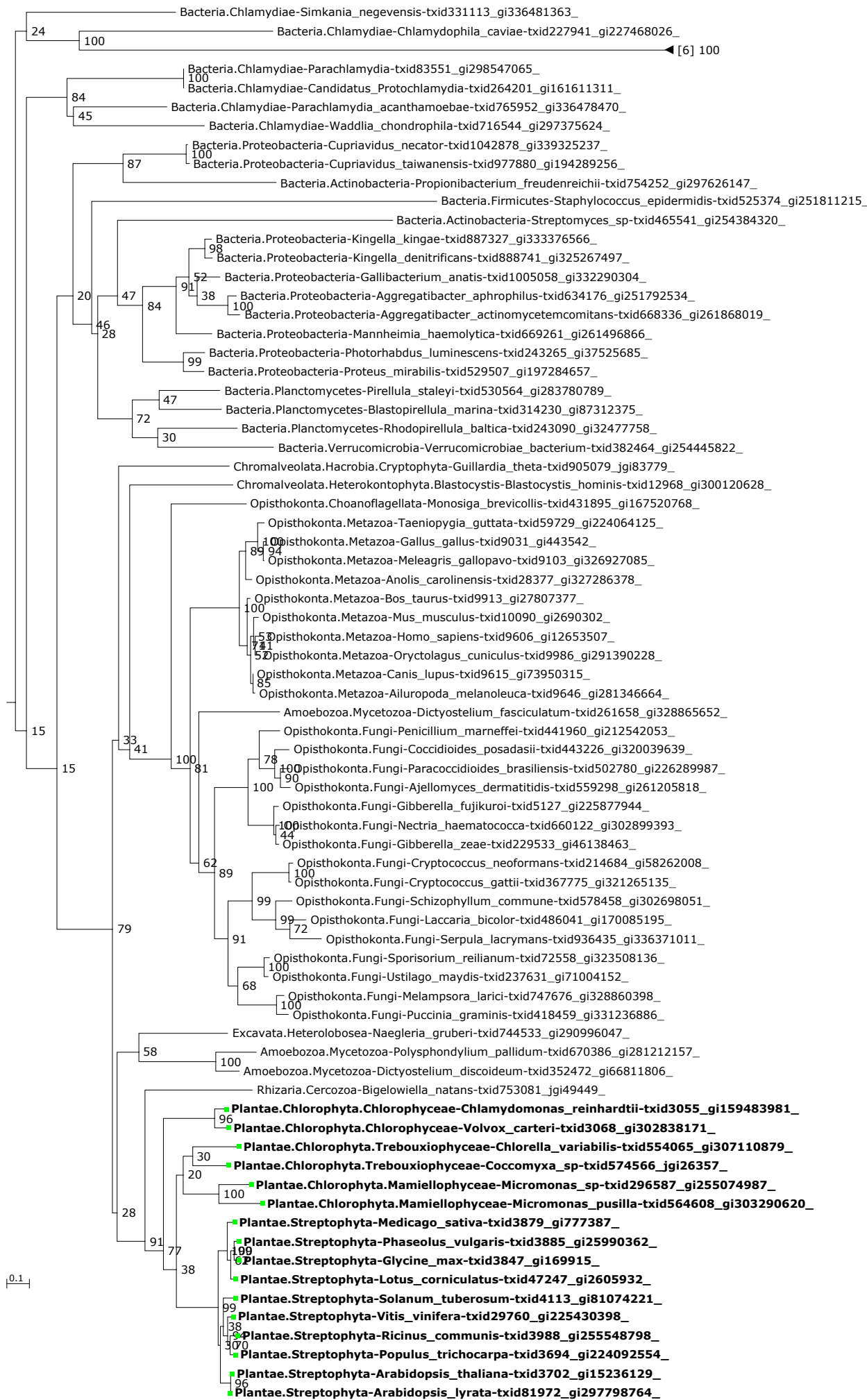


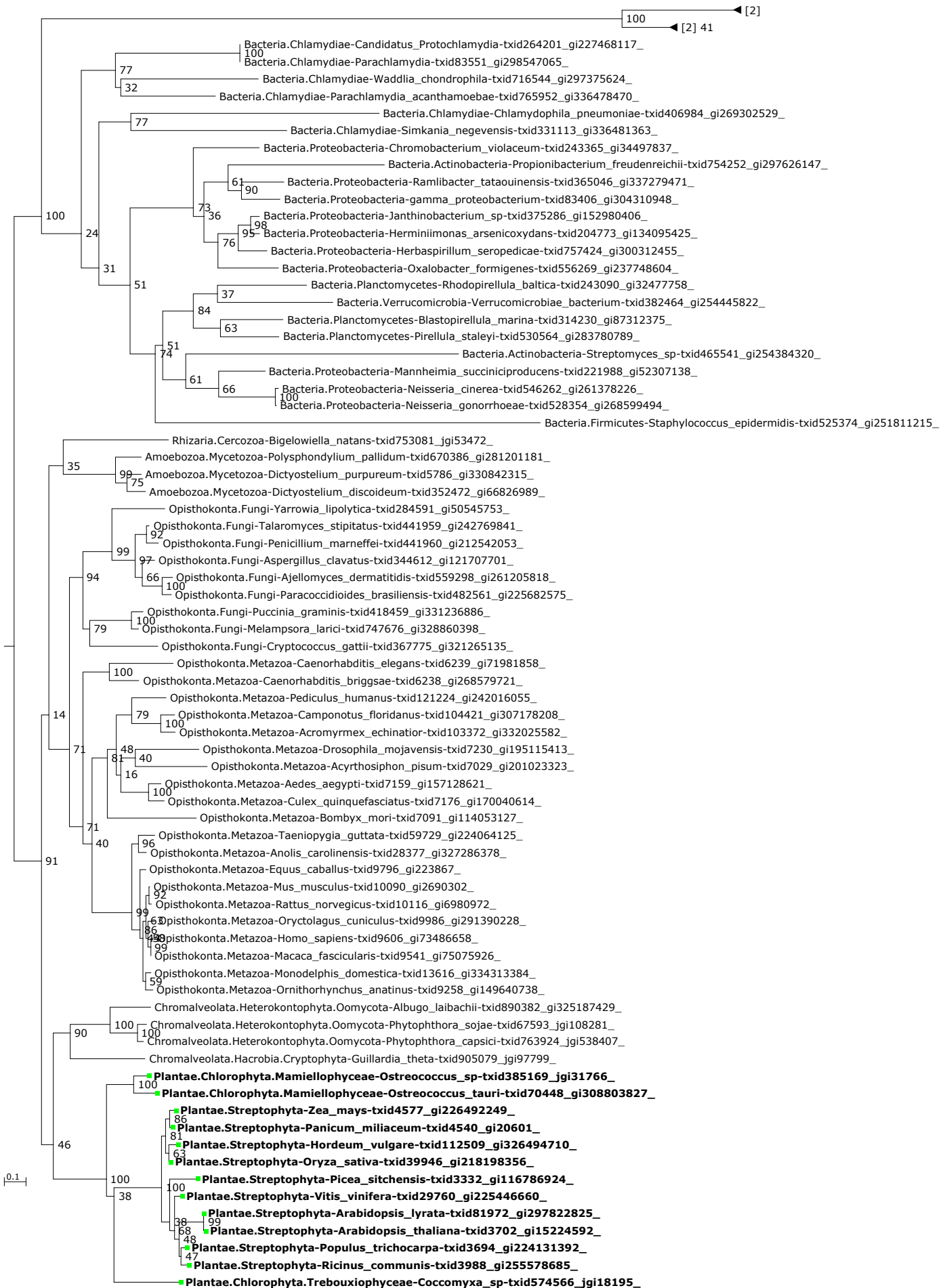


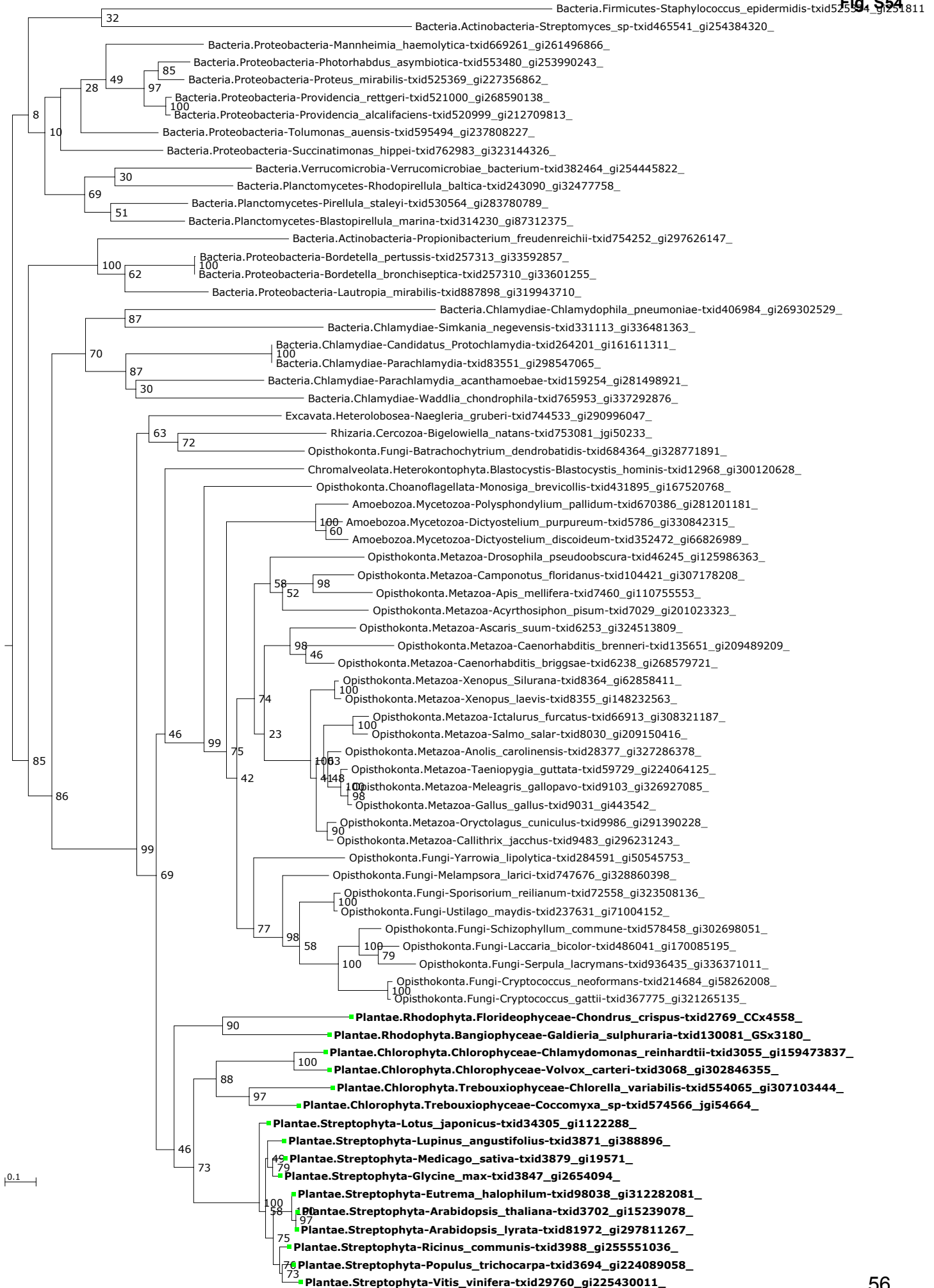






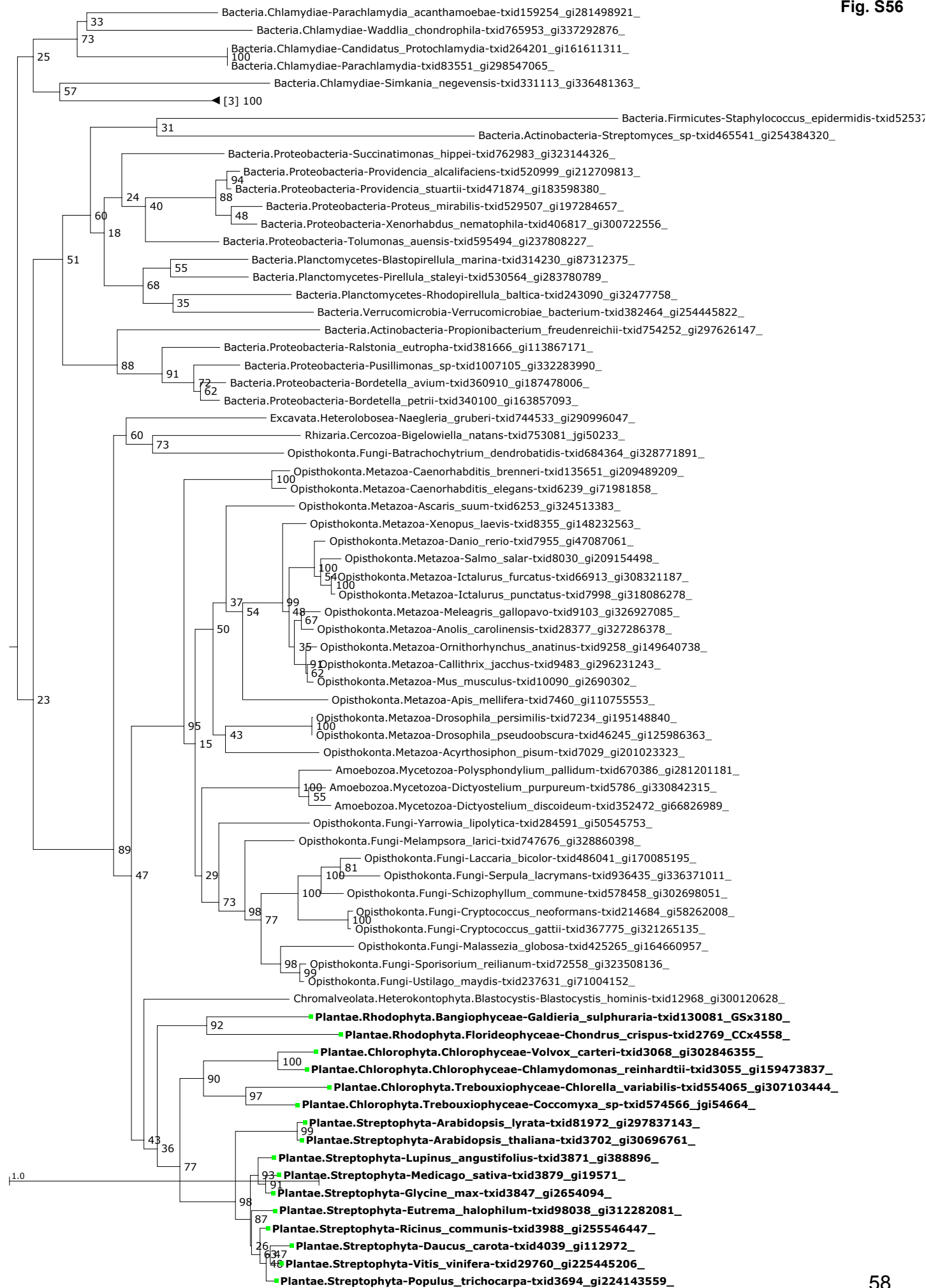


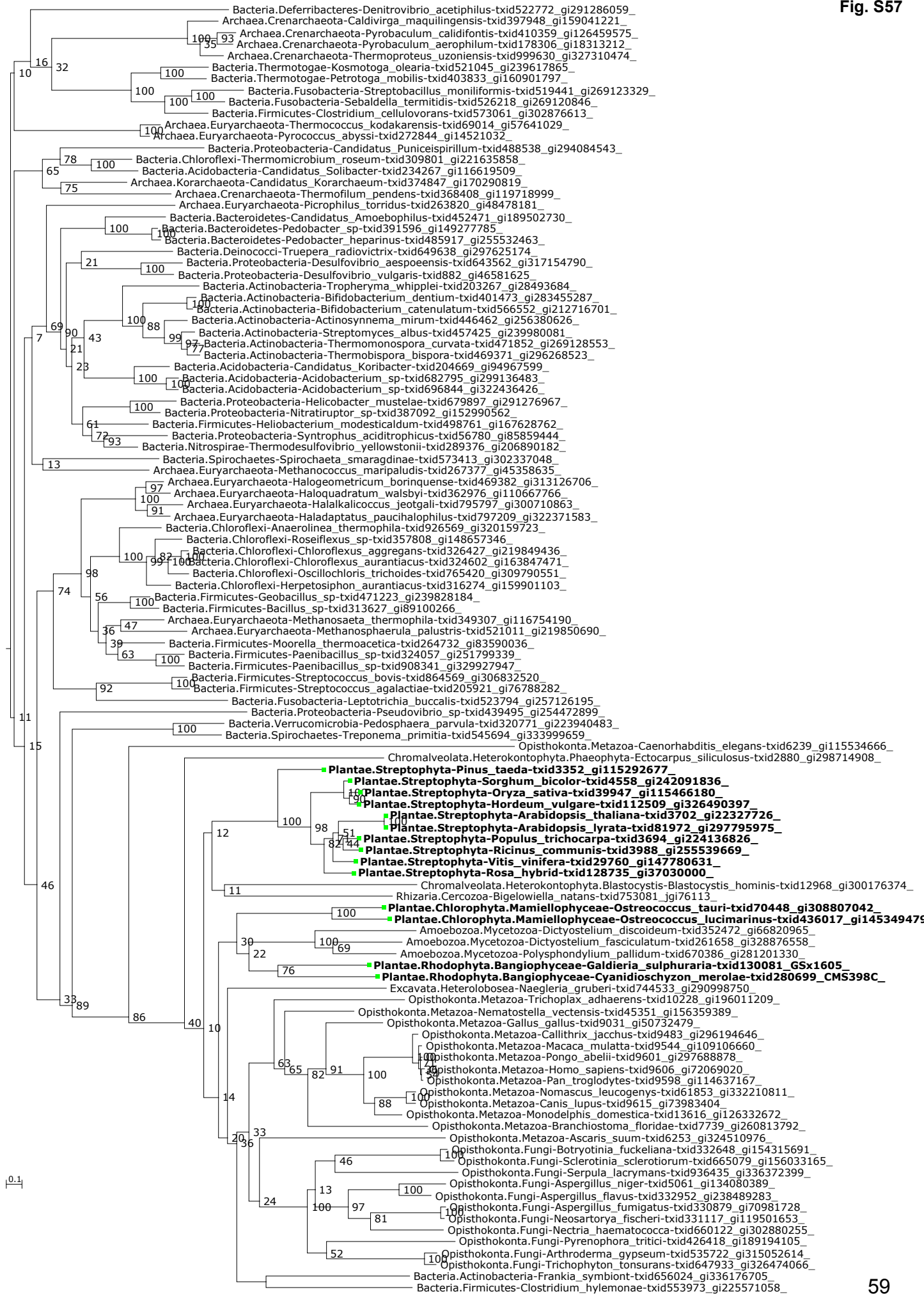


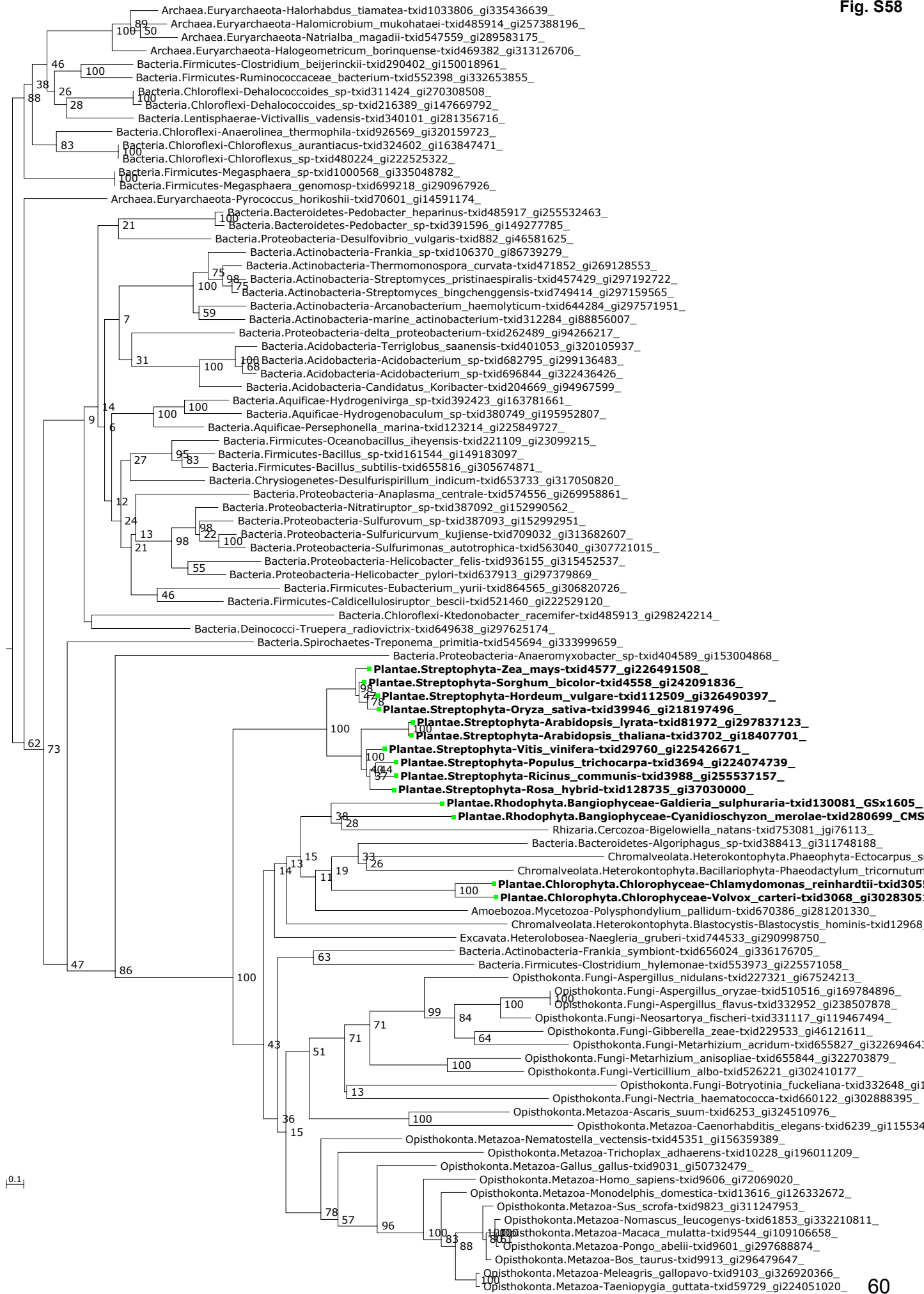




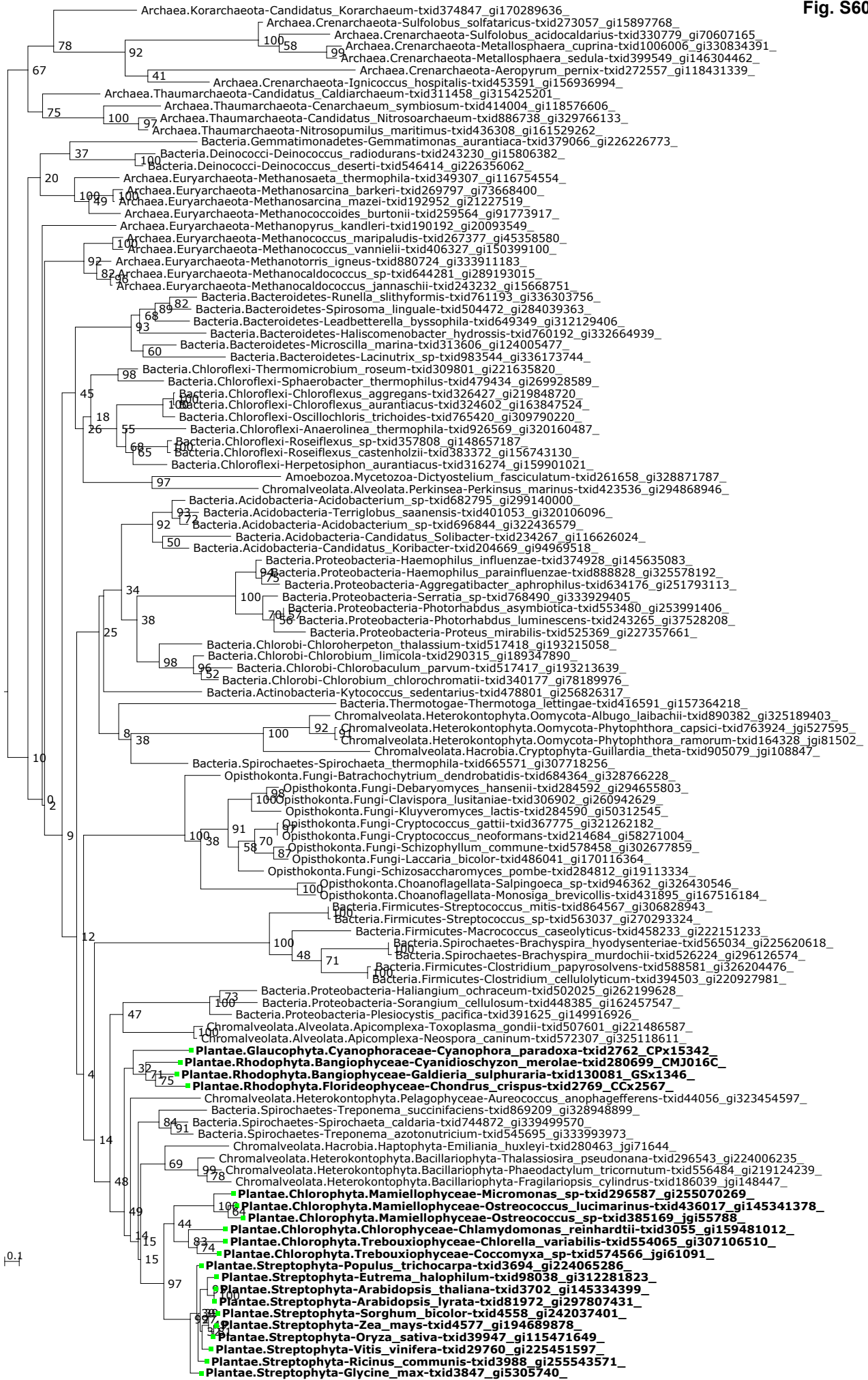


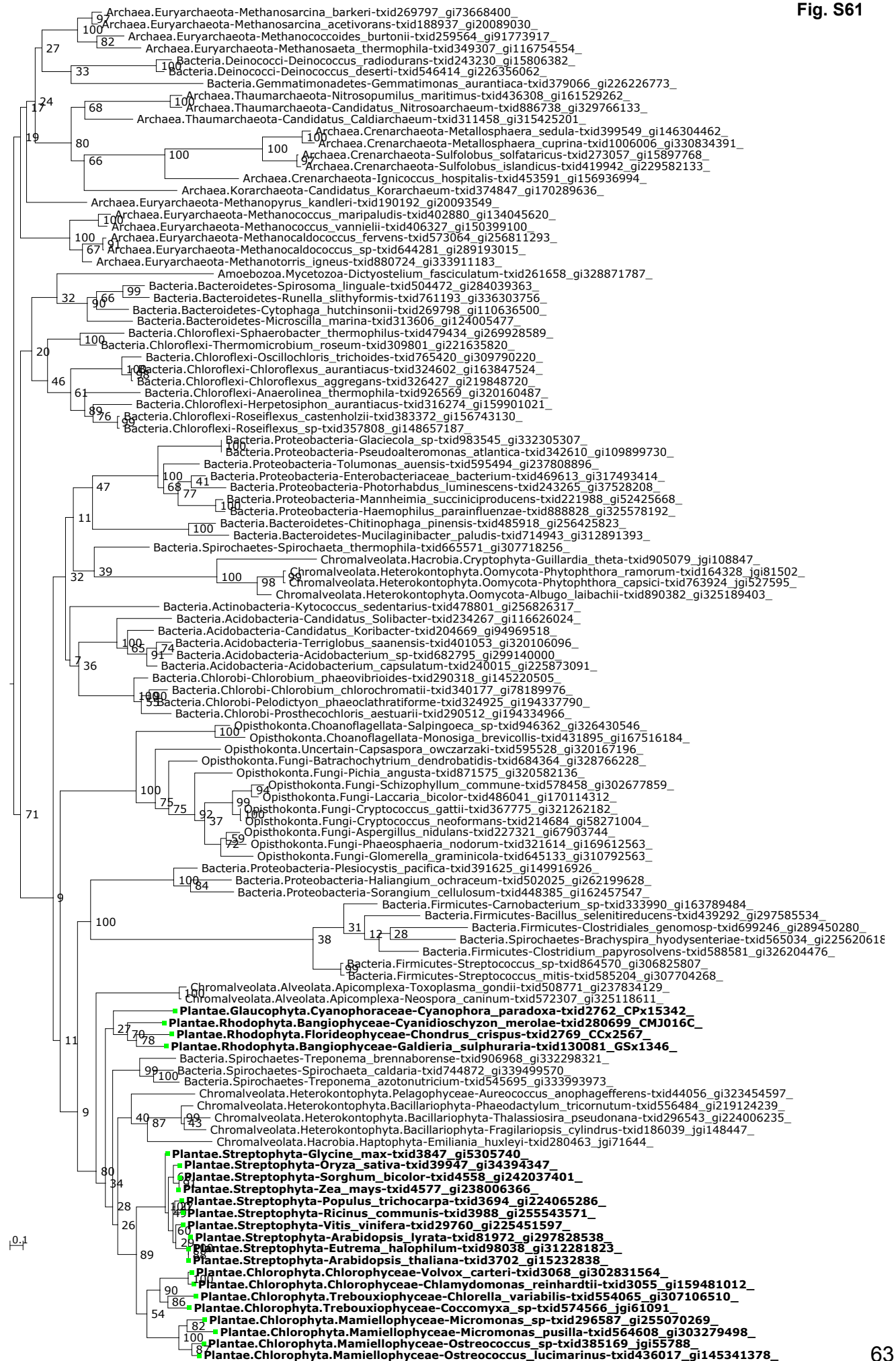


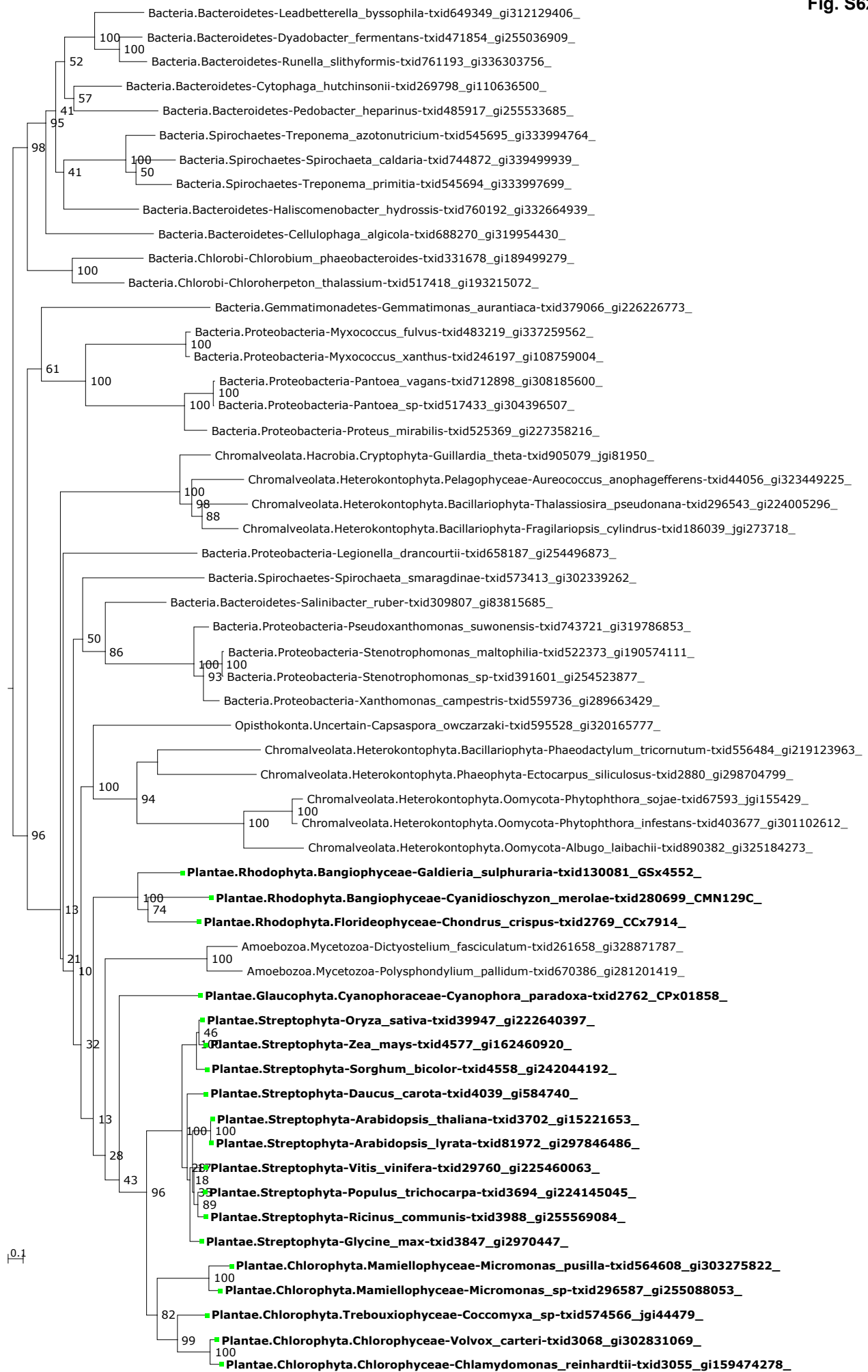




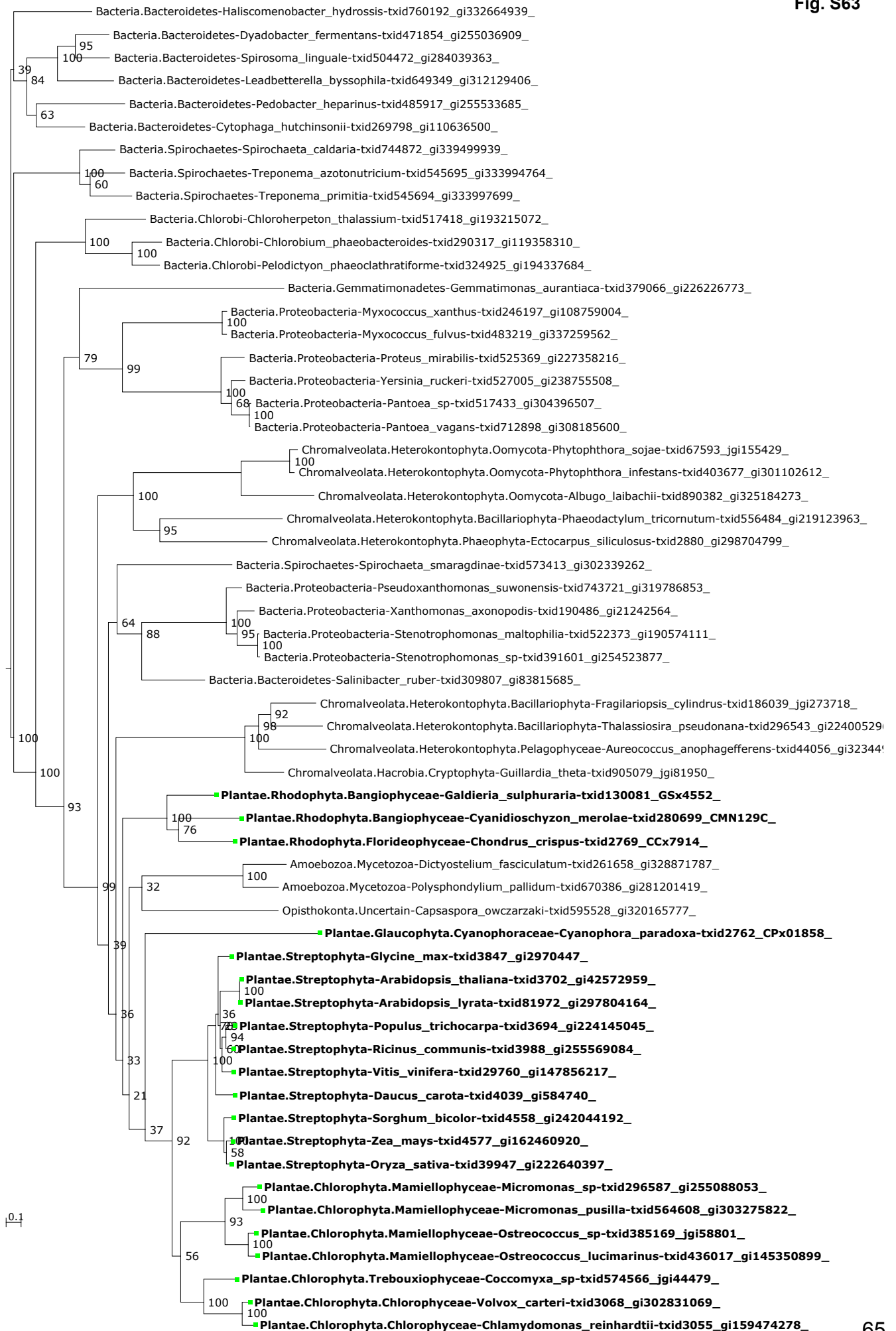


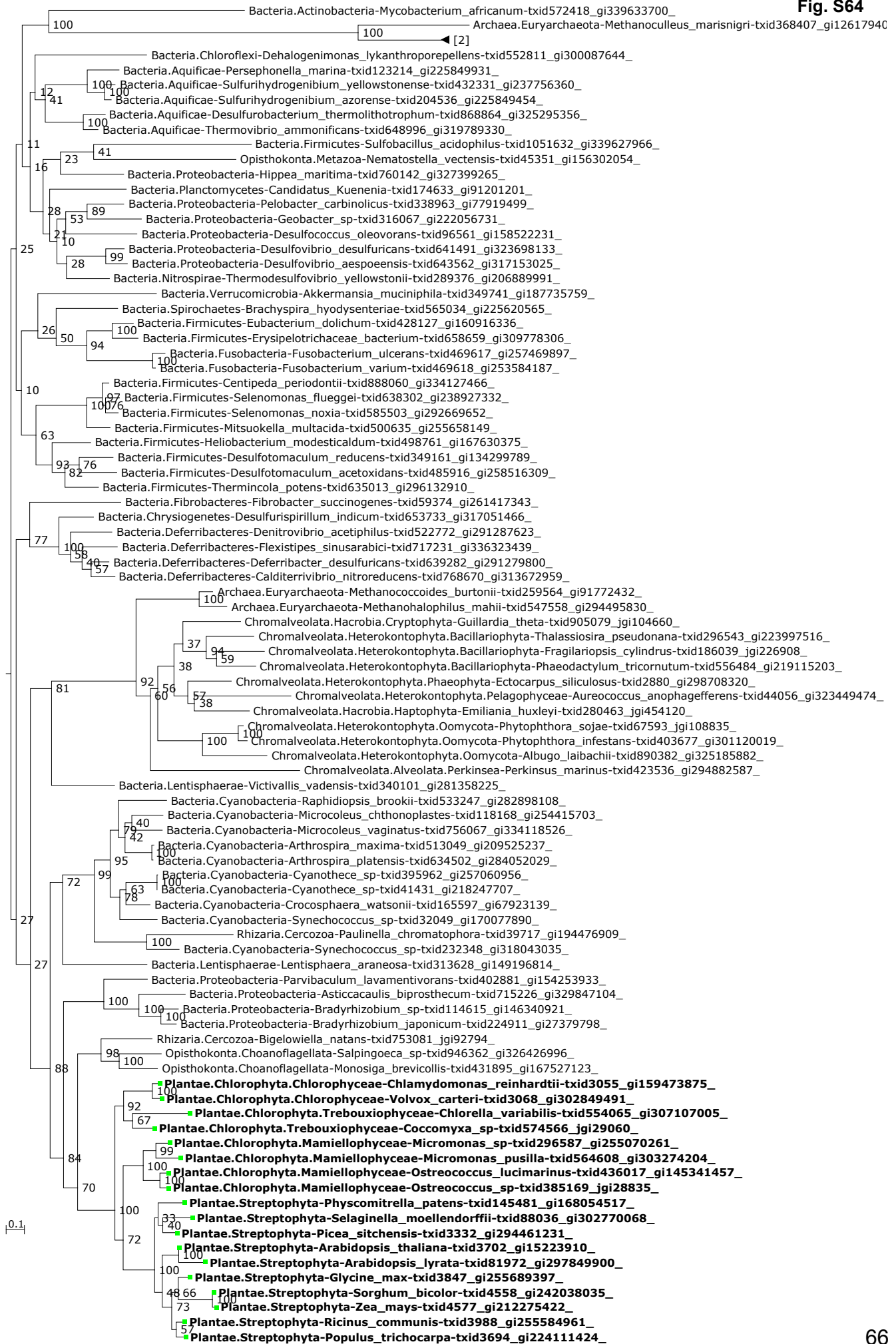


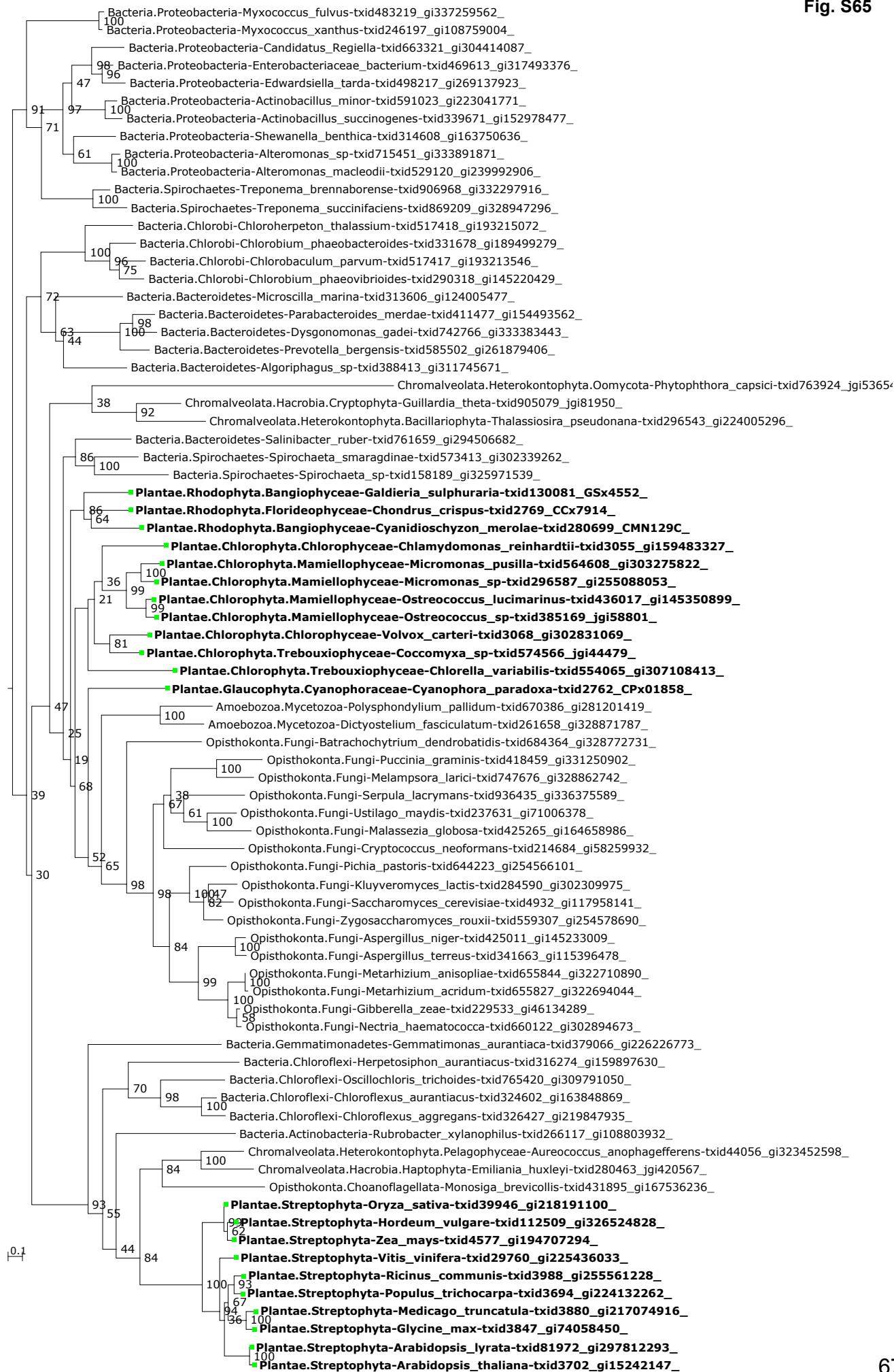


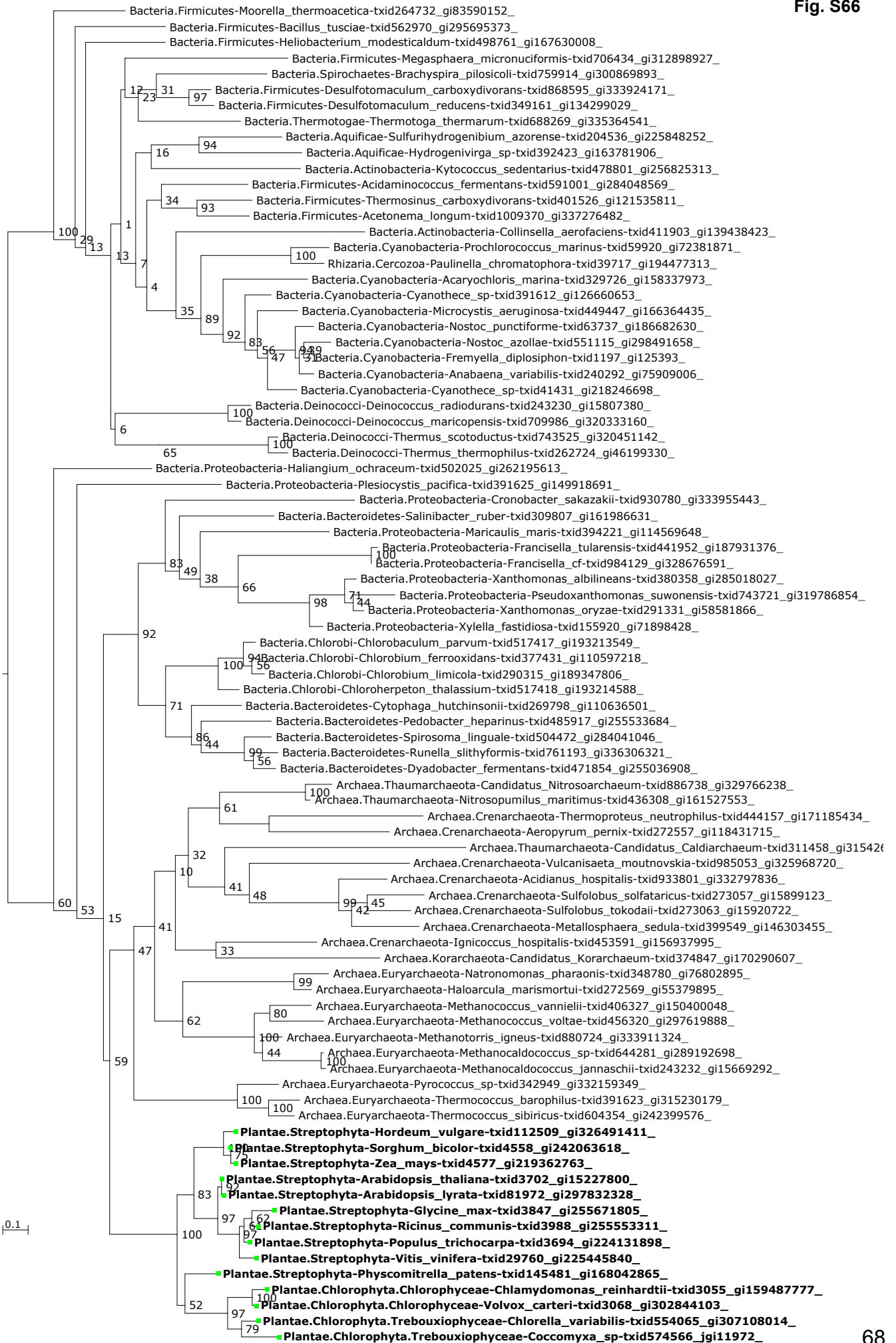


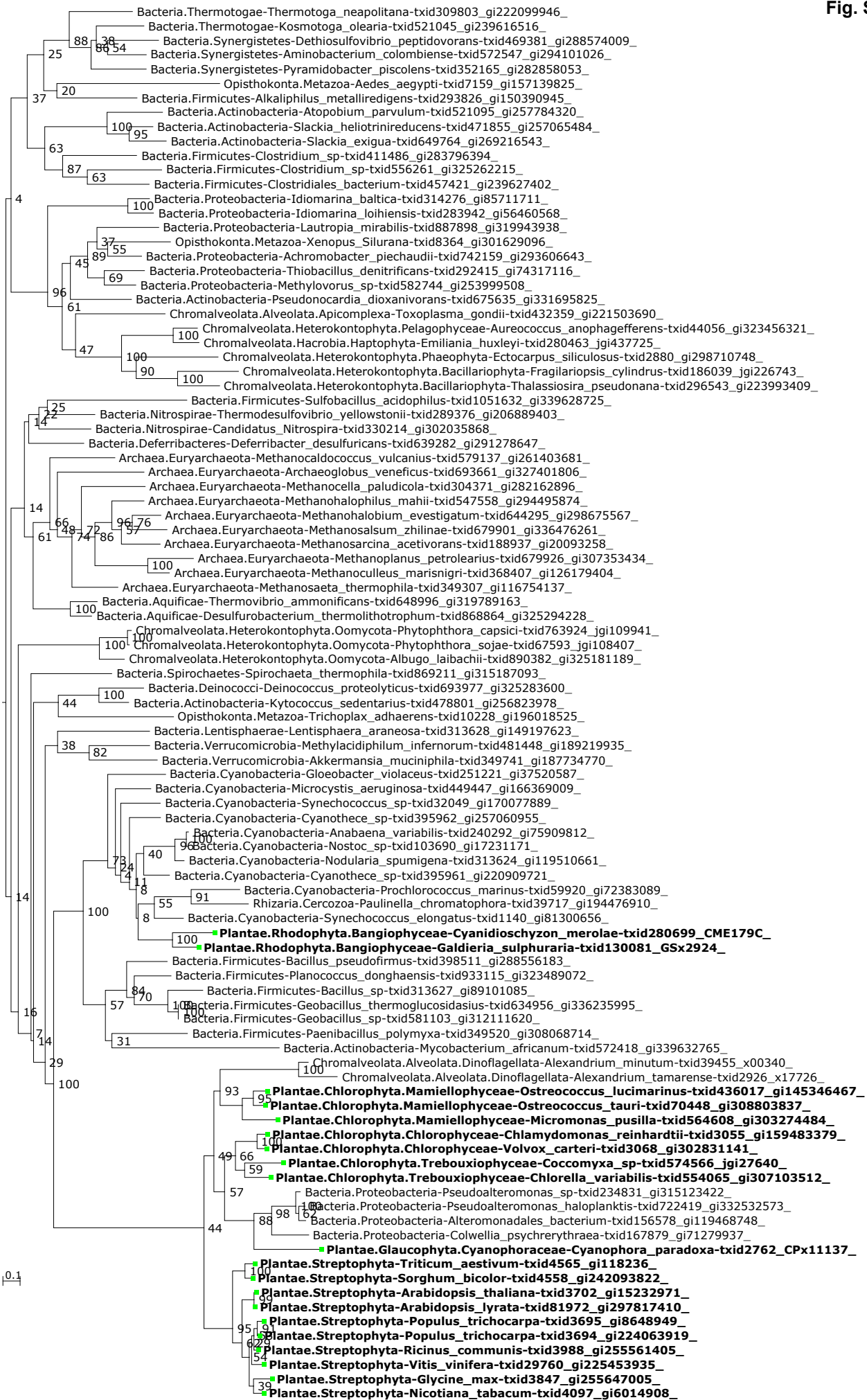


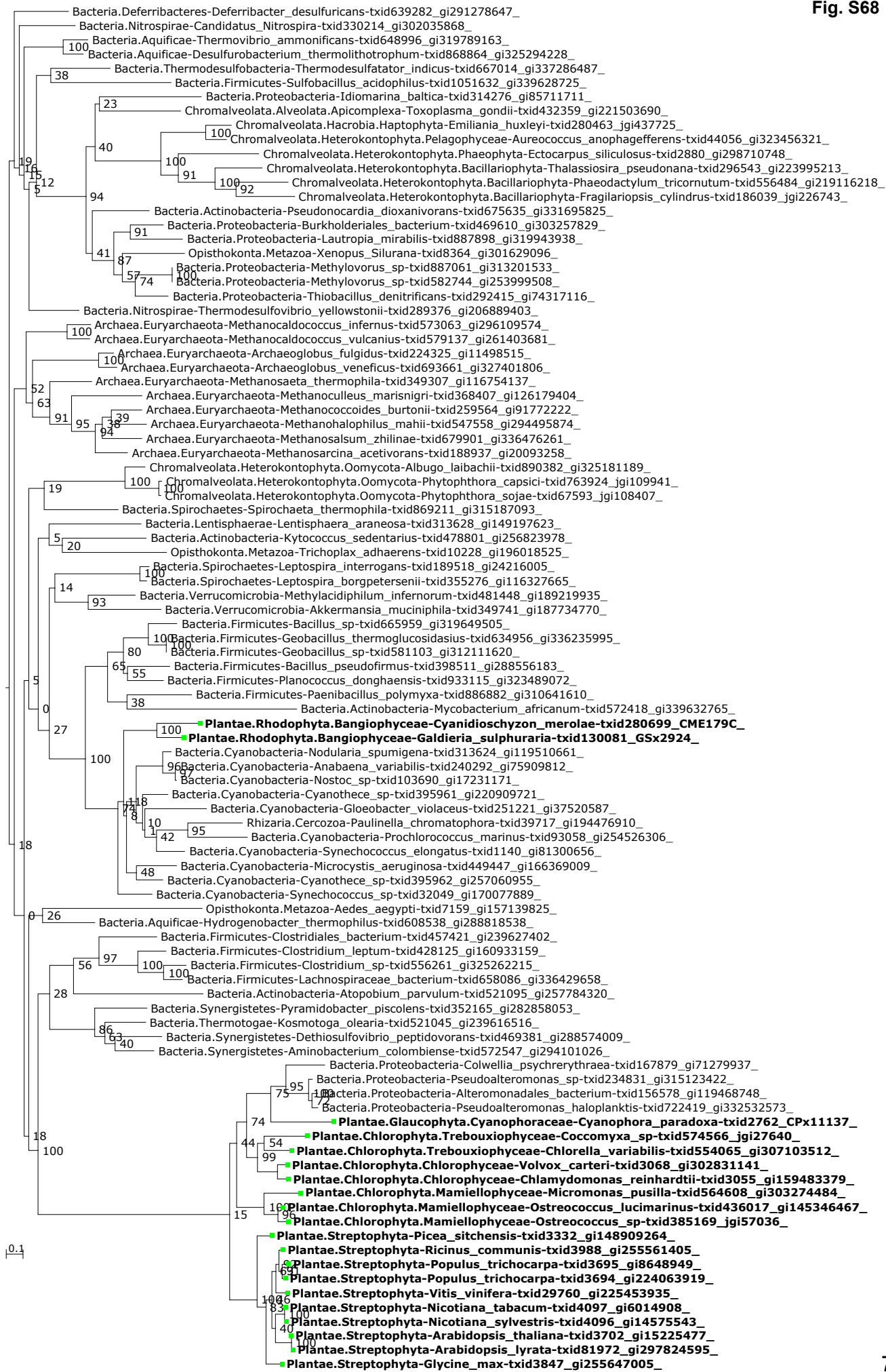


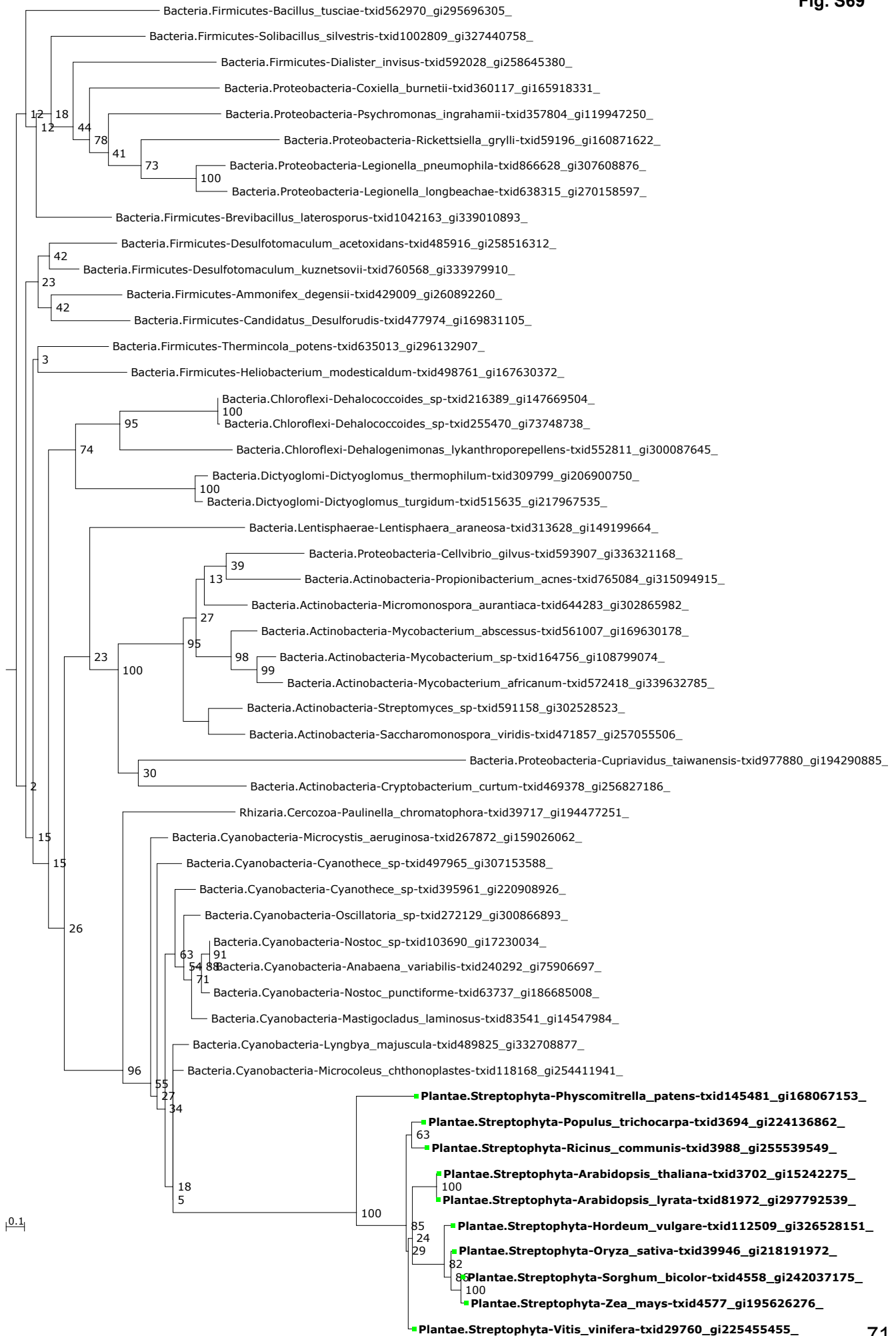


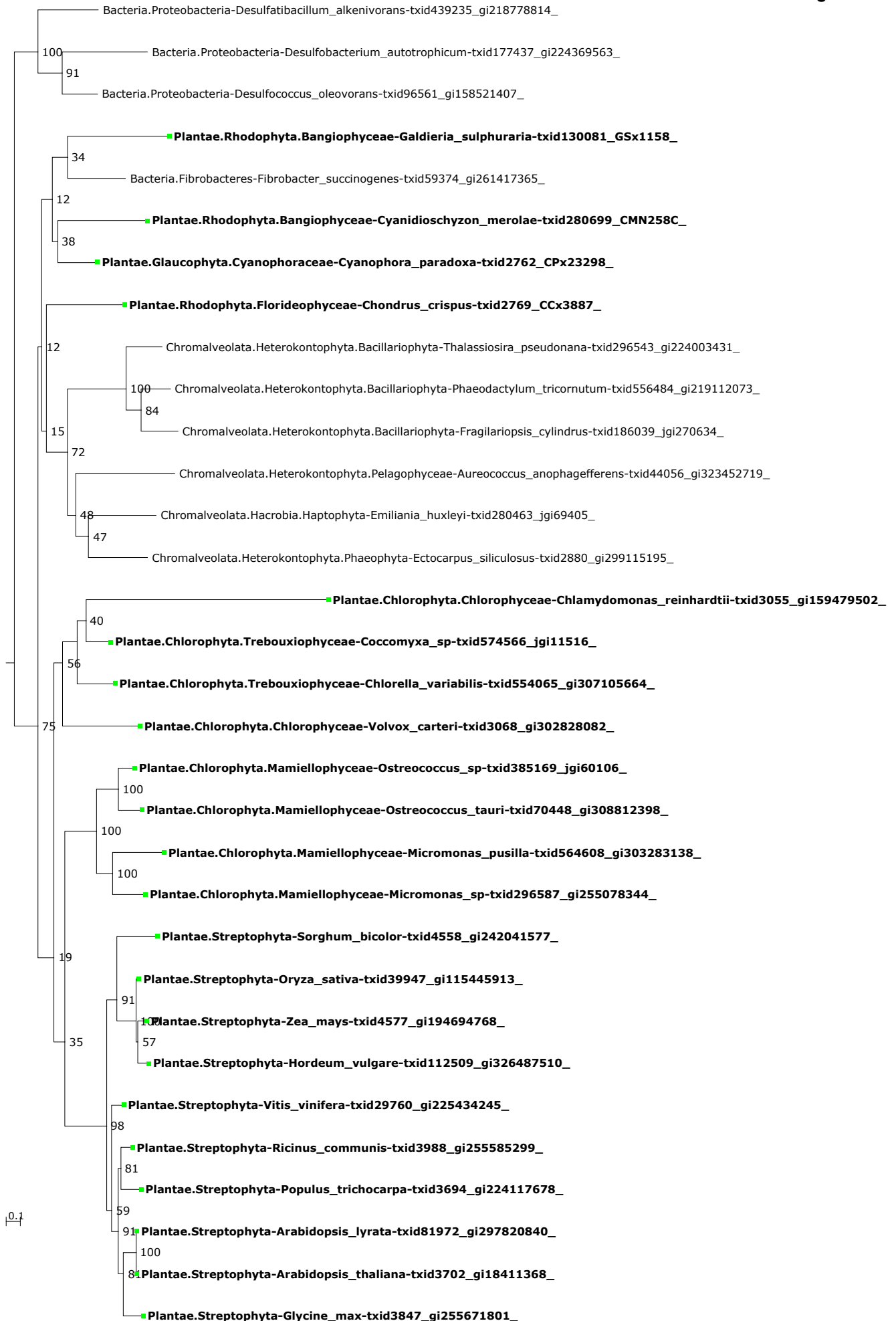




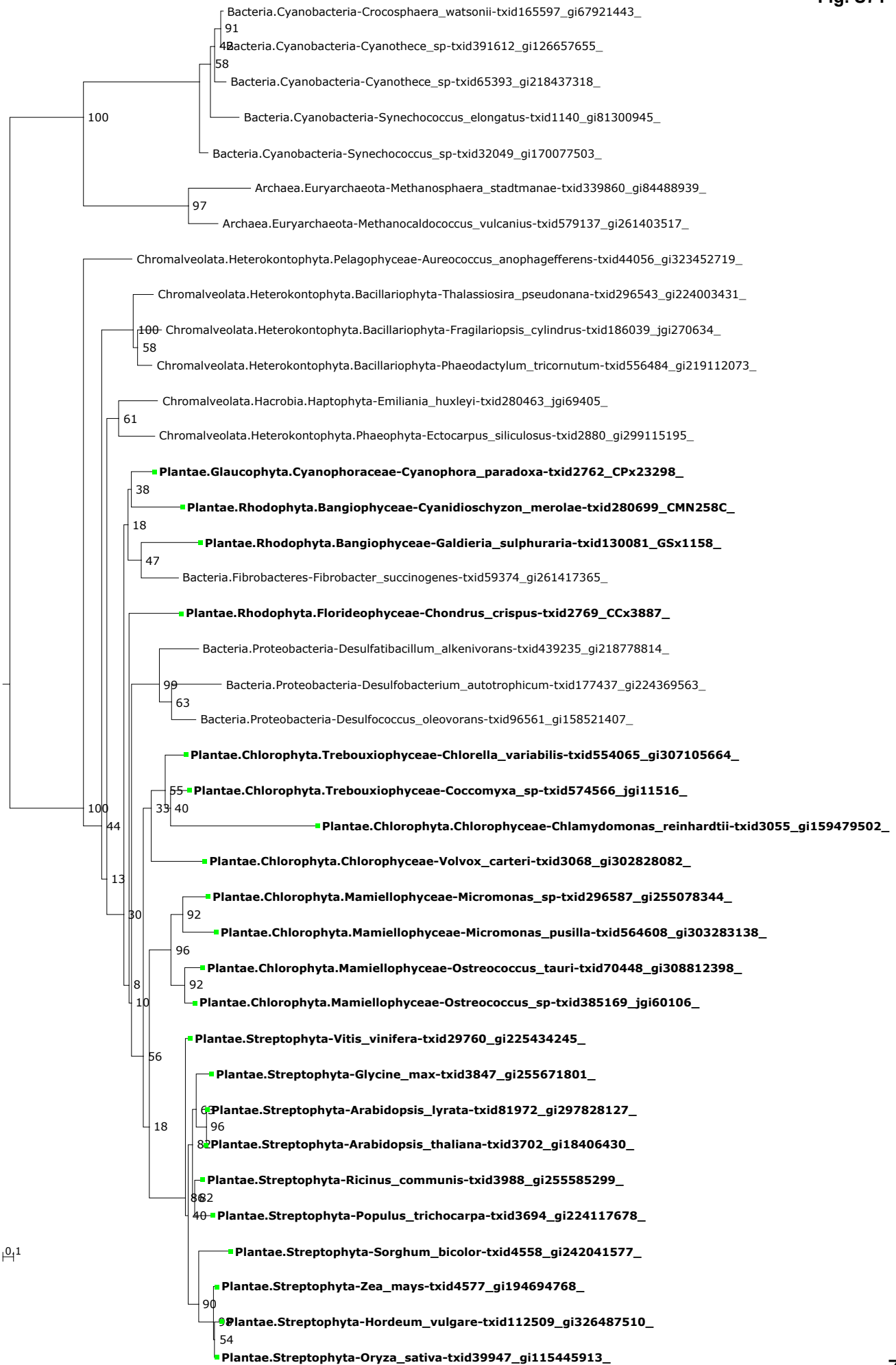




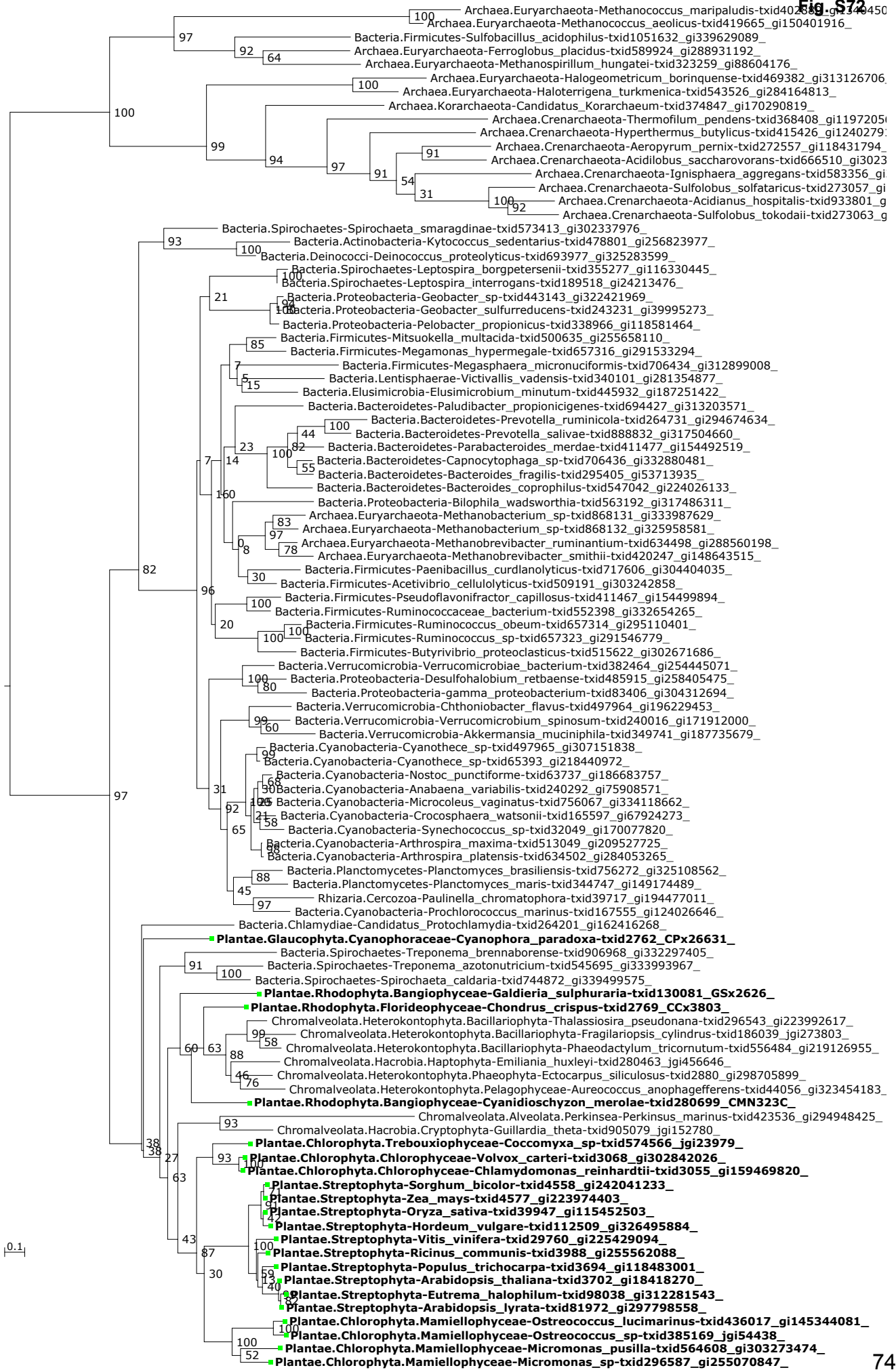


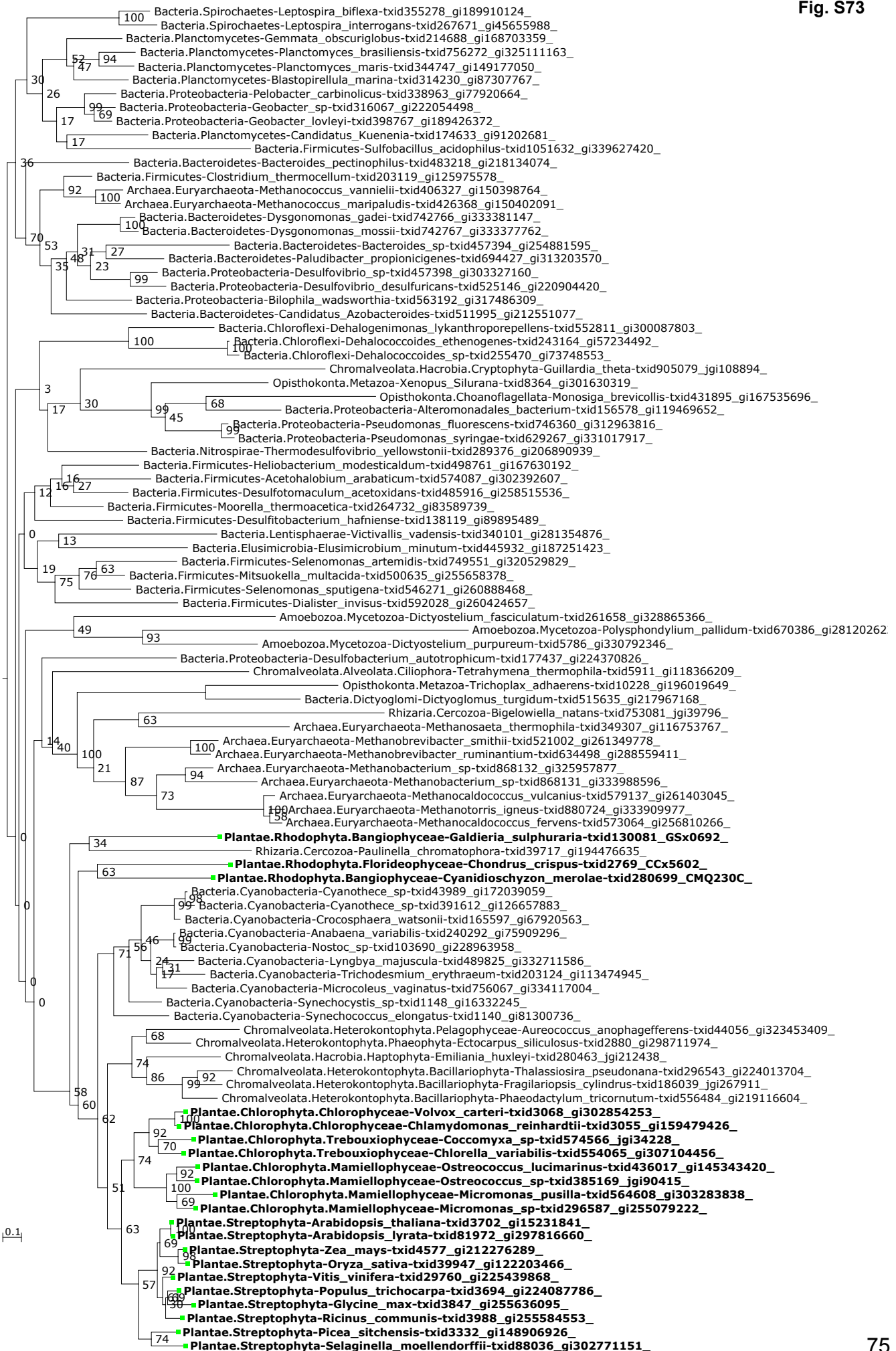






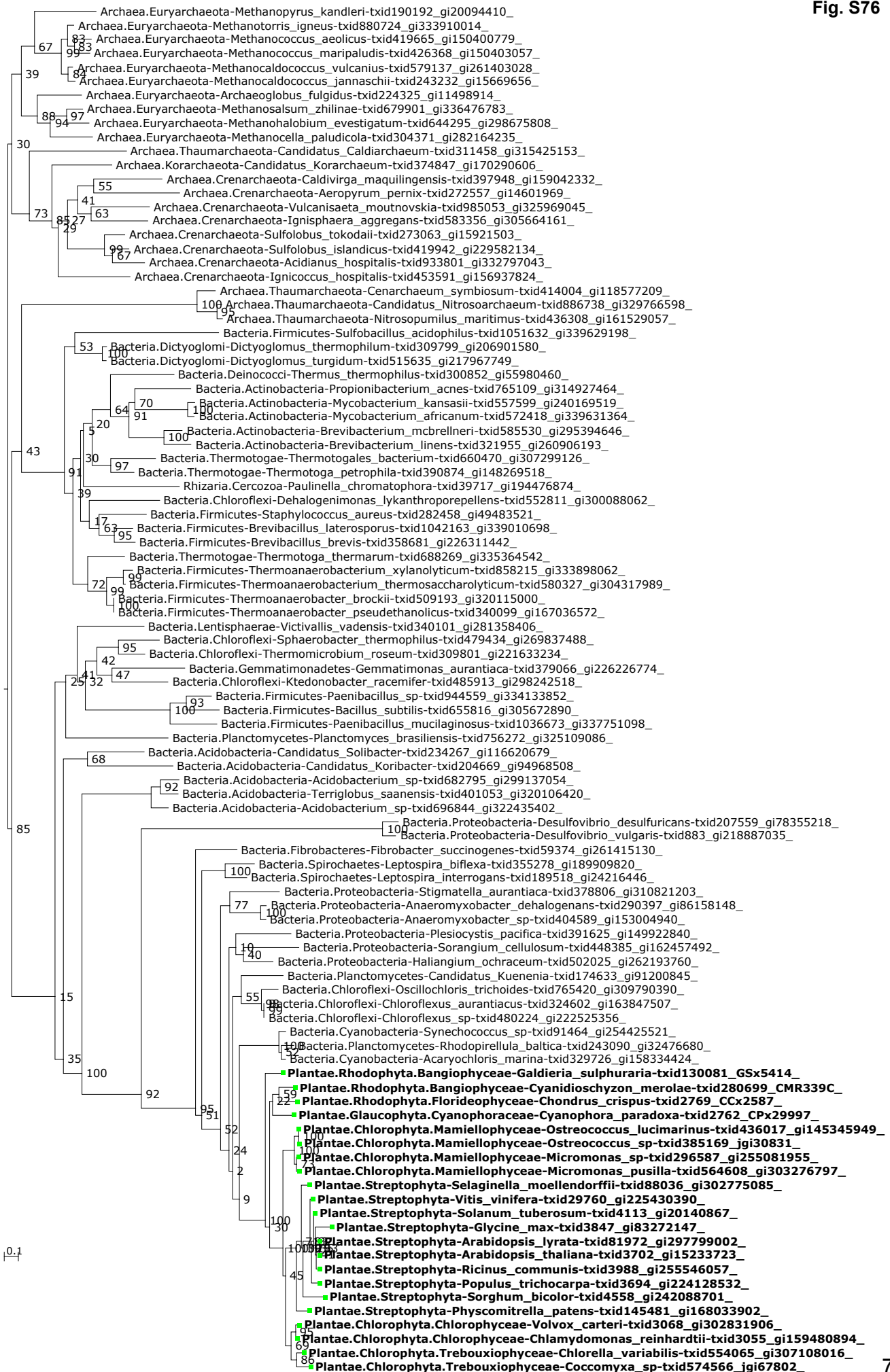
0,1

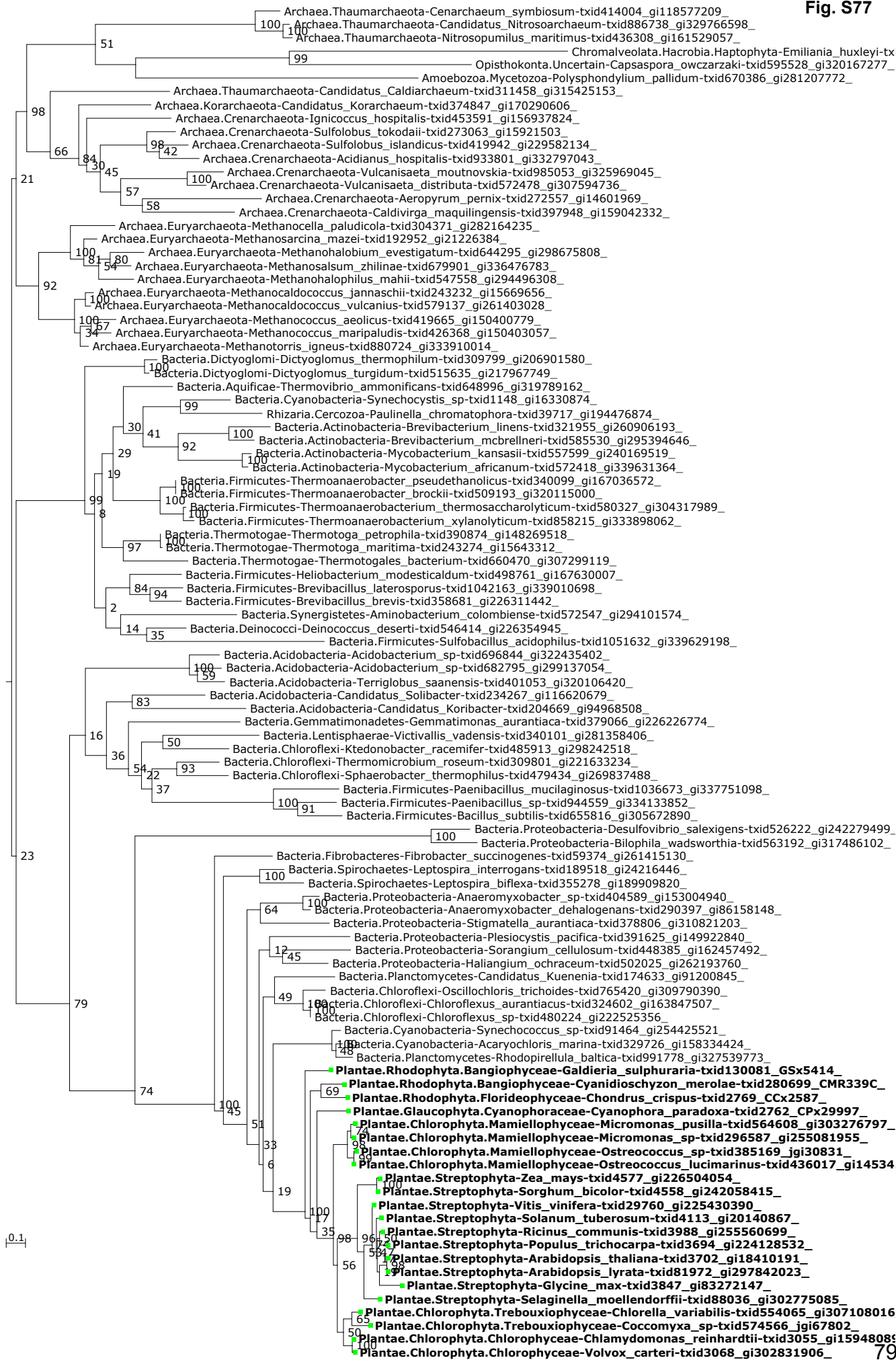


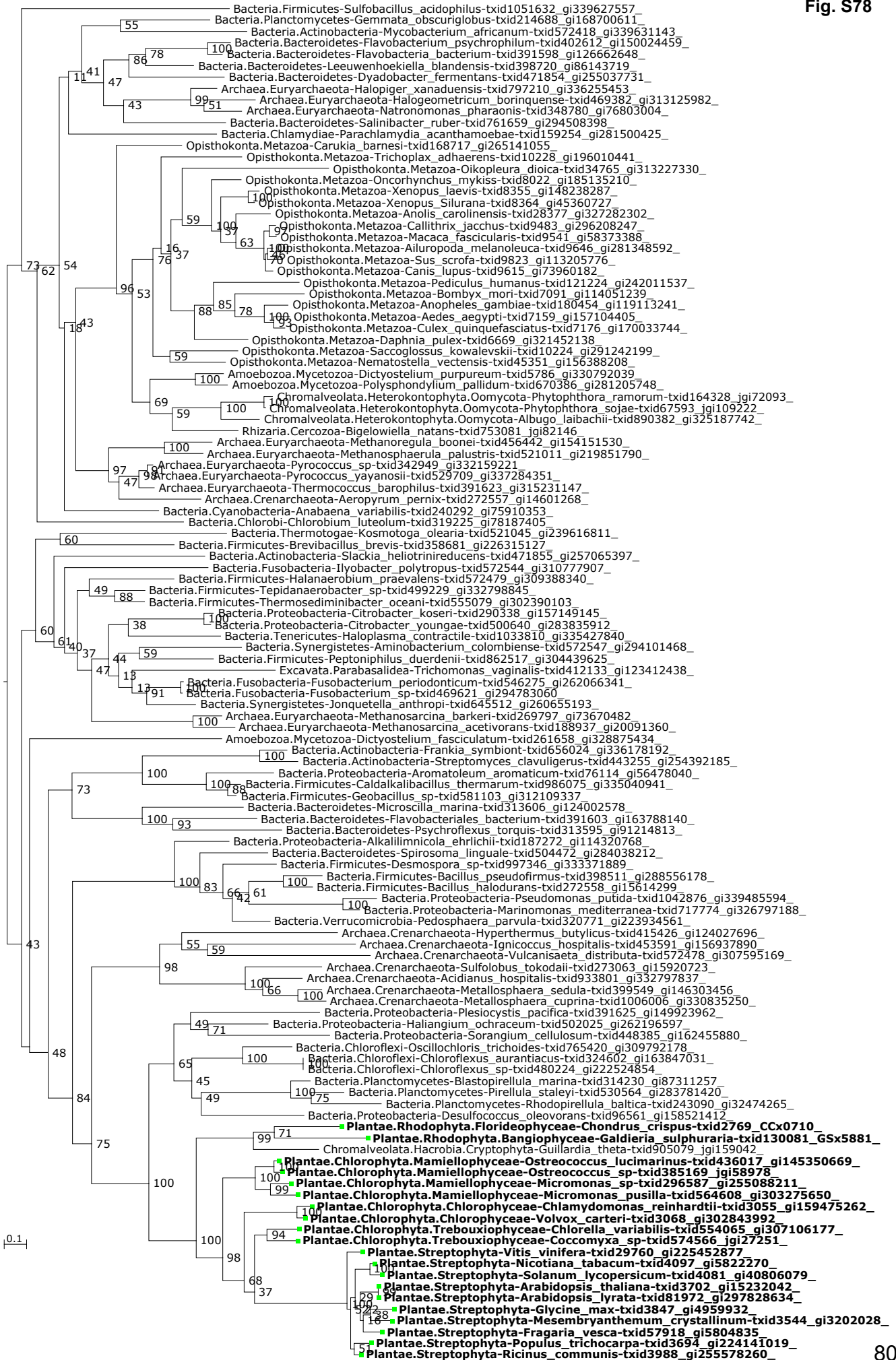




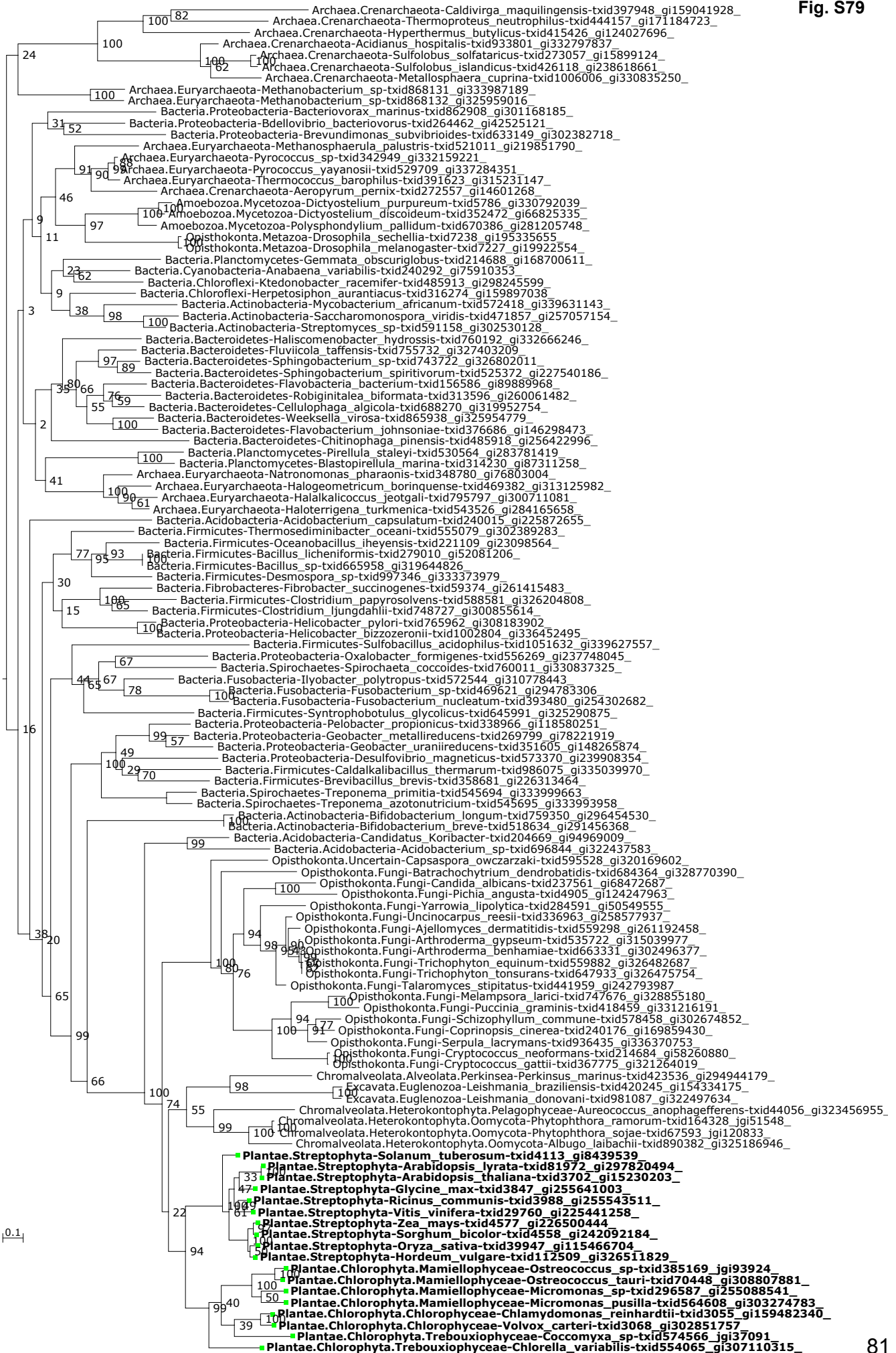




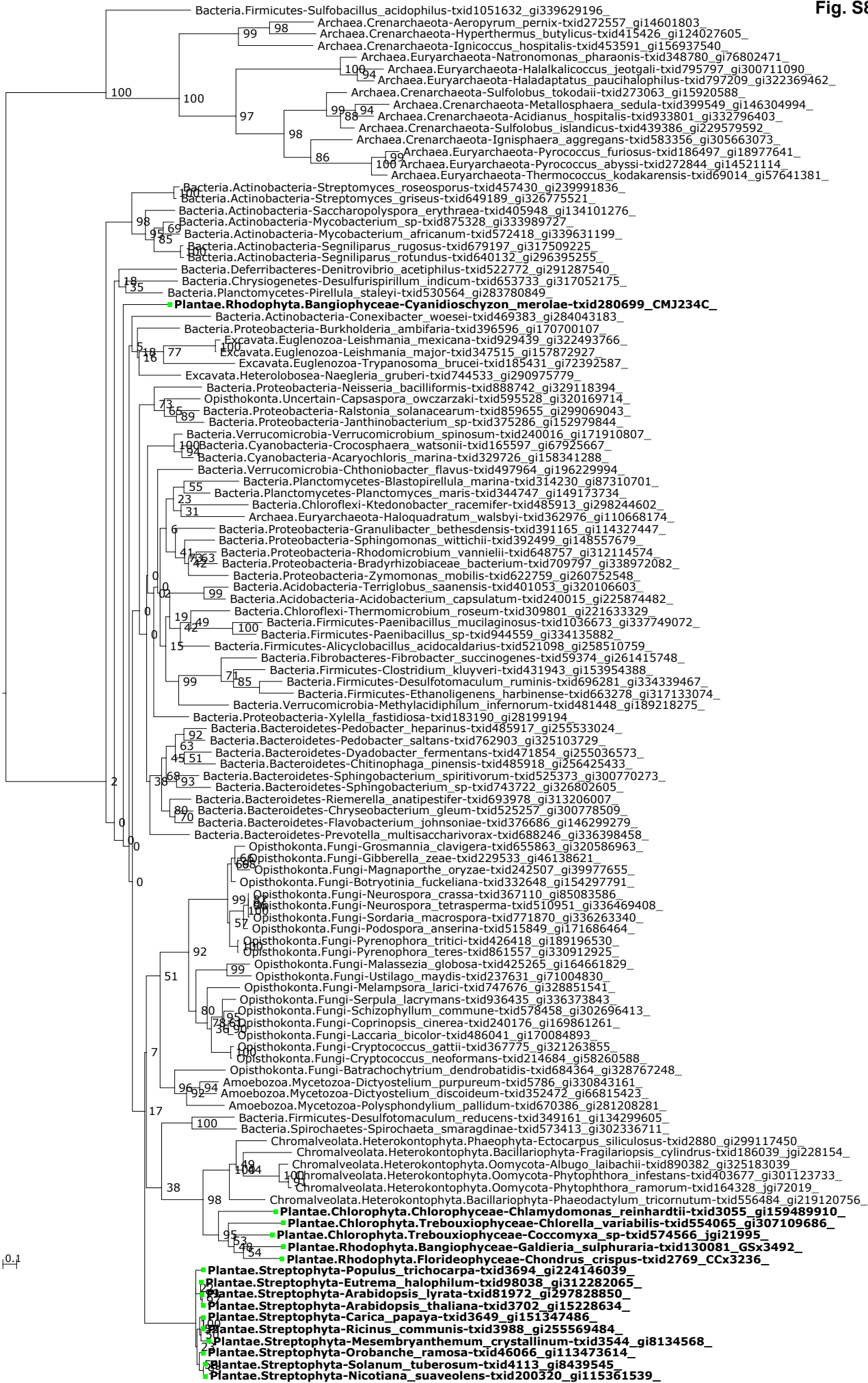




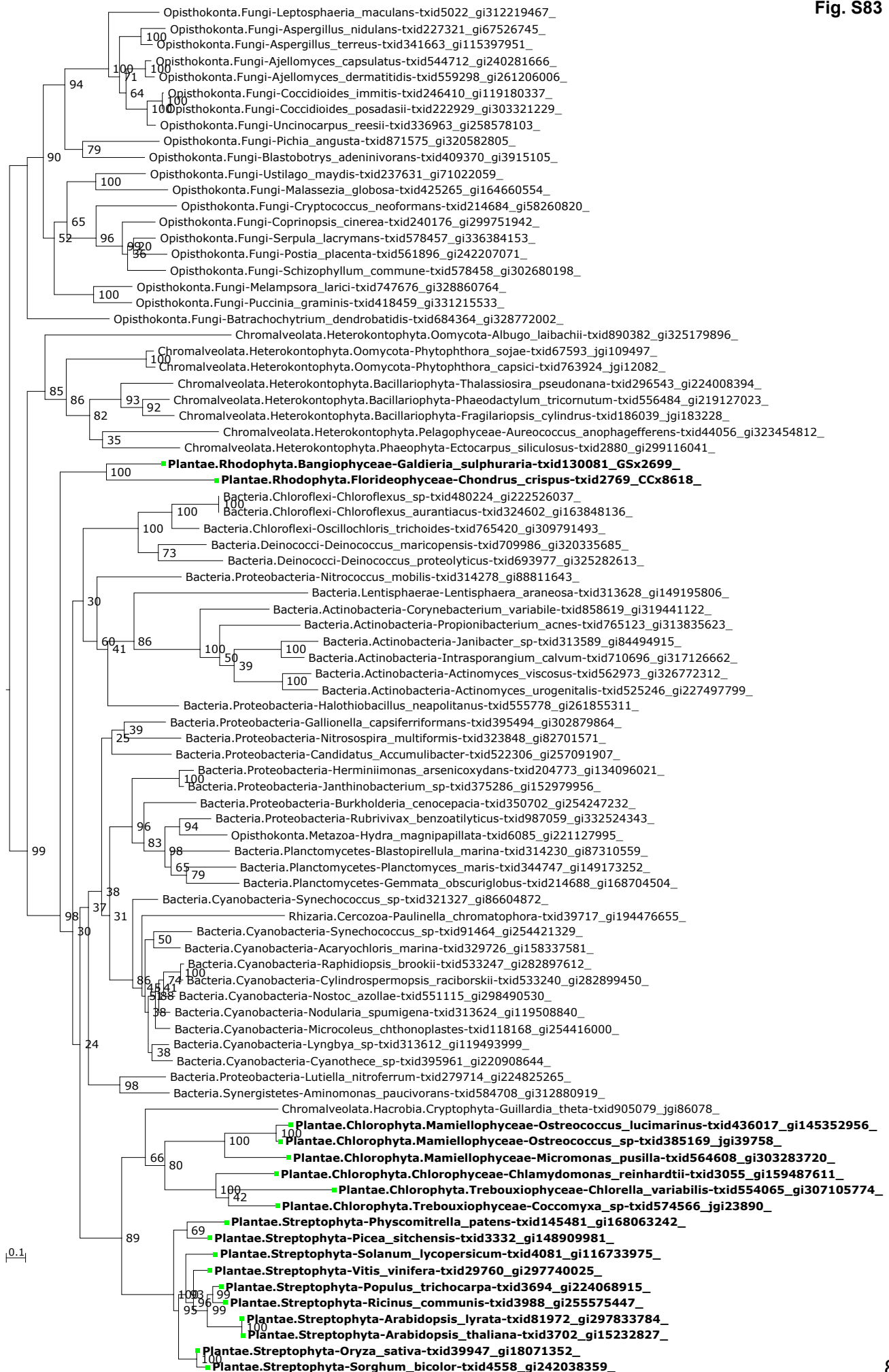


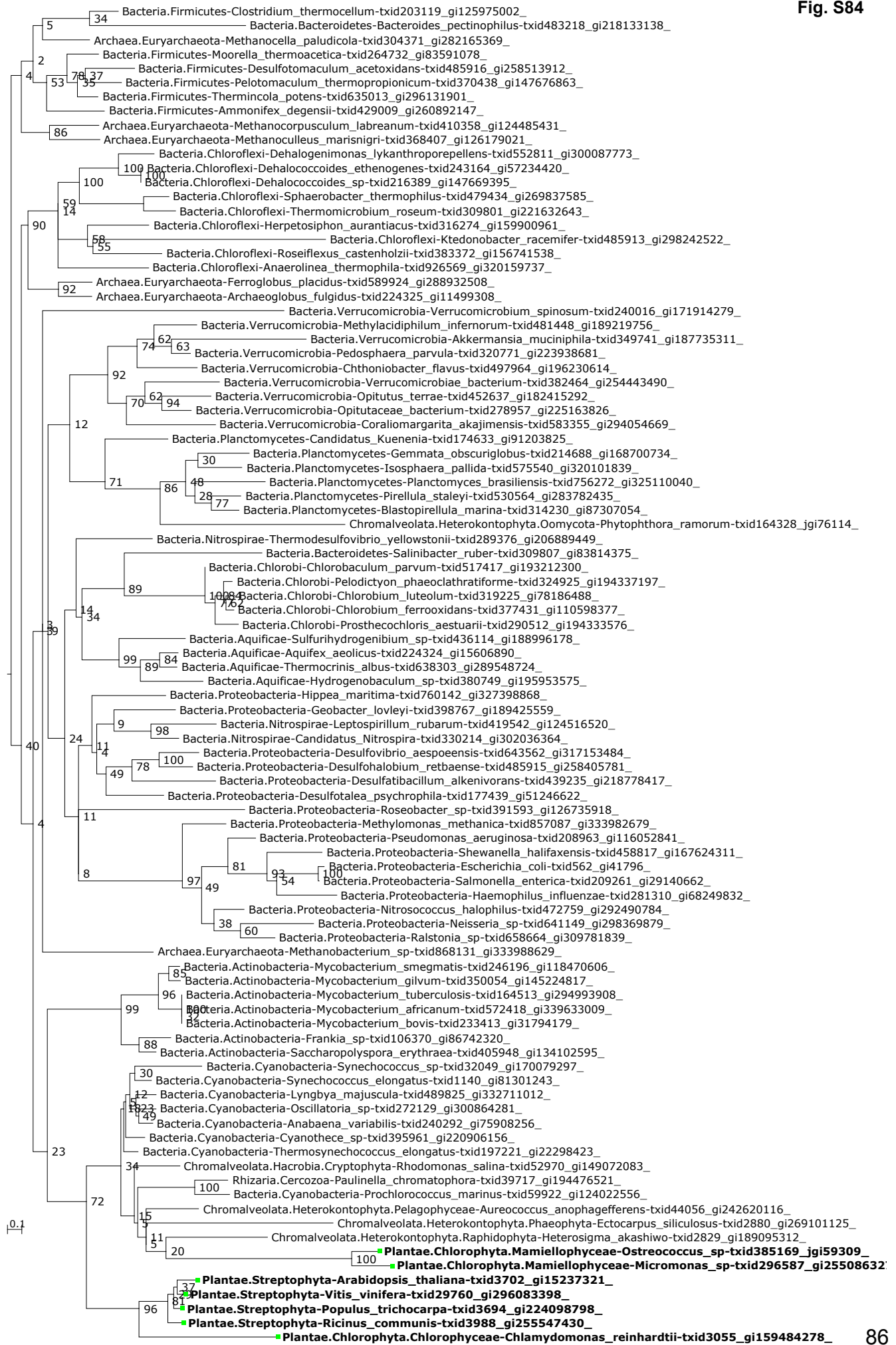








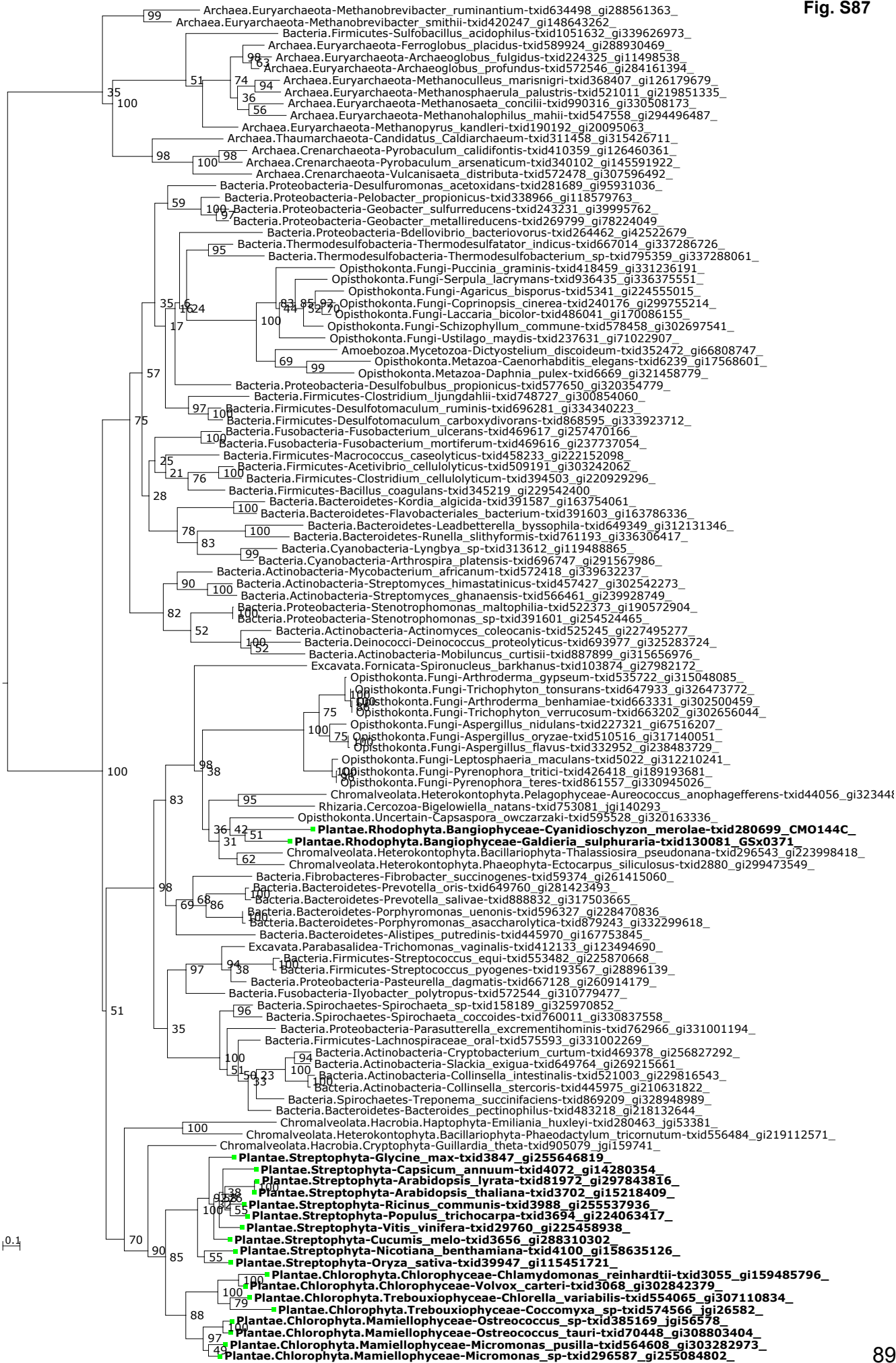


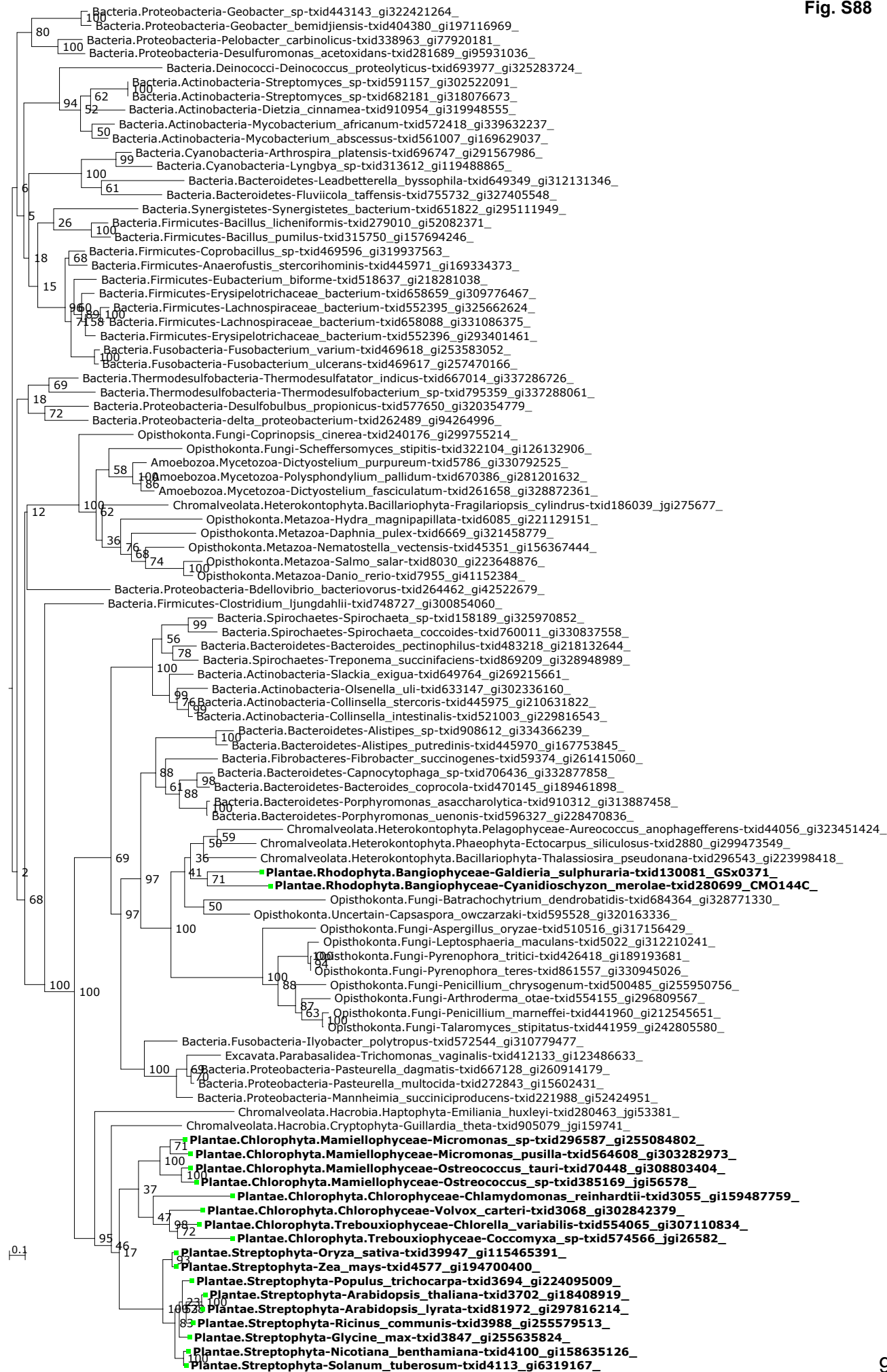


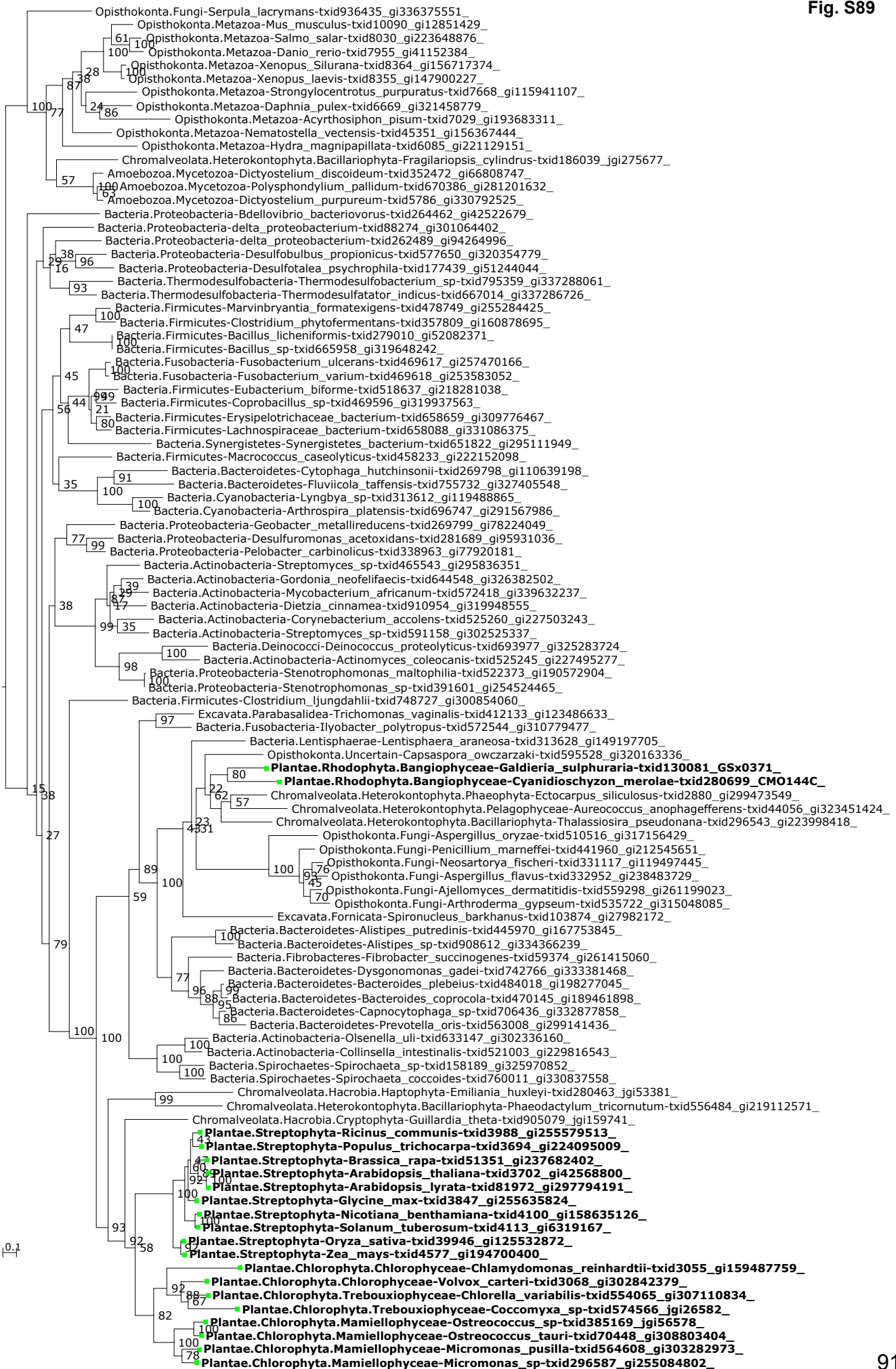




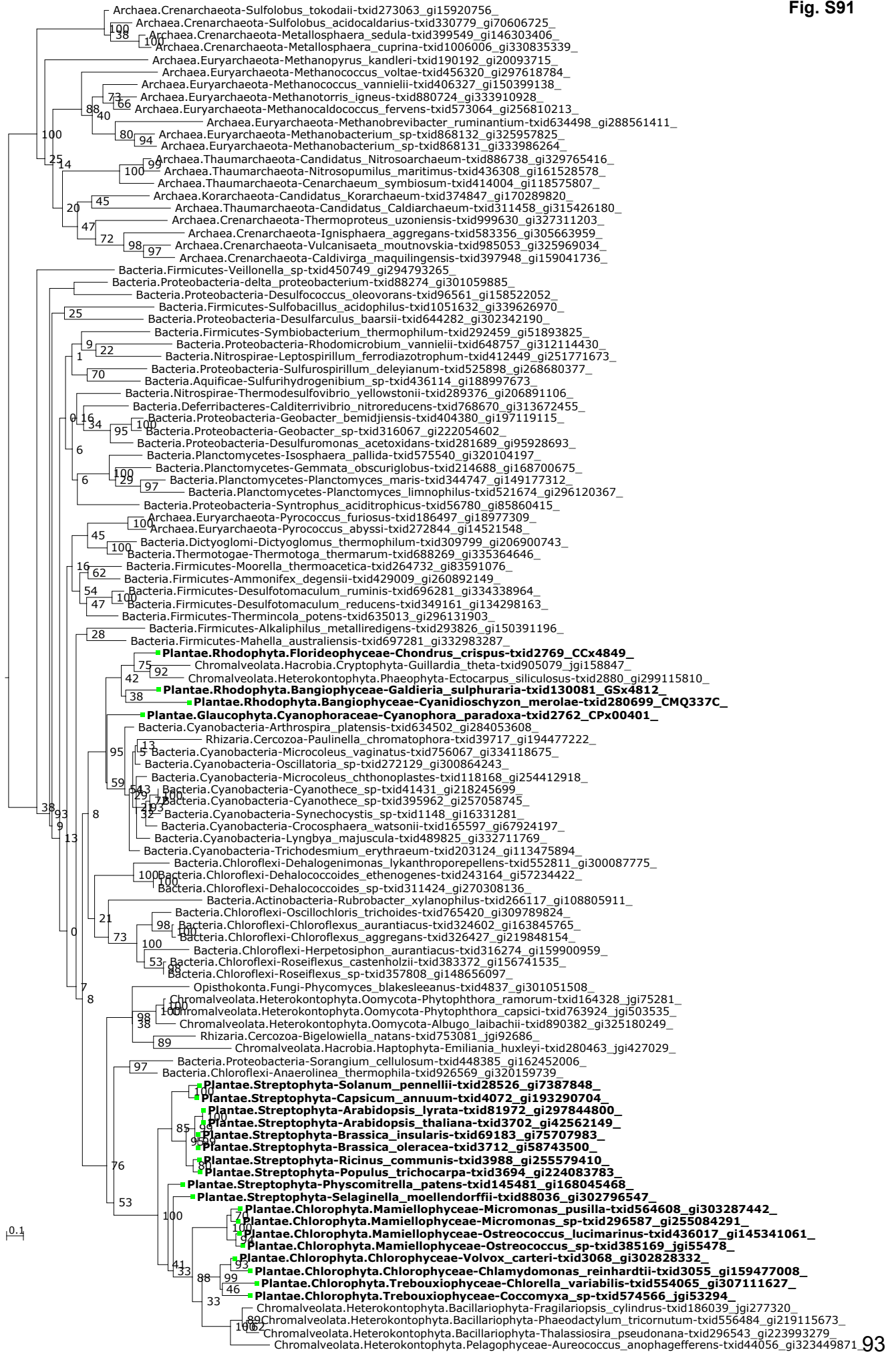


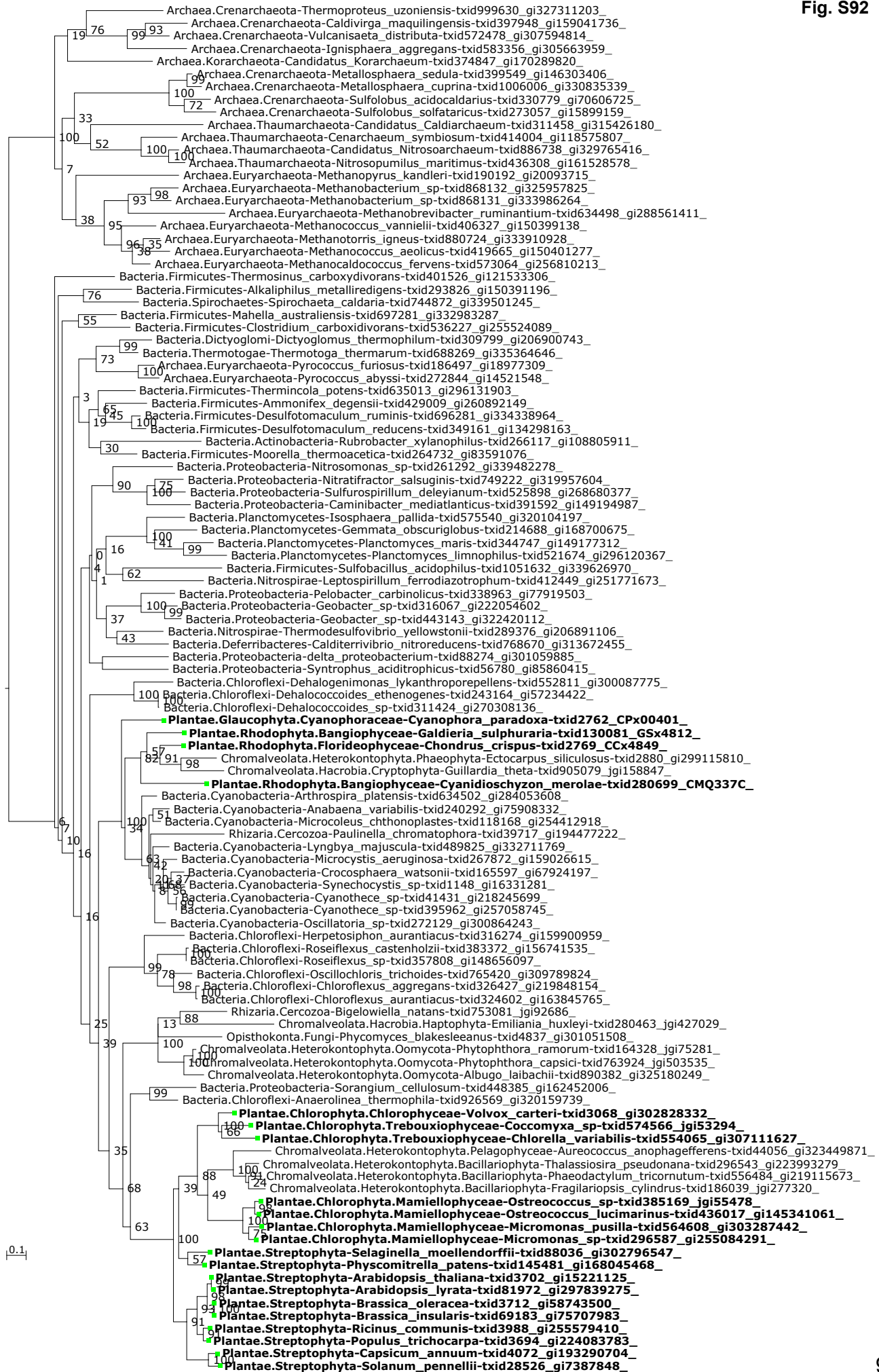










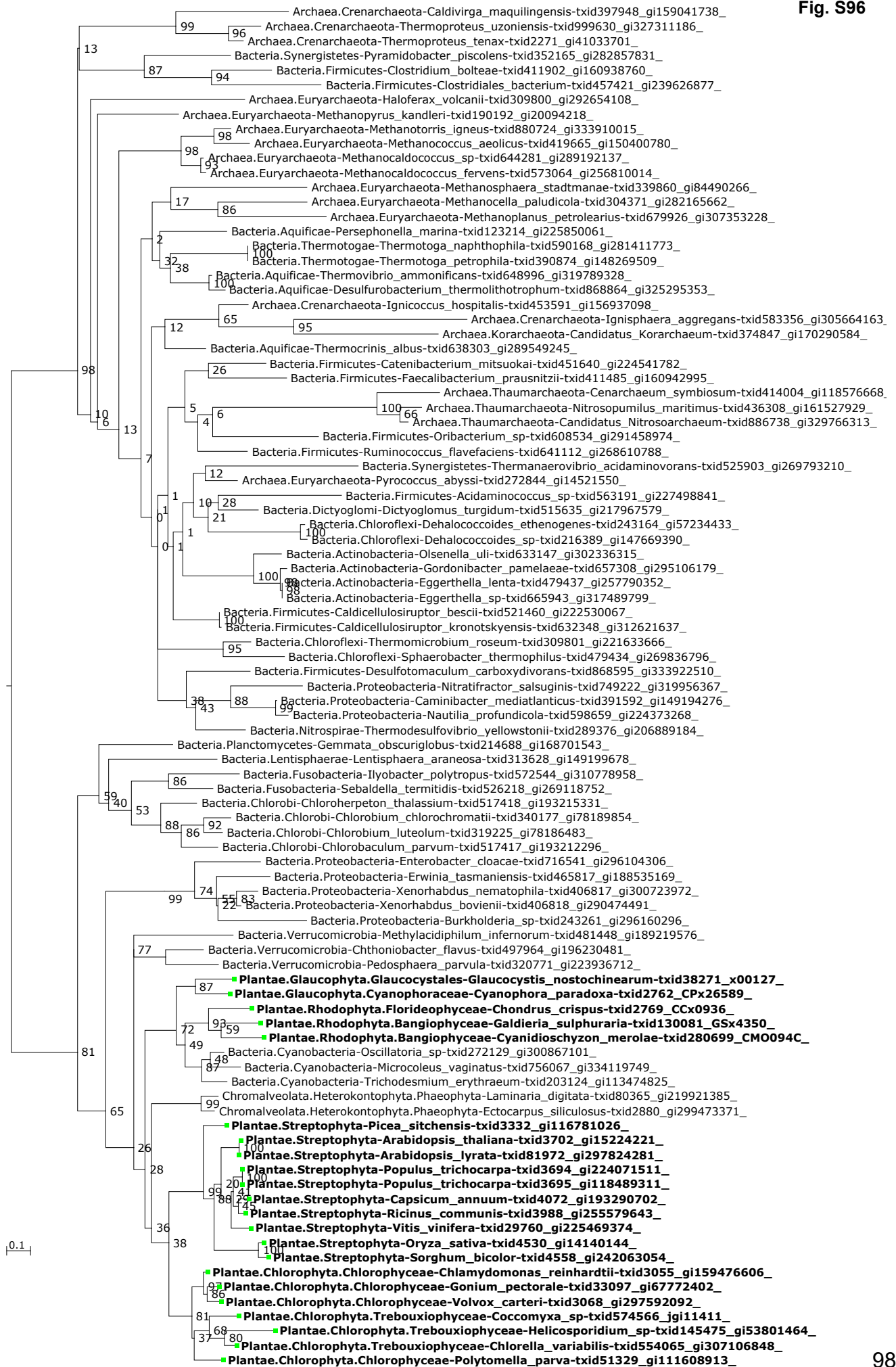


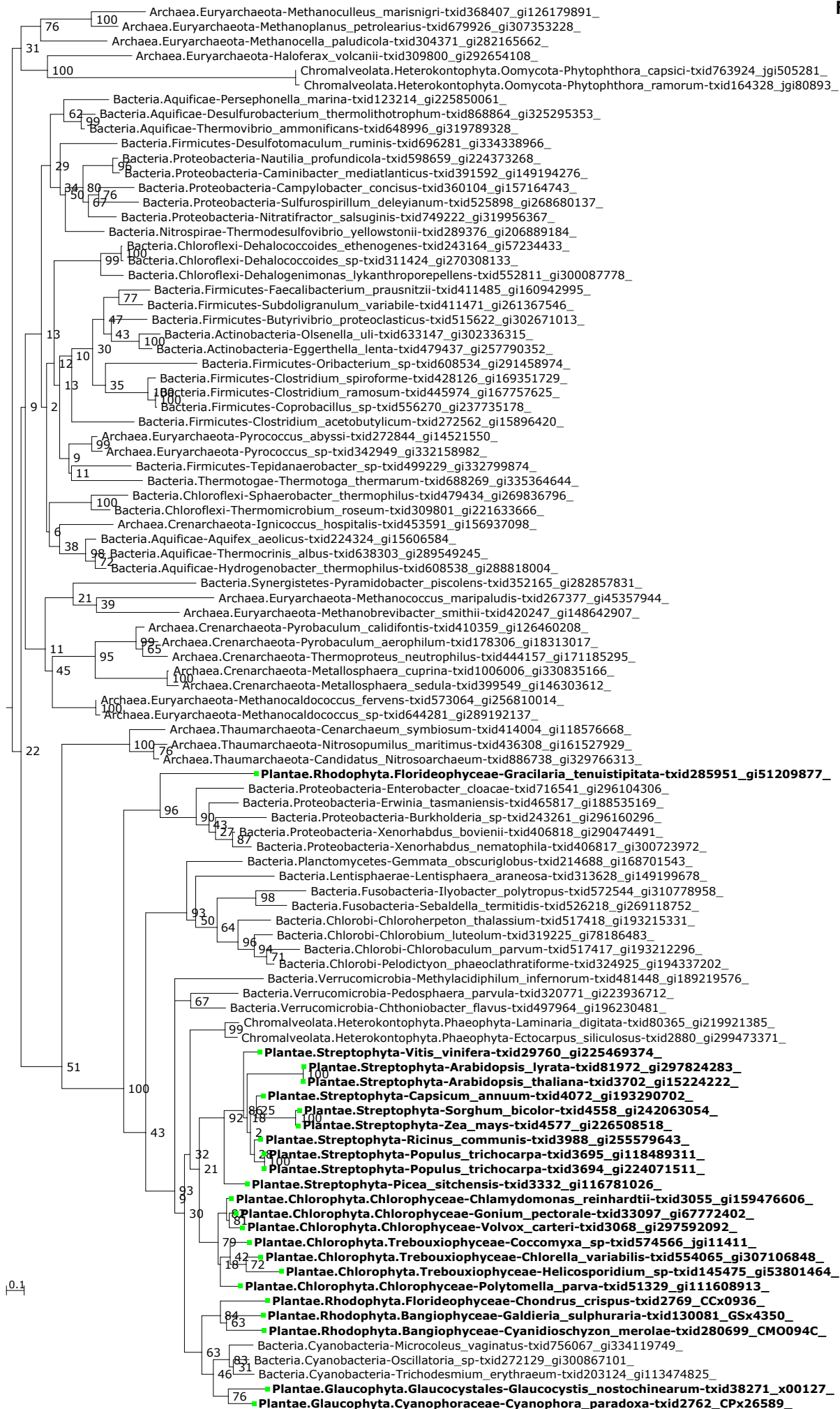


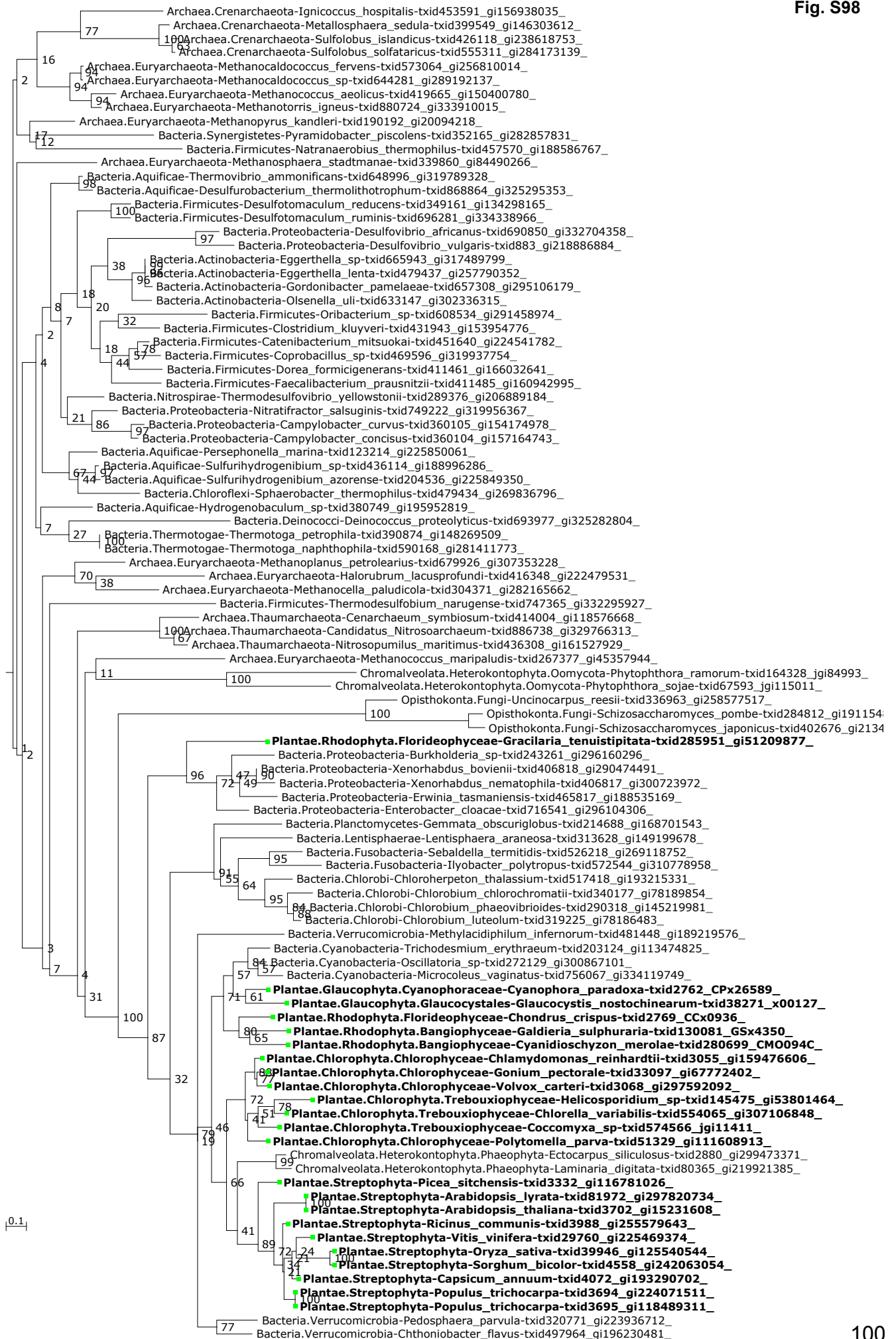


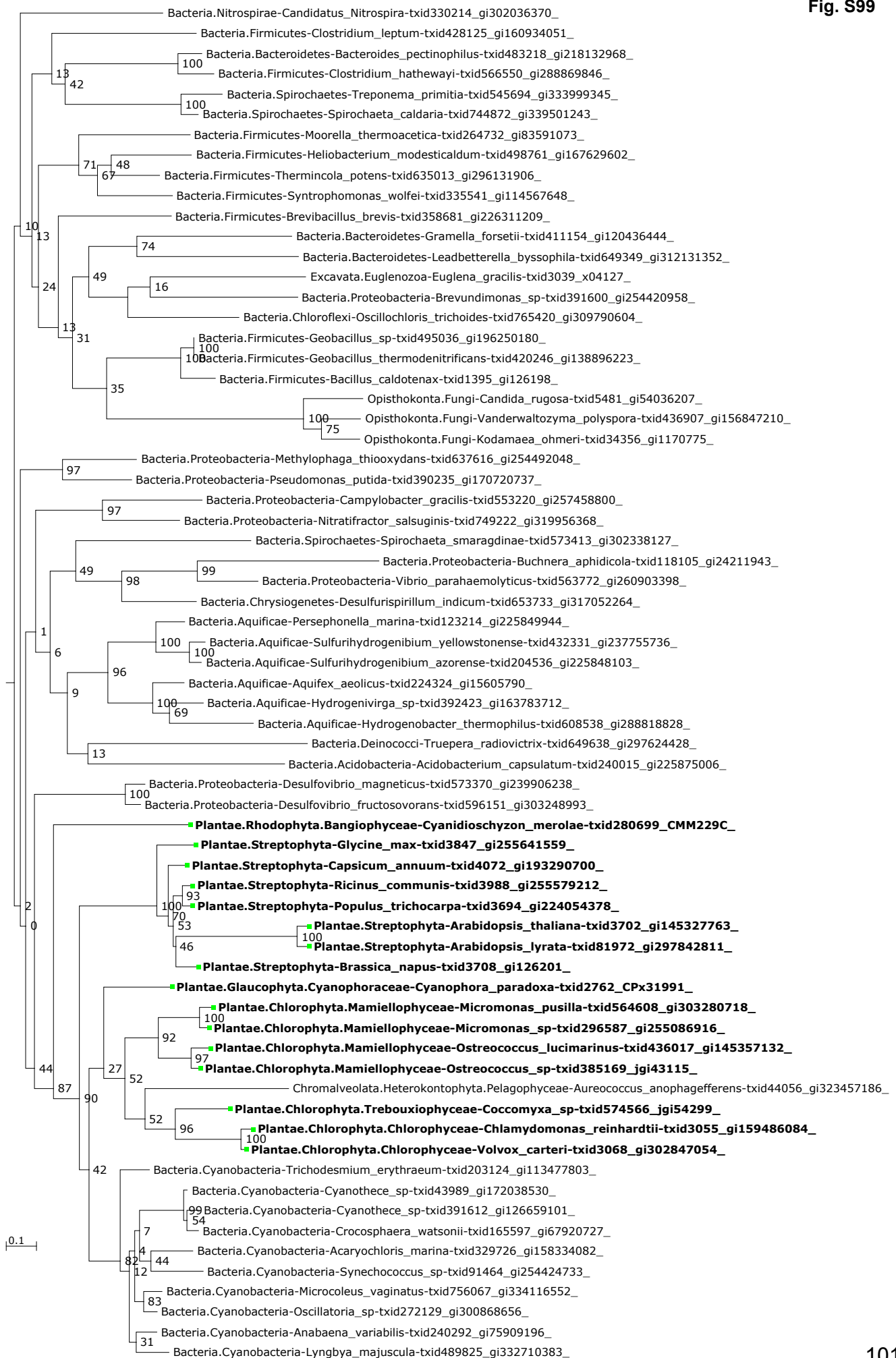


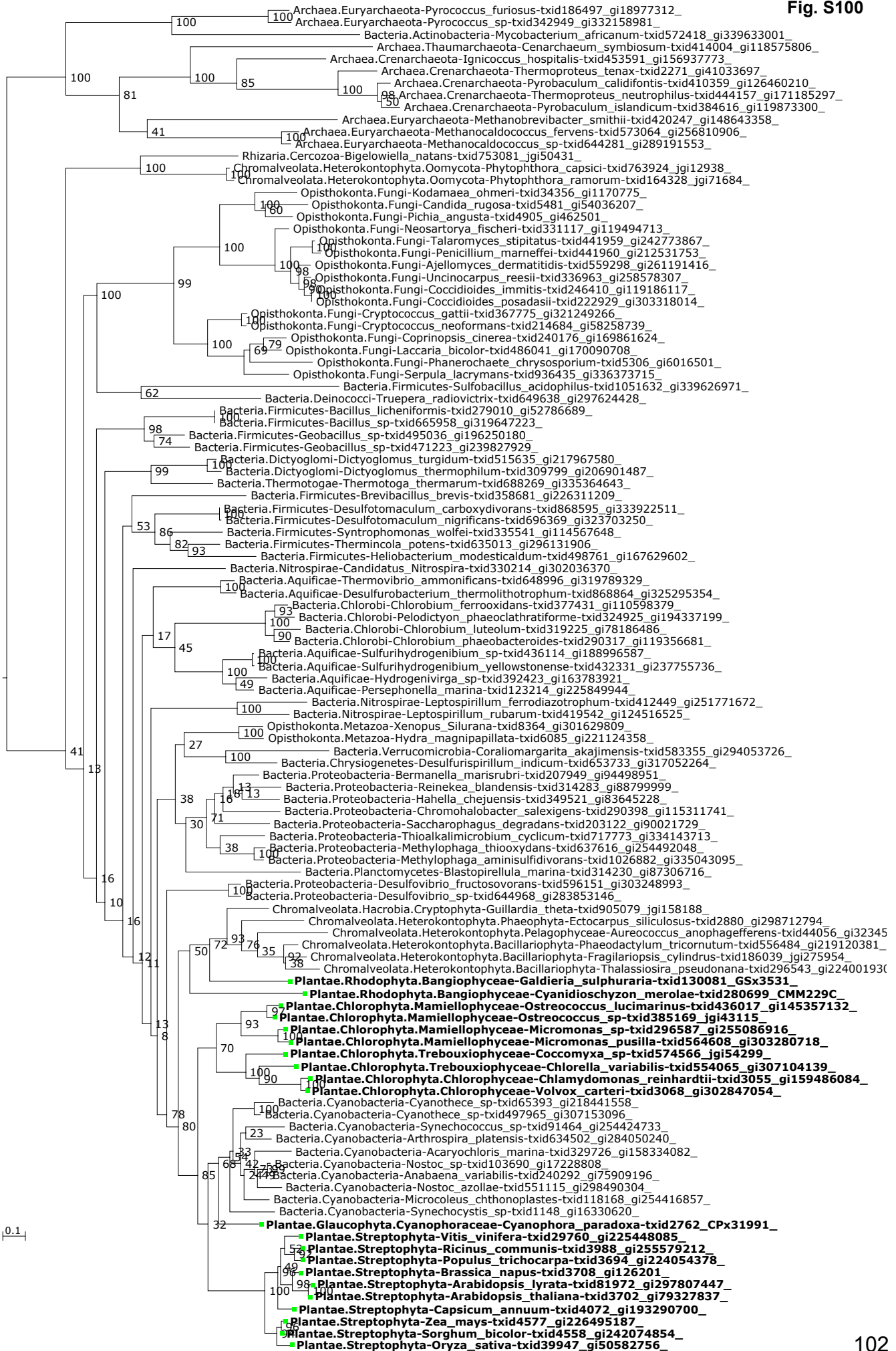










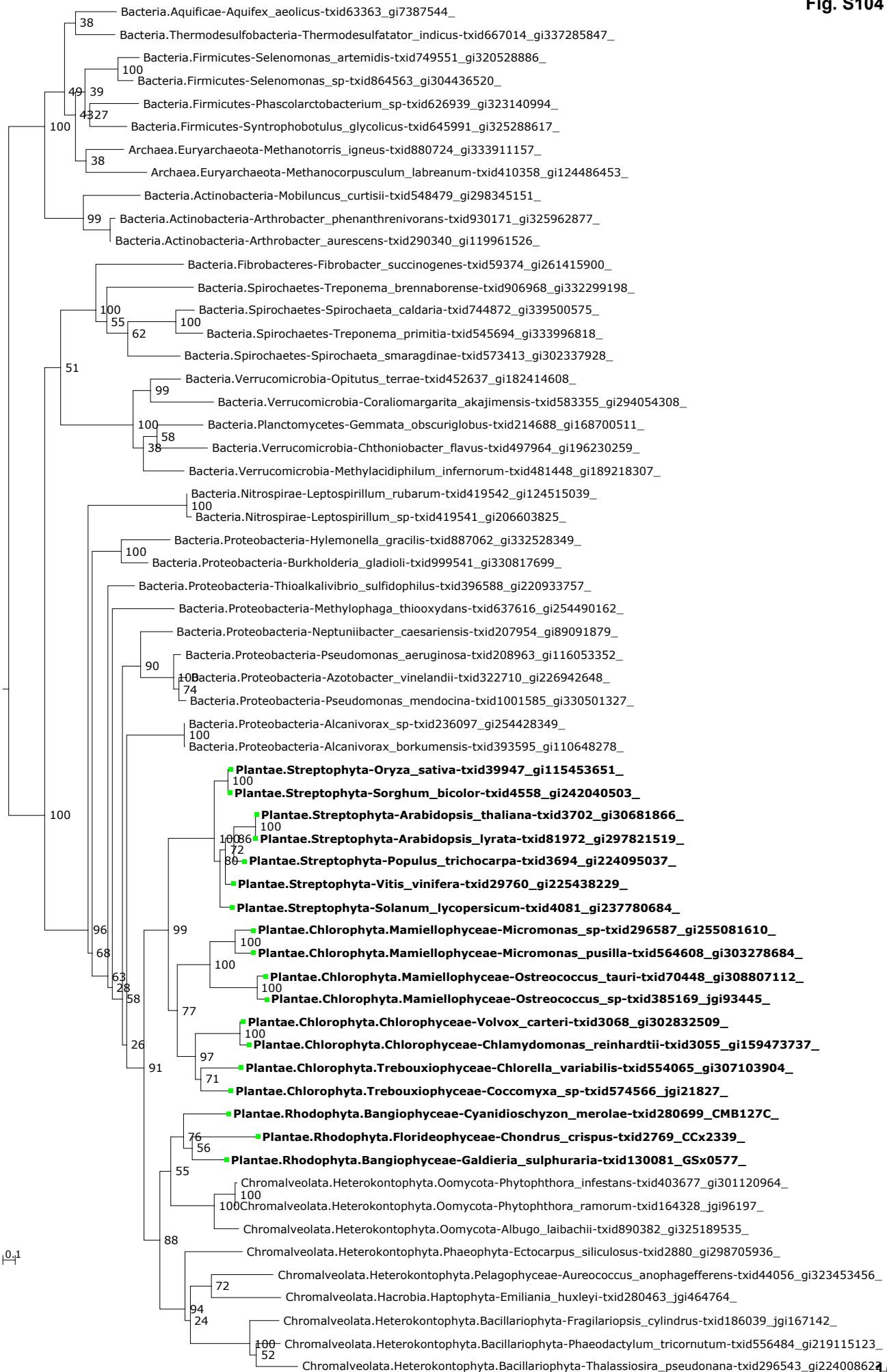




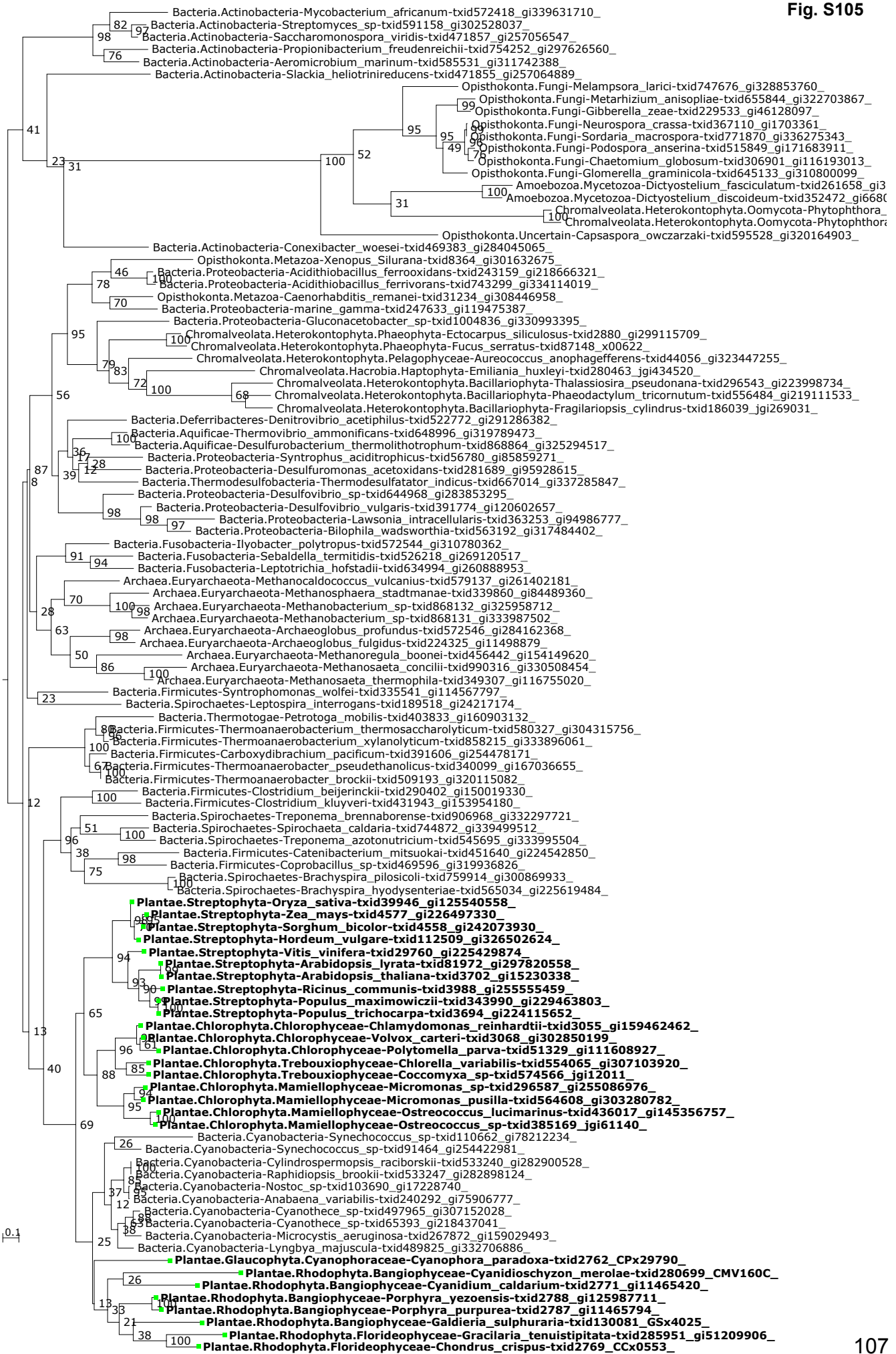


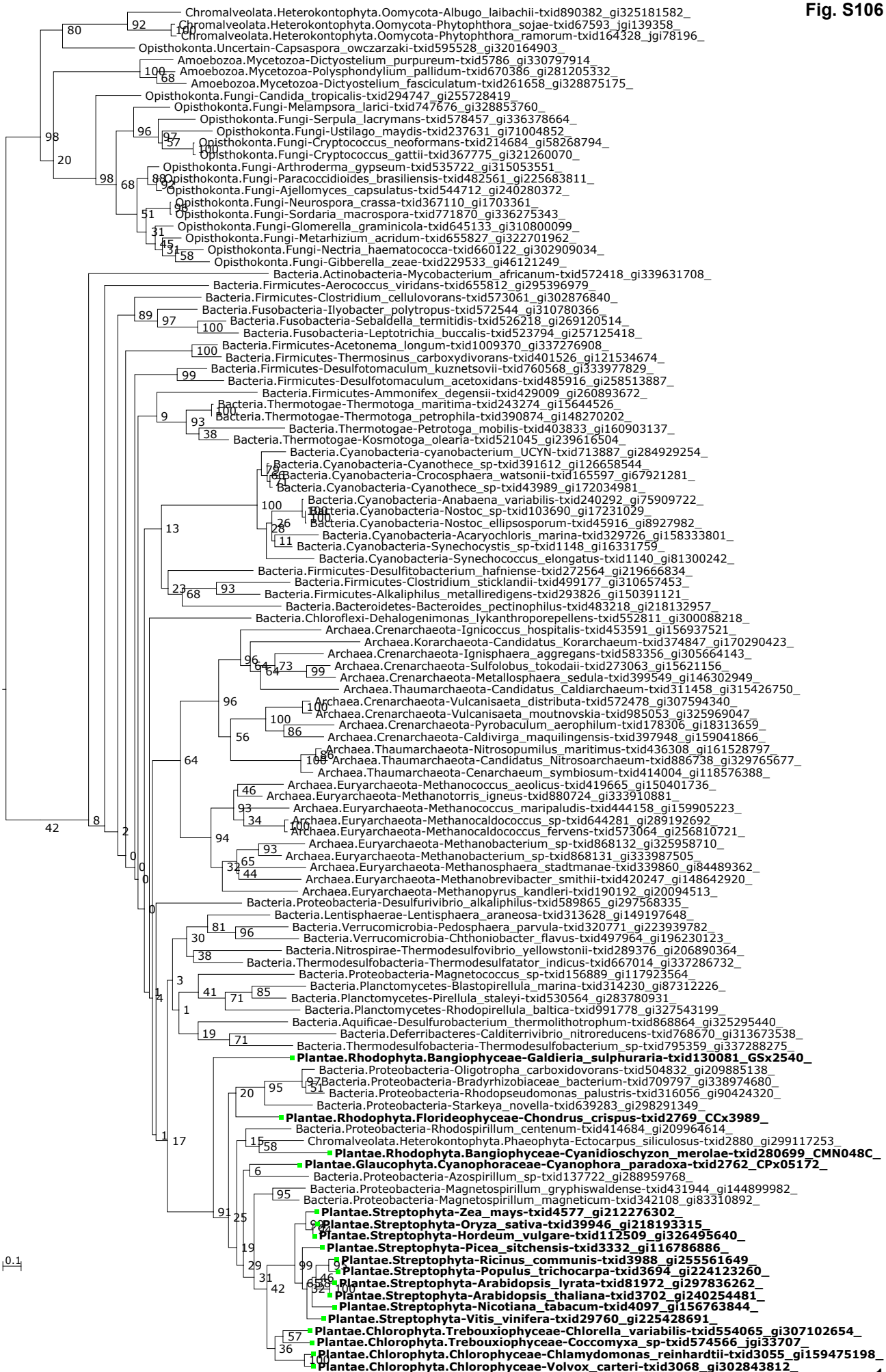


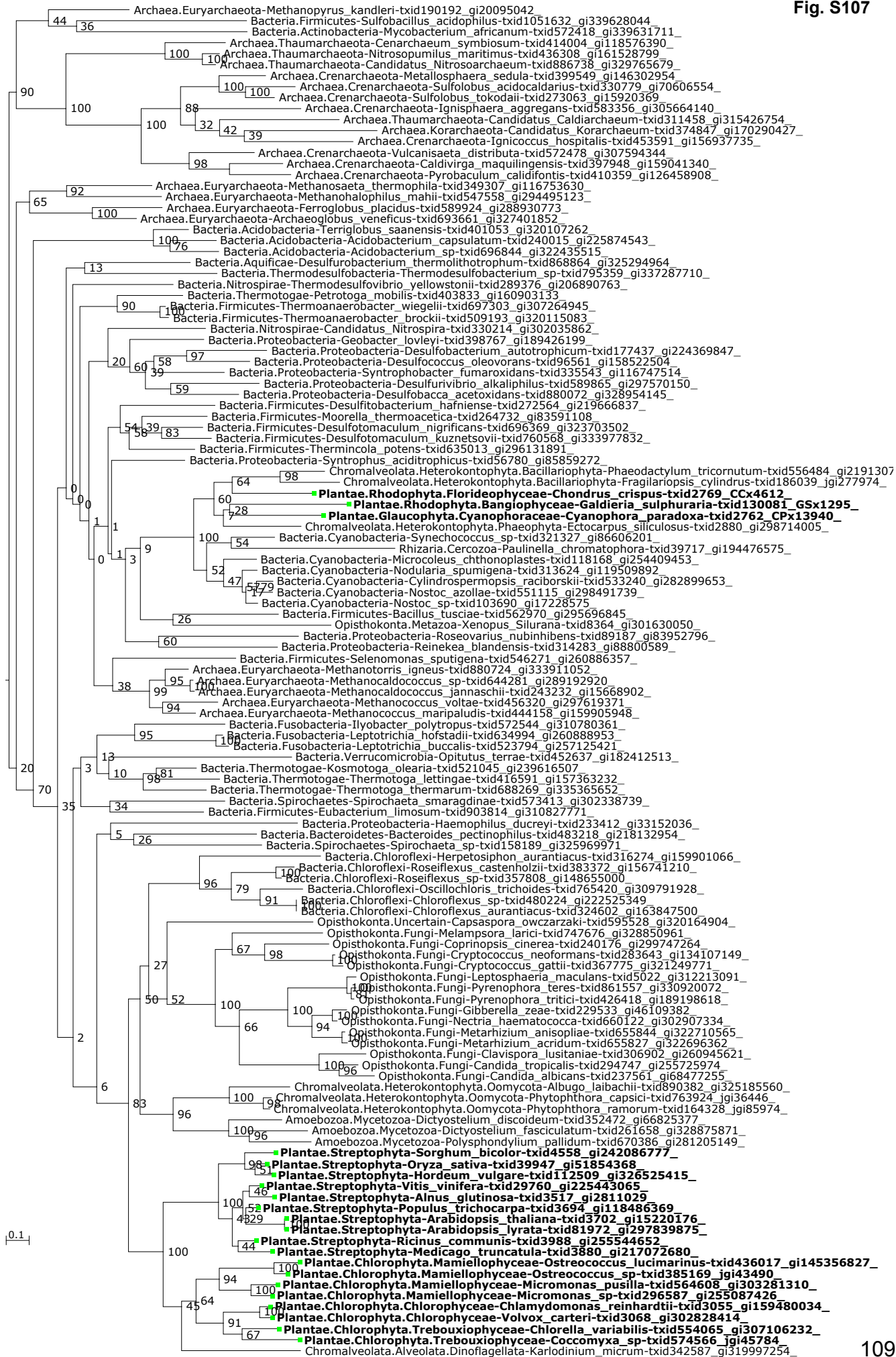




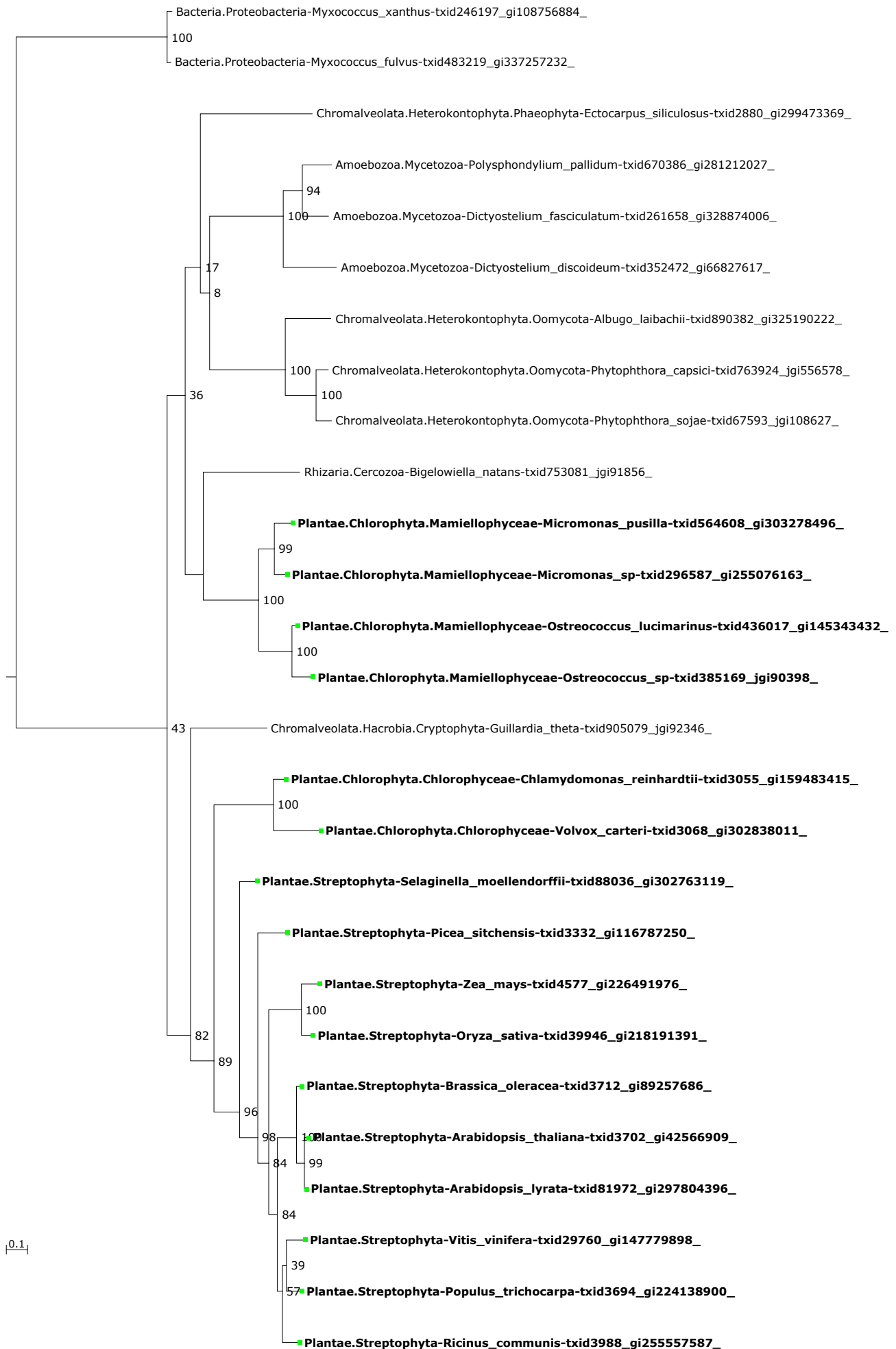
0.1







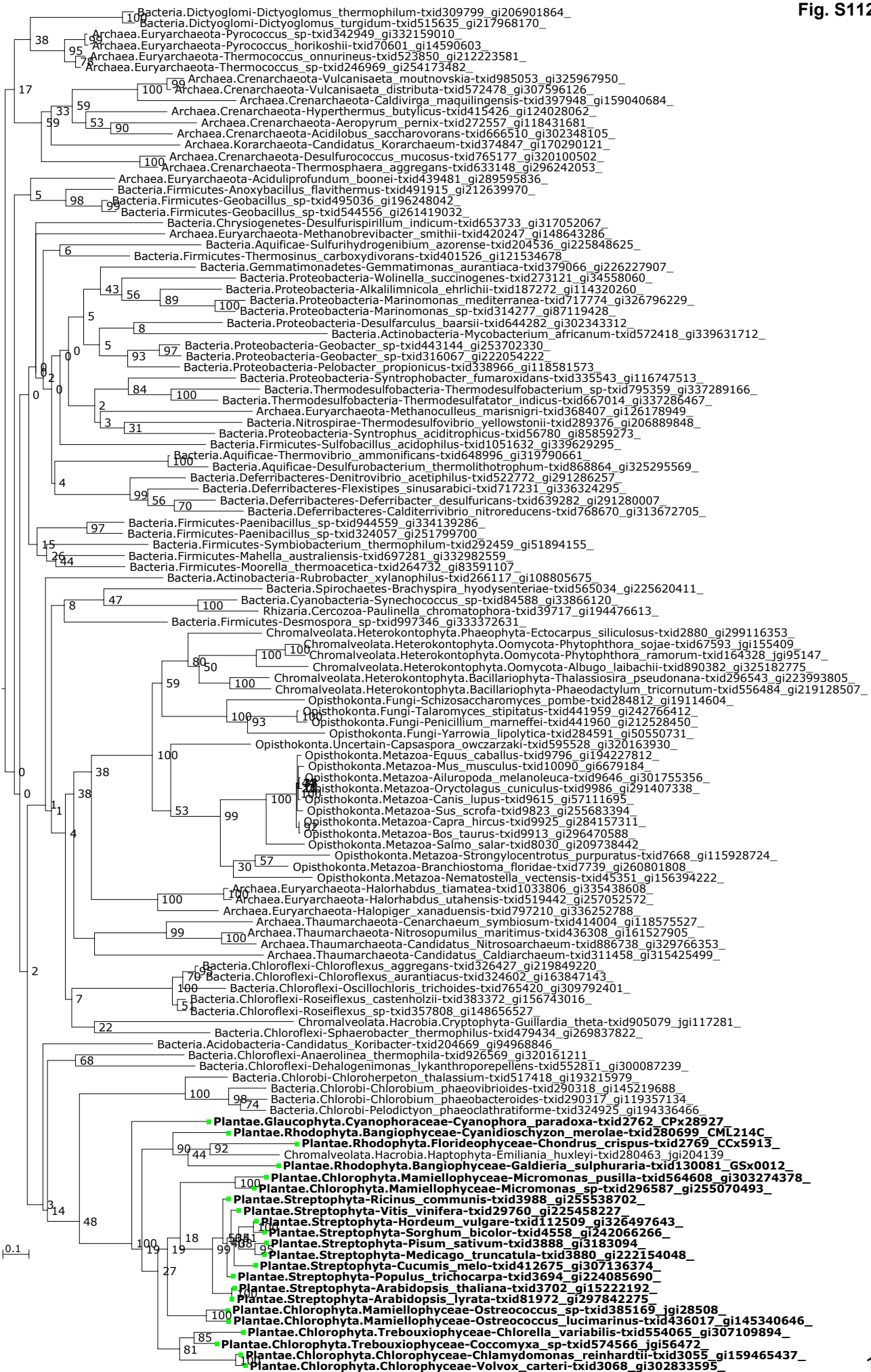




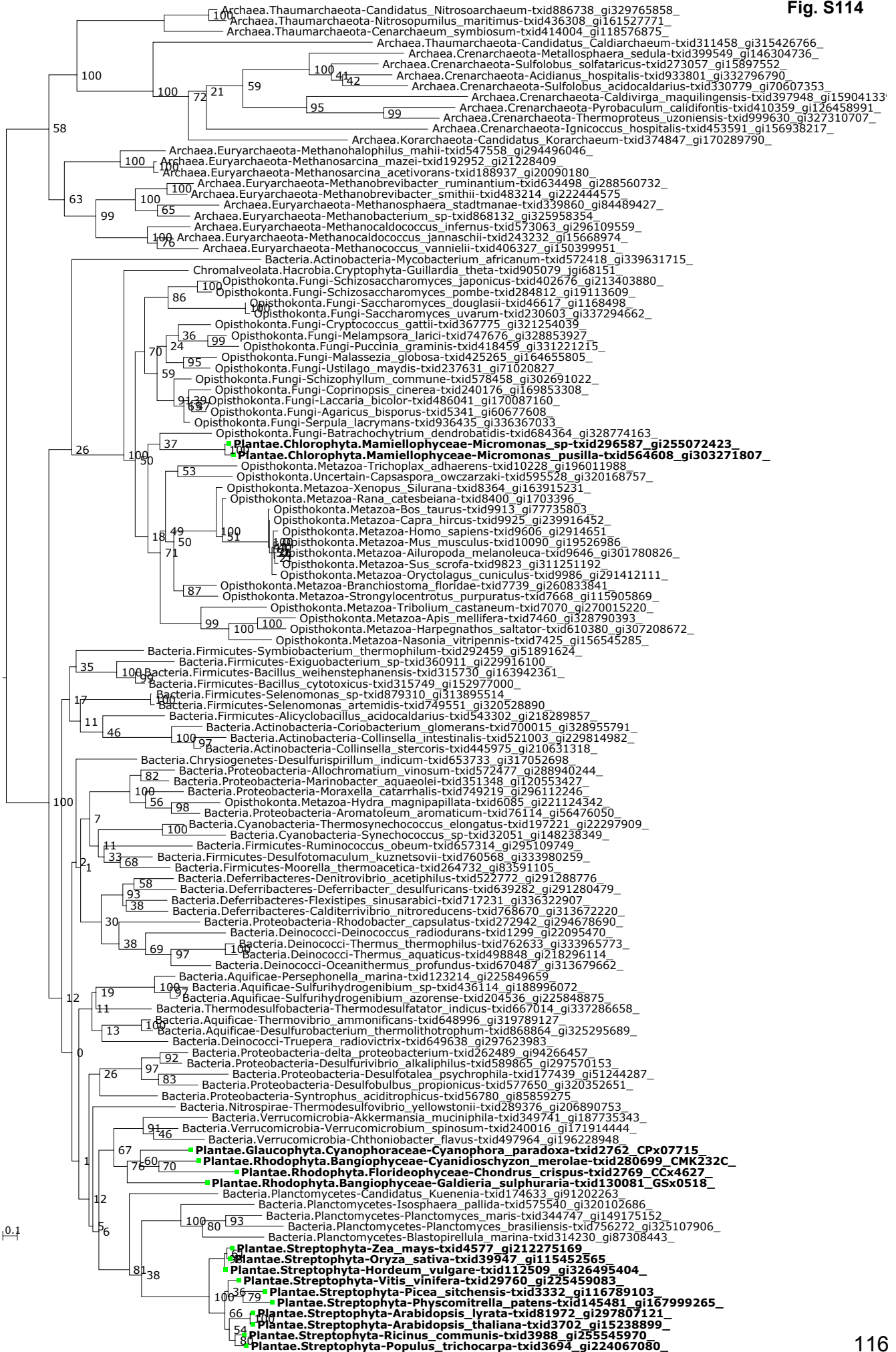




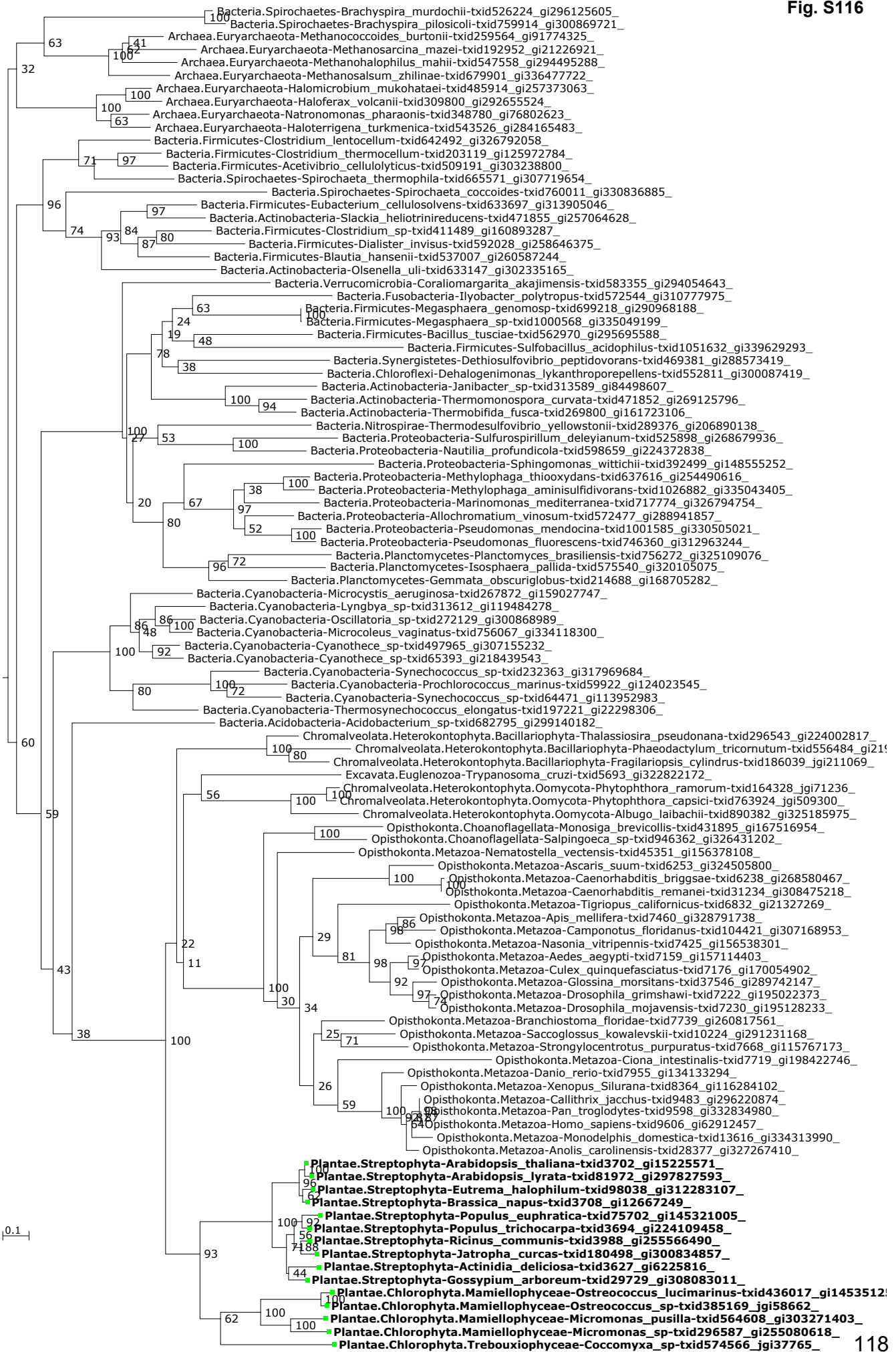


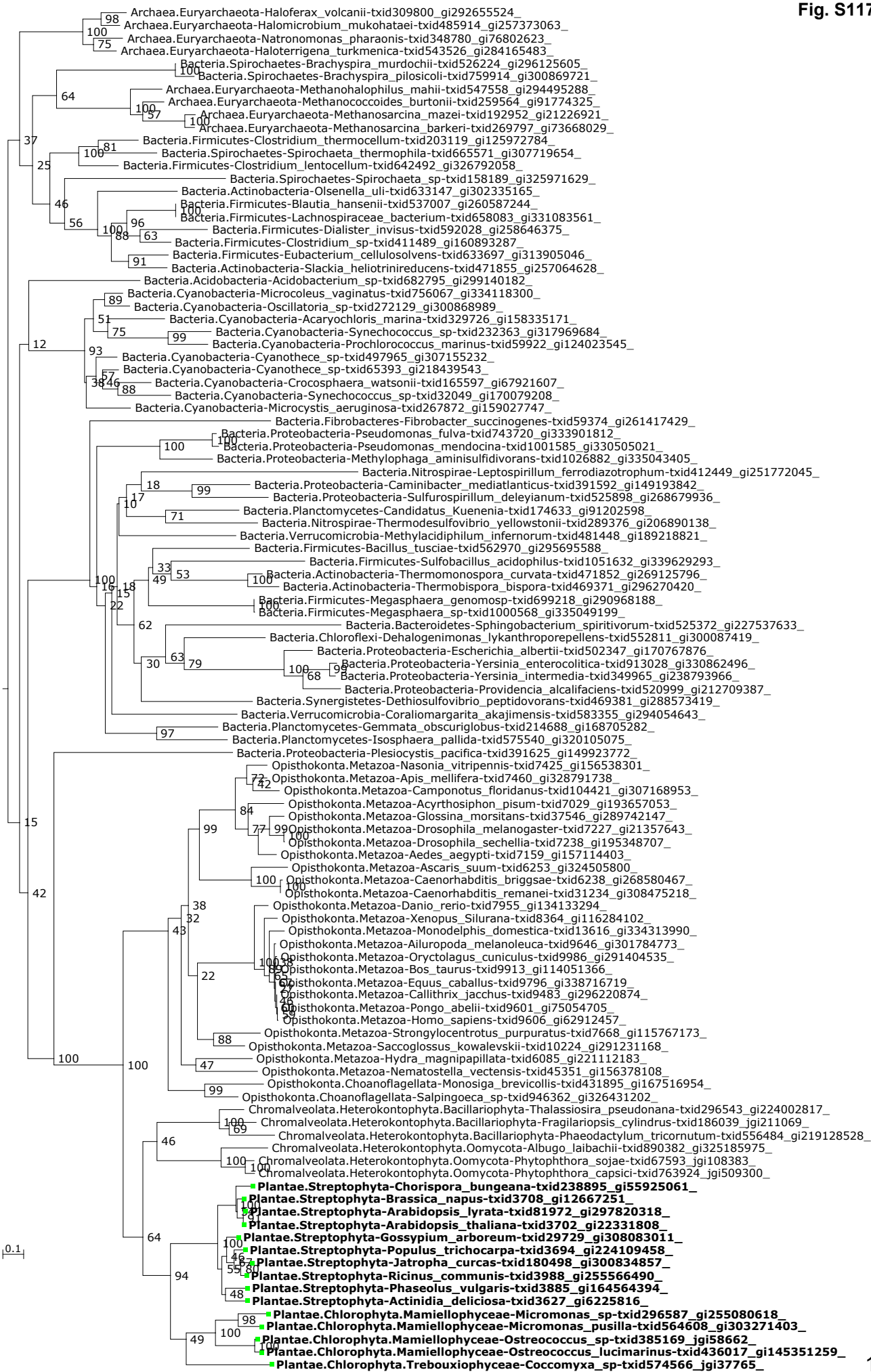


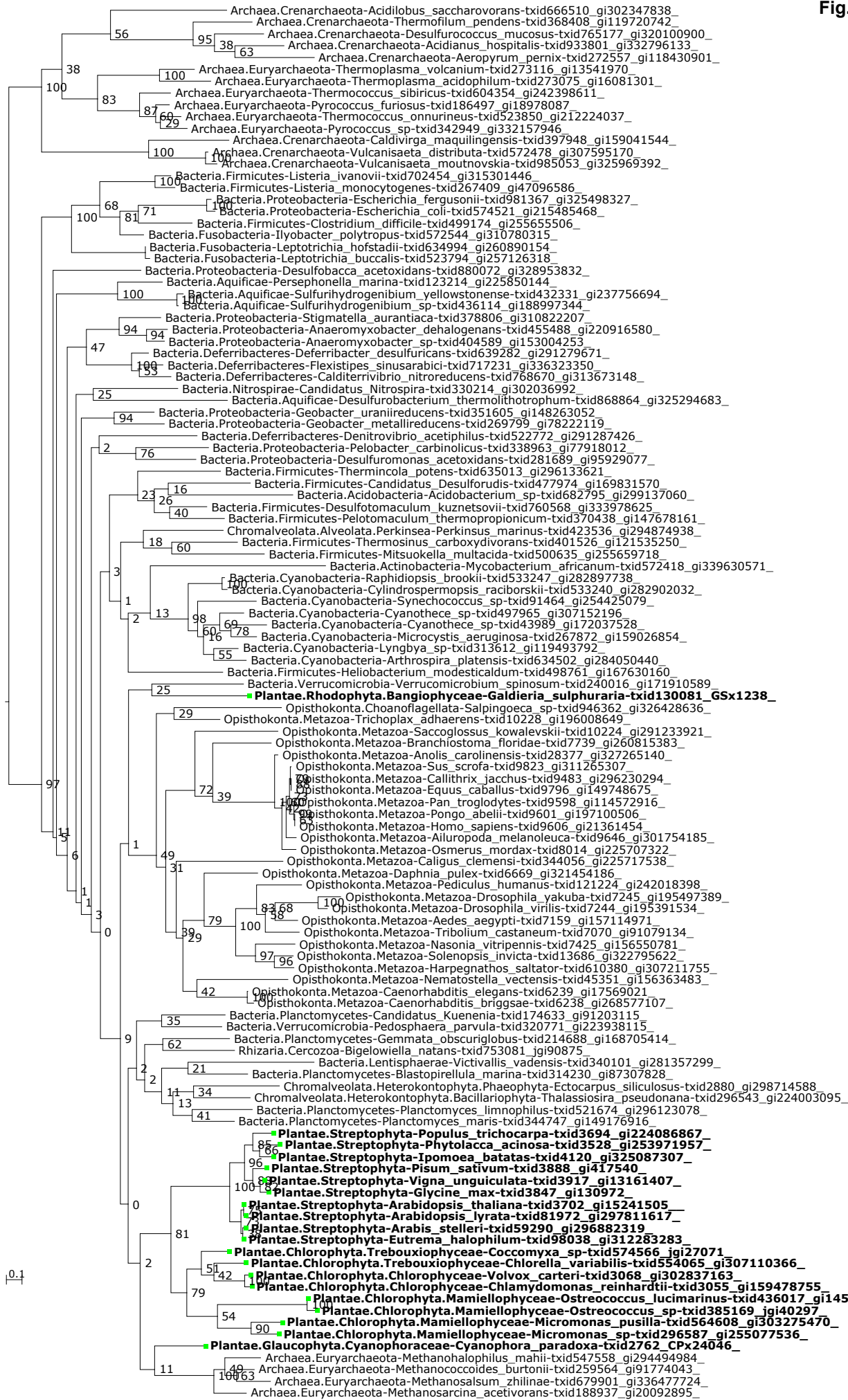




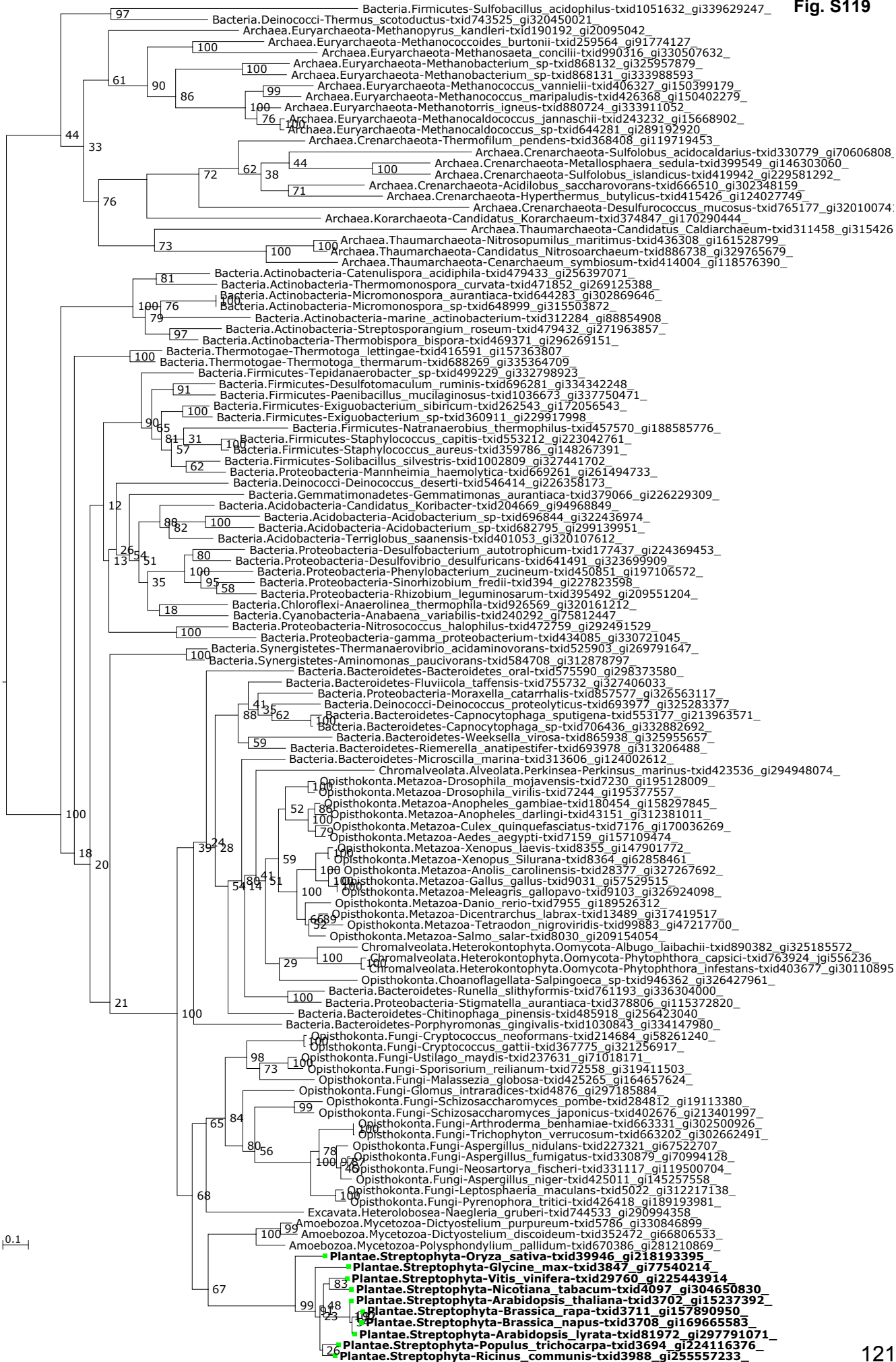


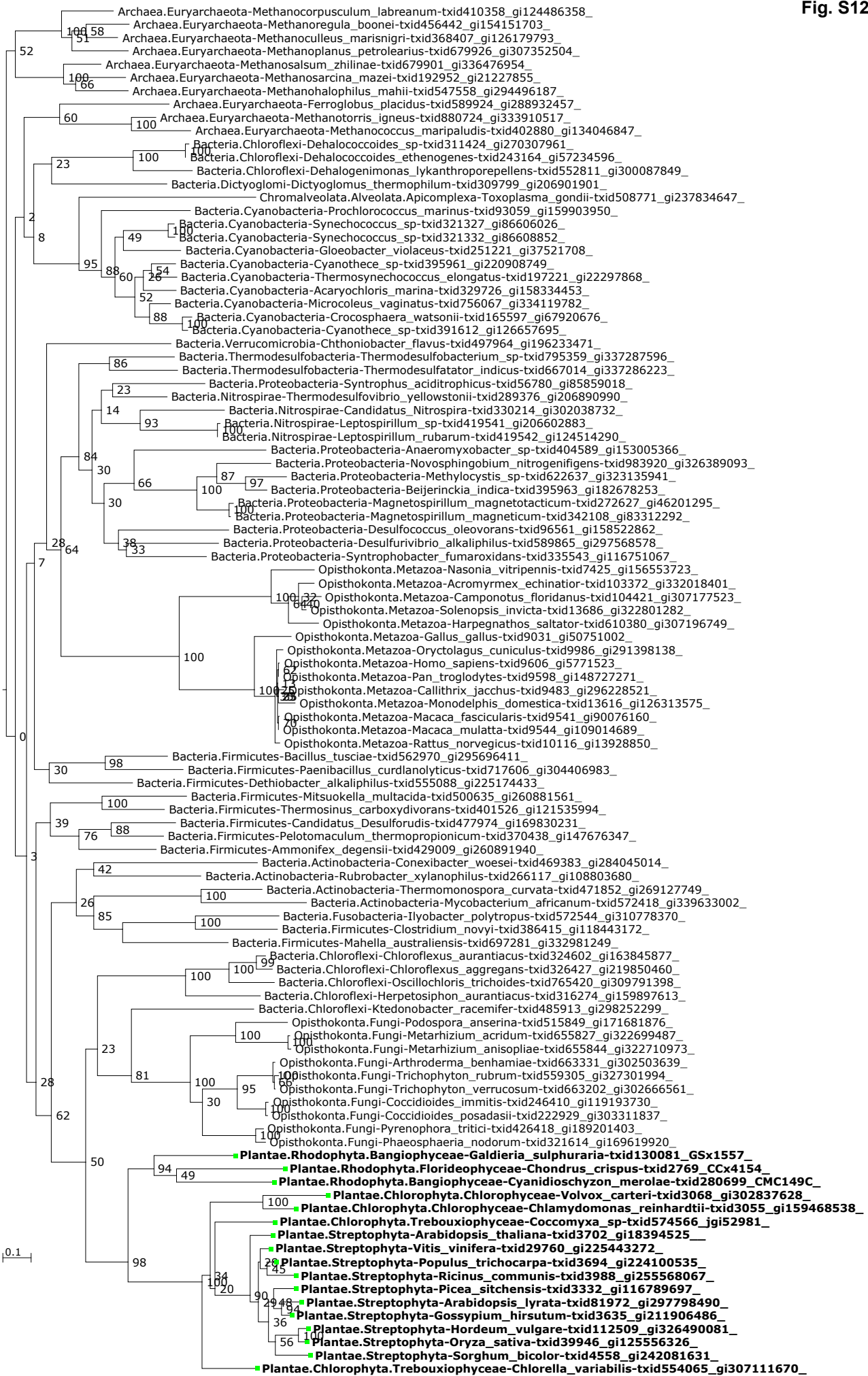




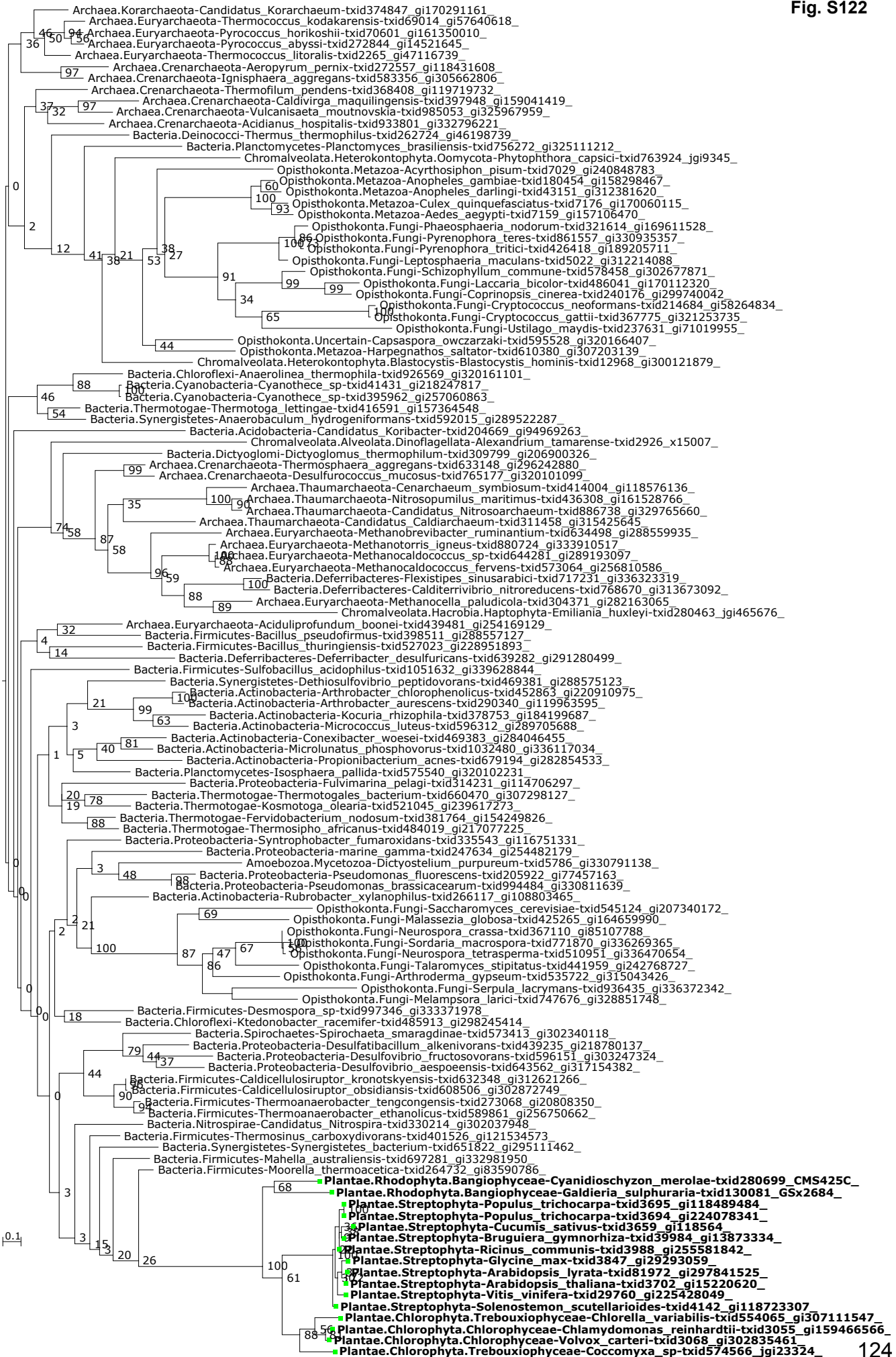


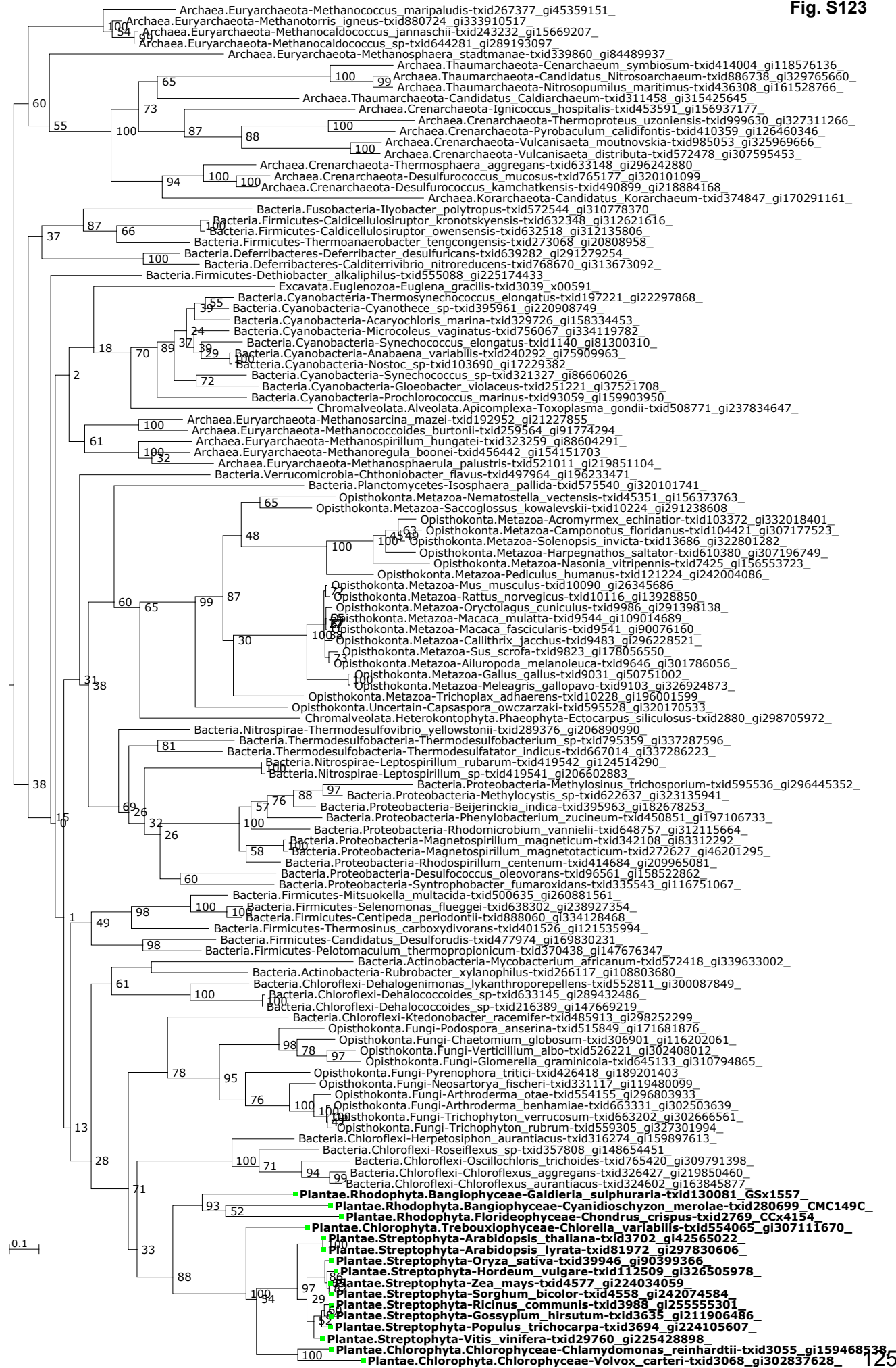


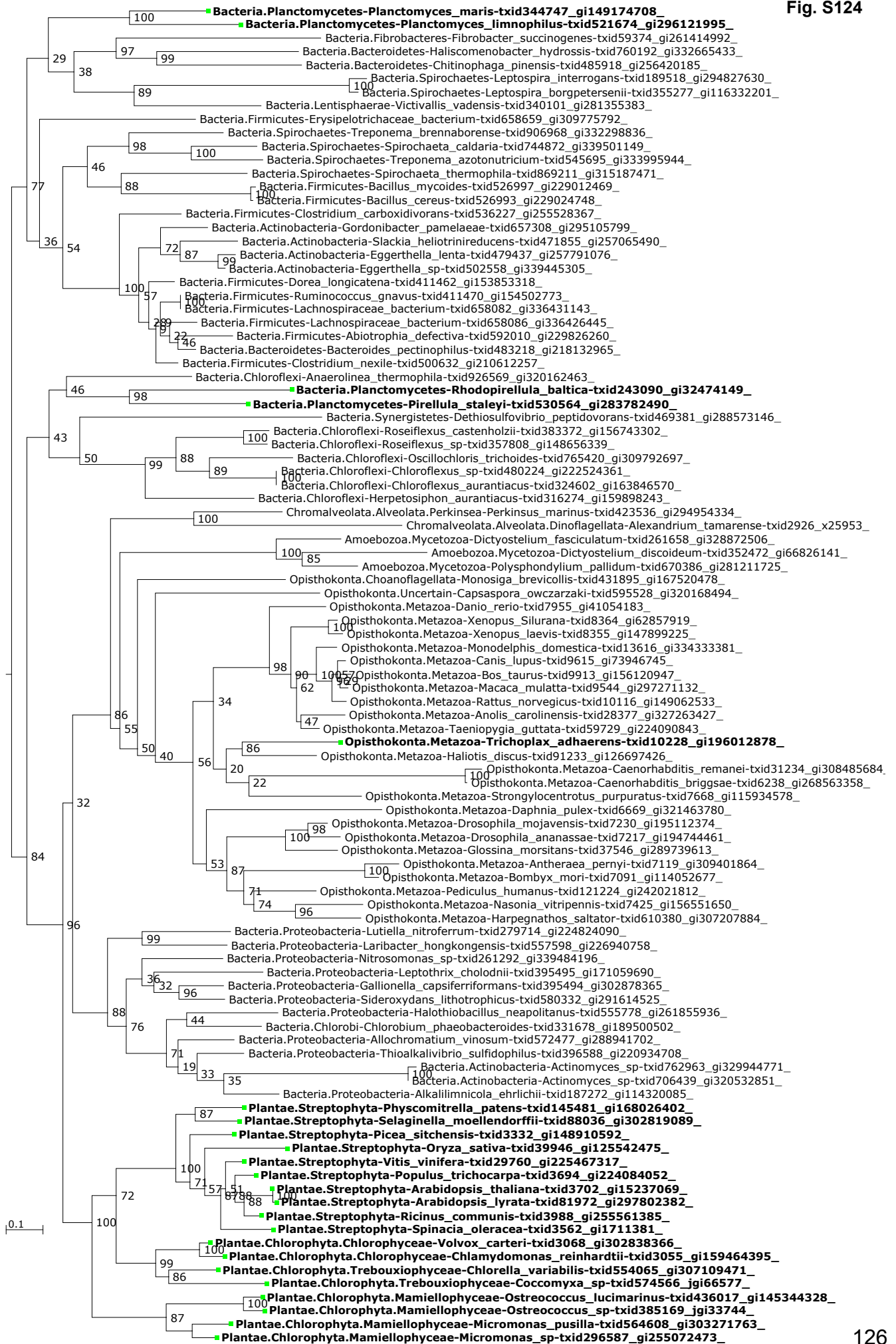


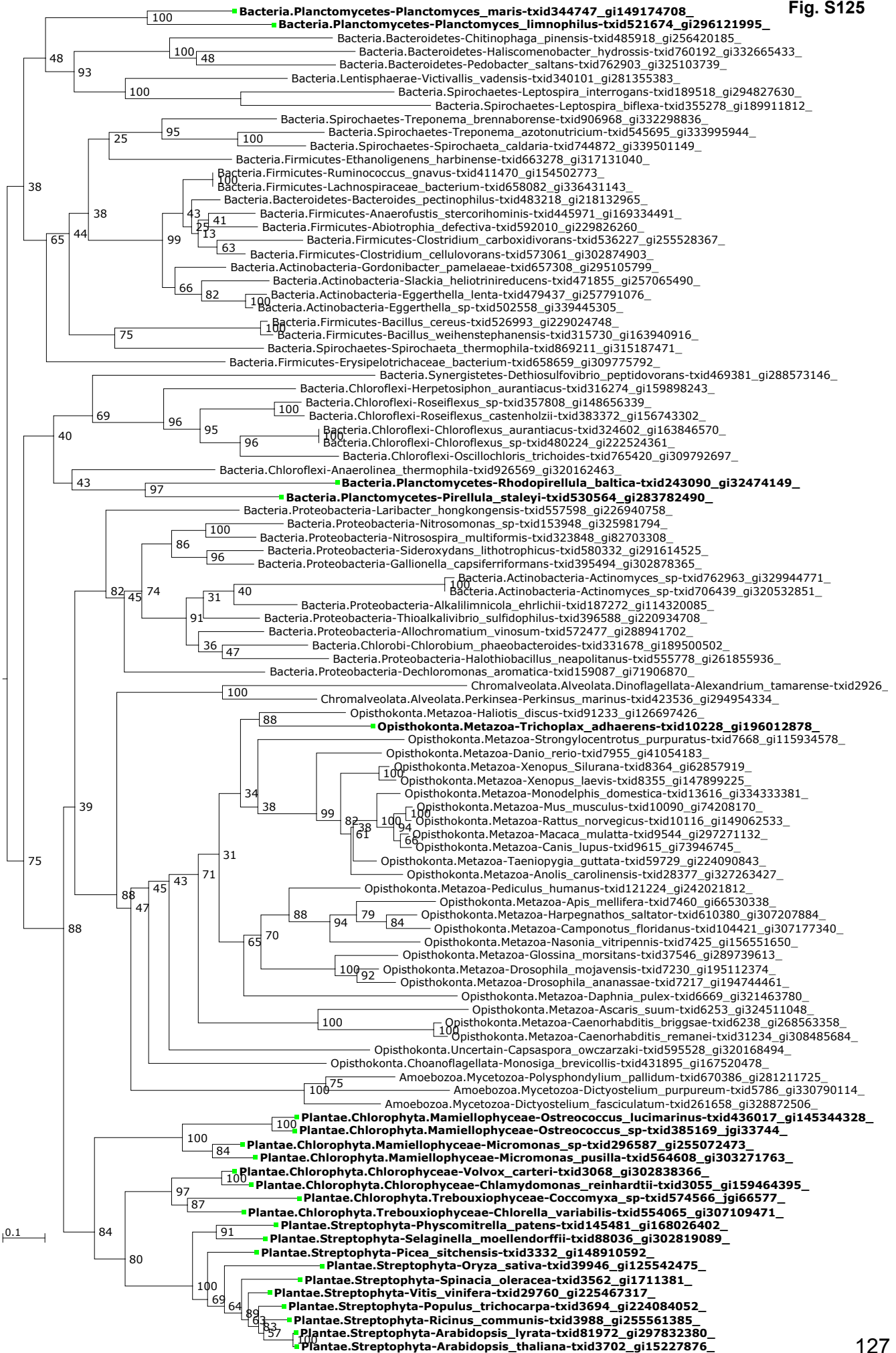


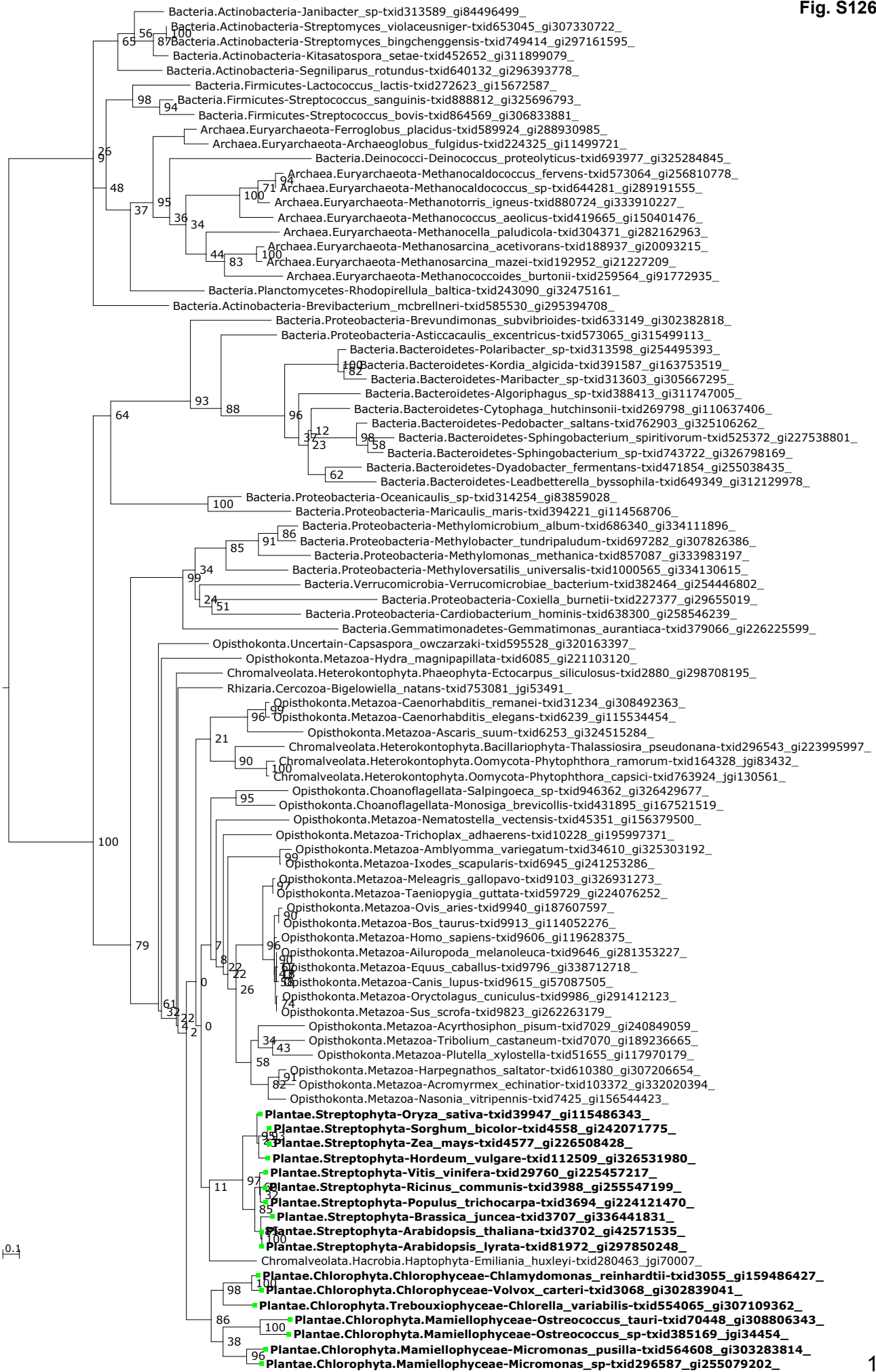




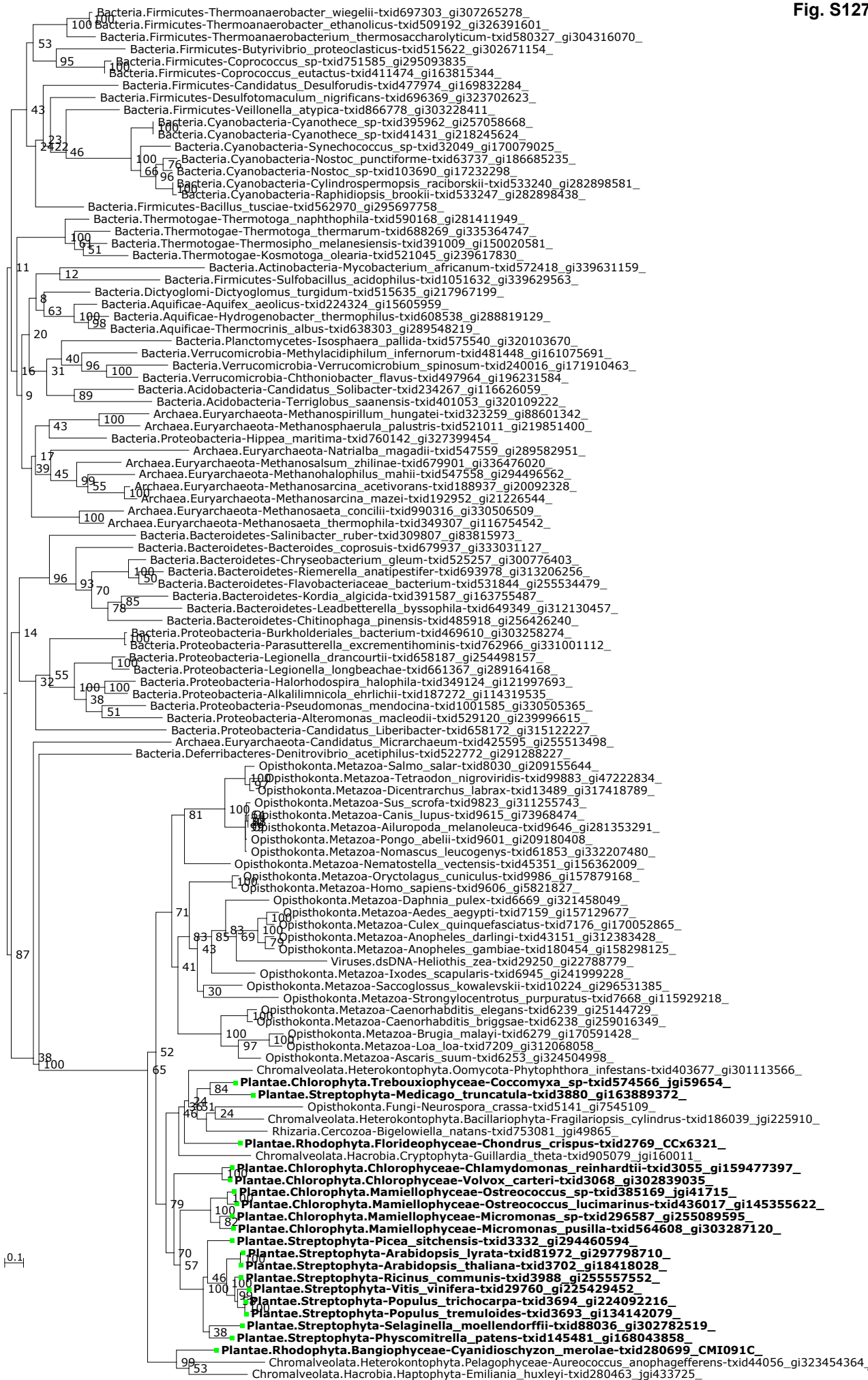


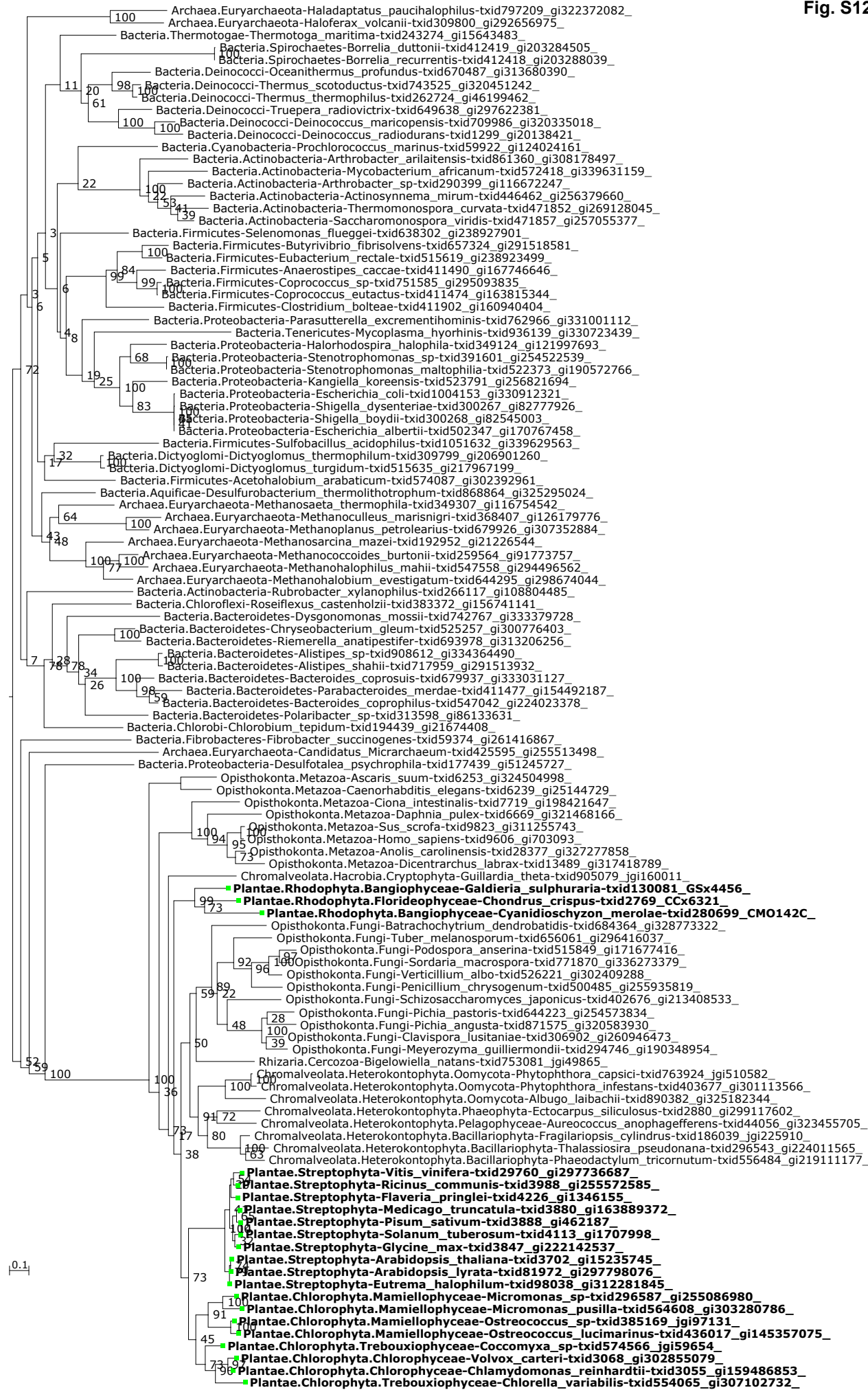












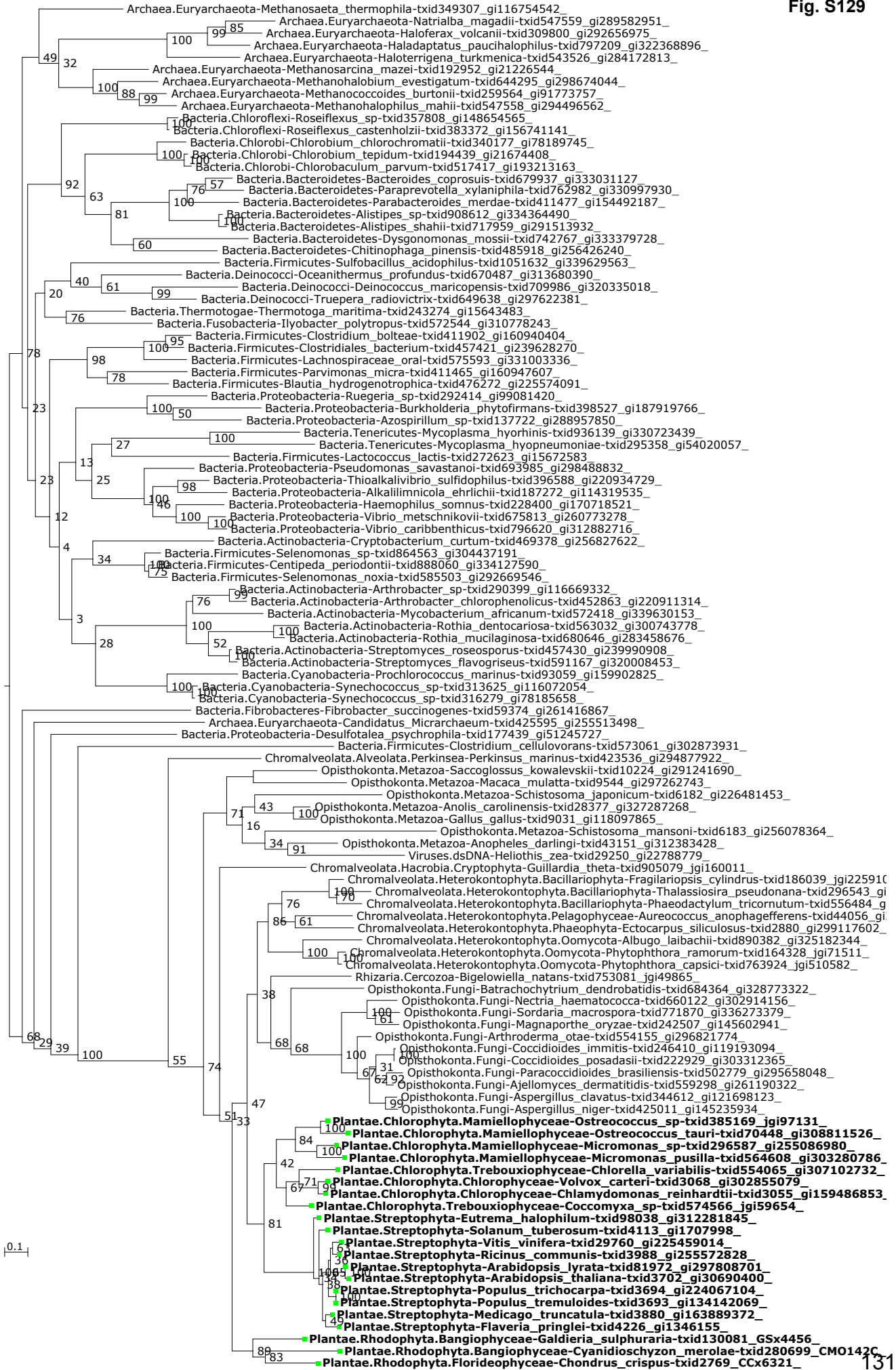
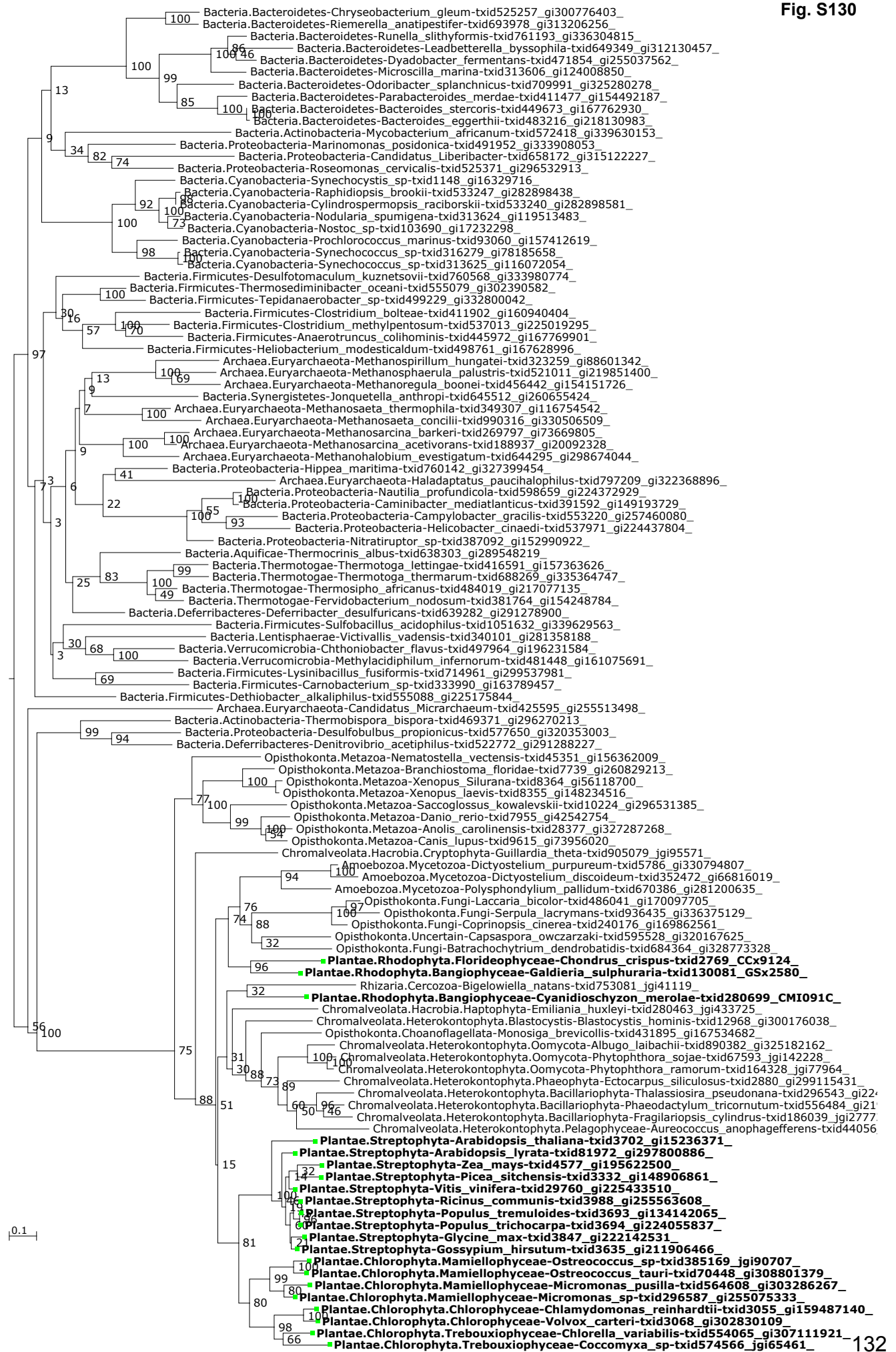
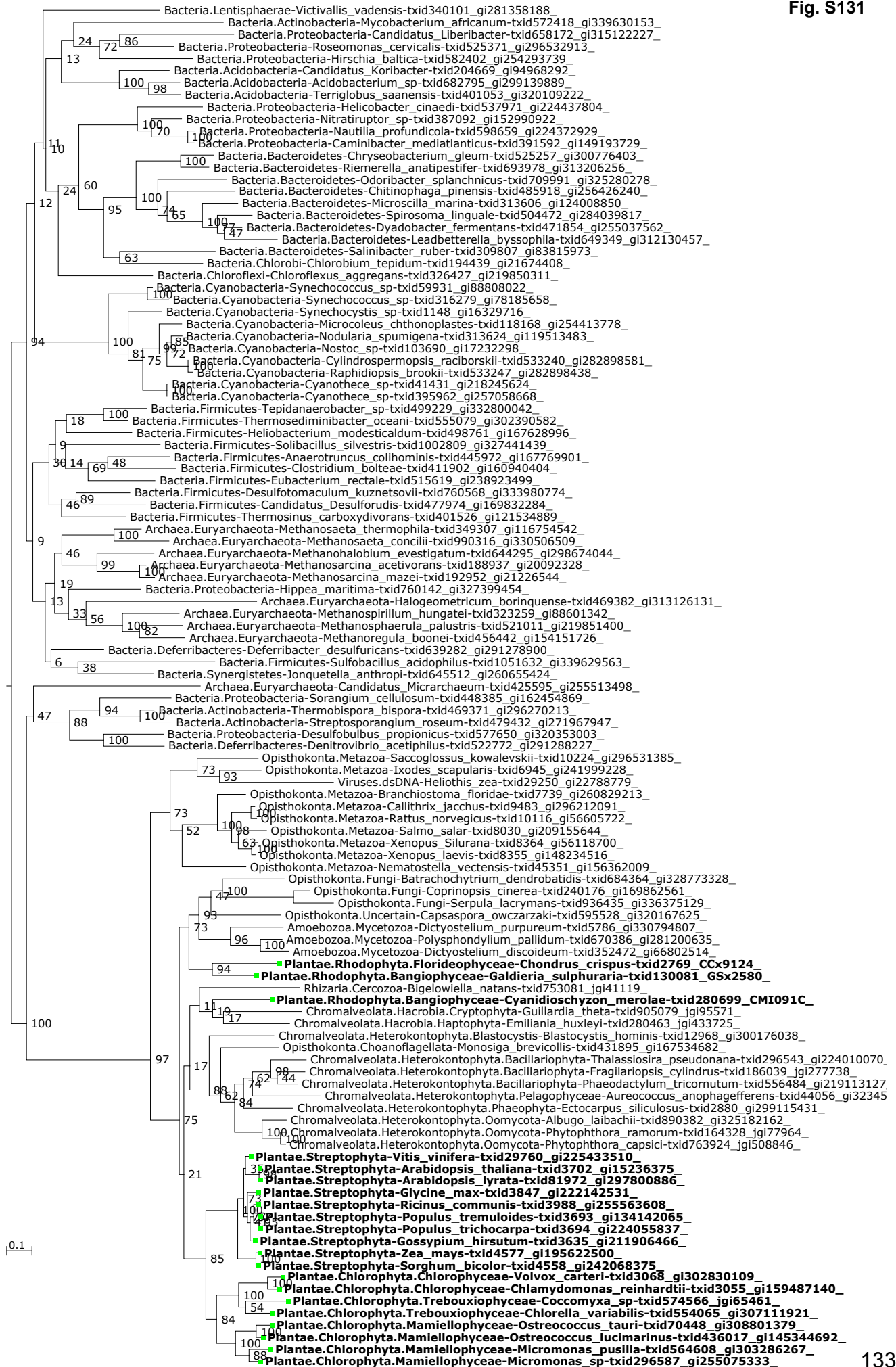
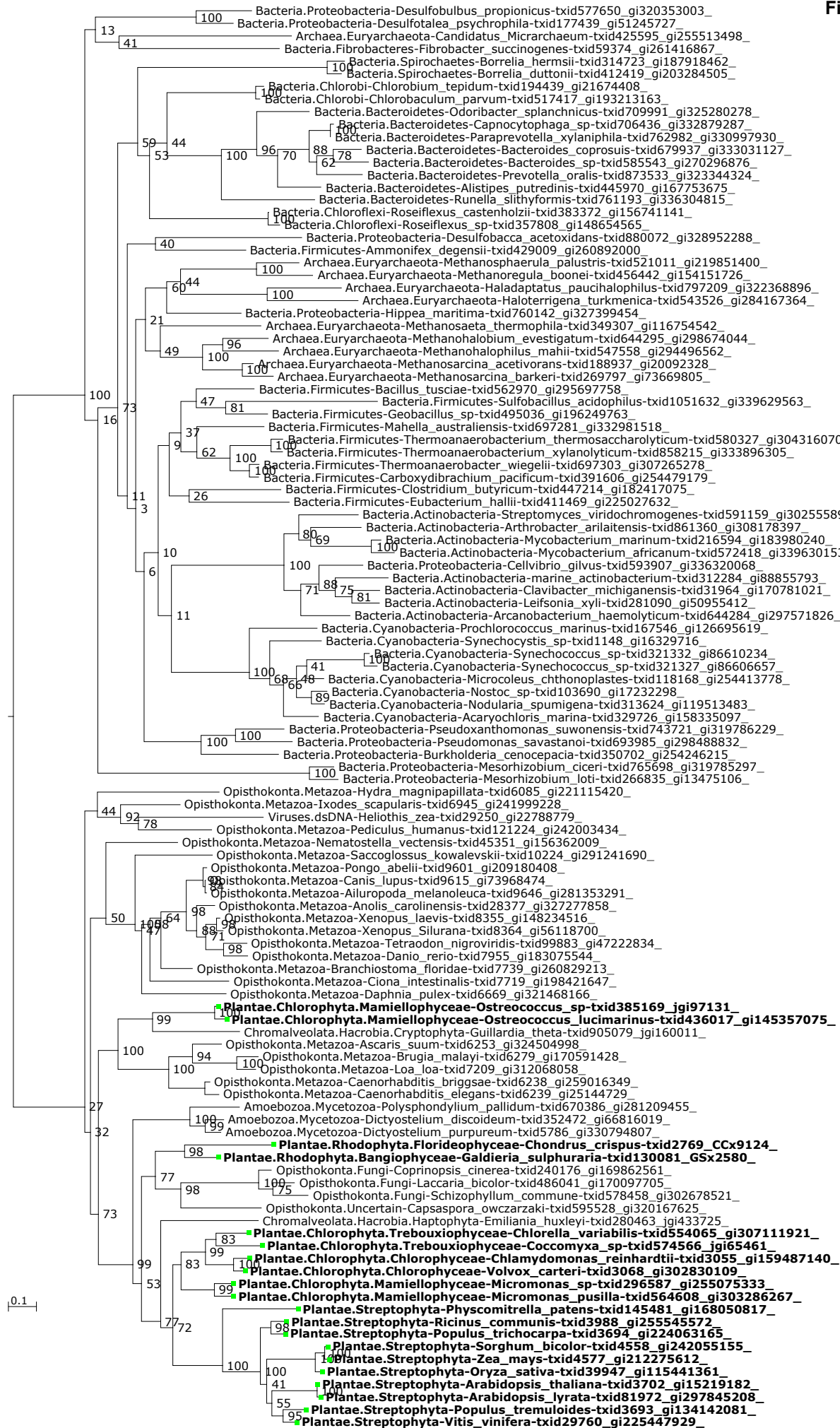
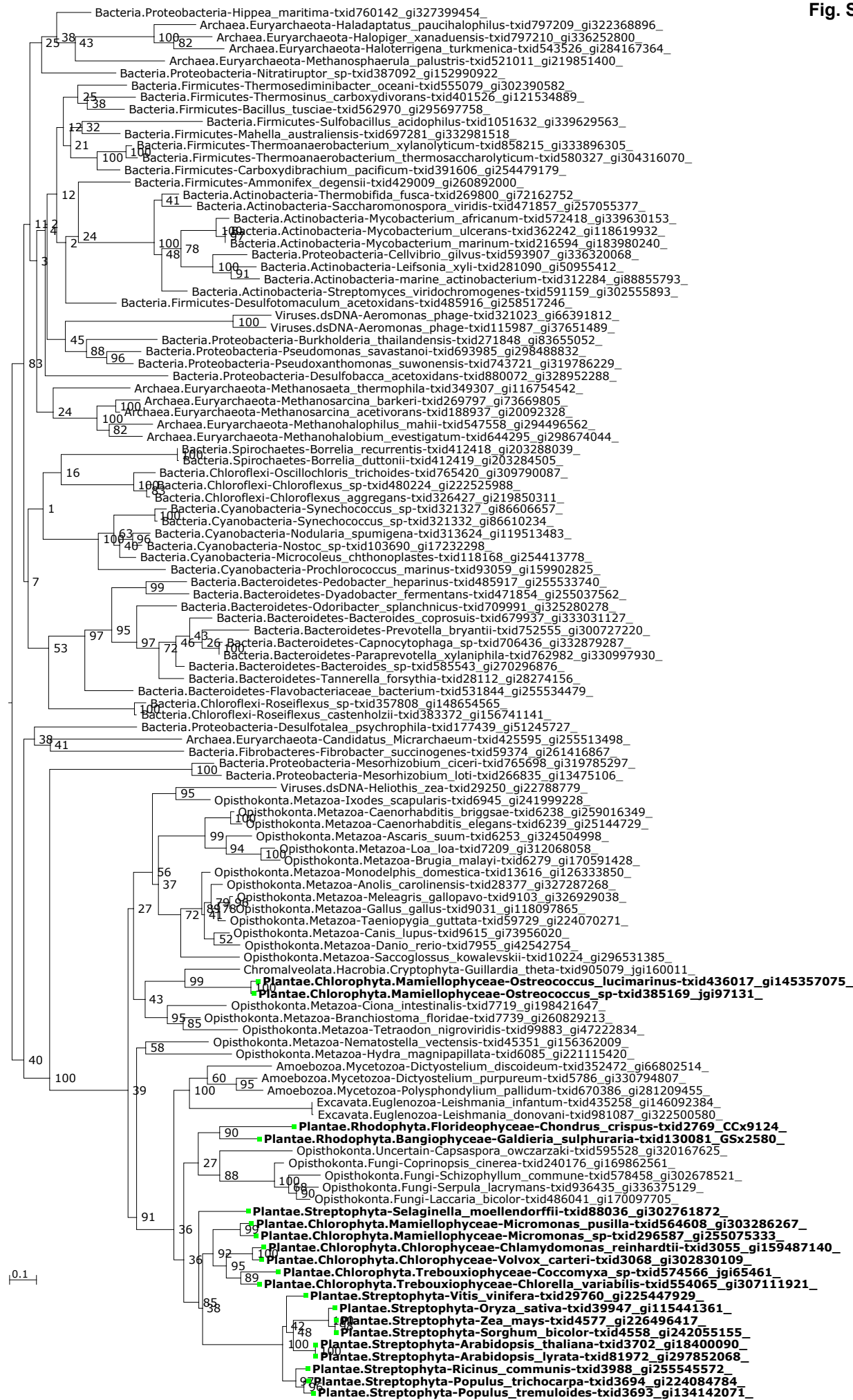


Fig. S130



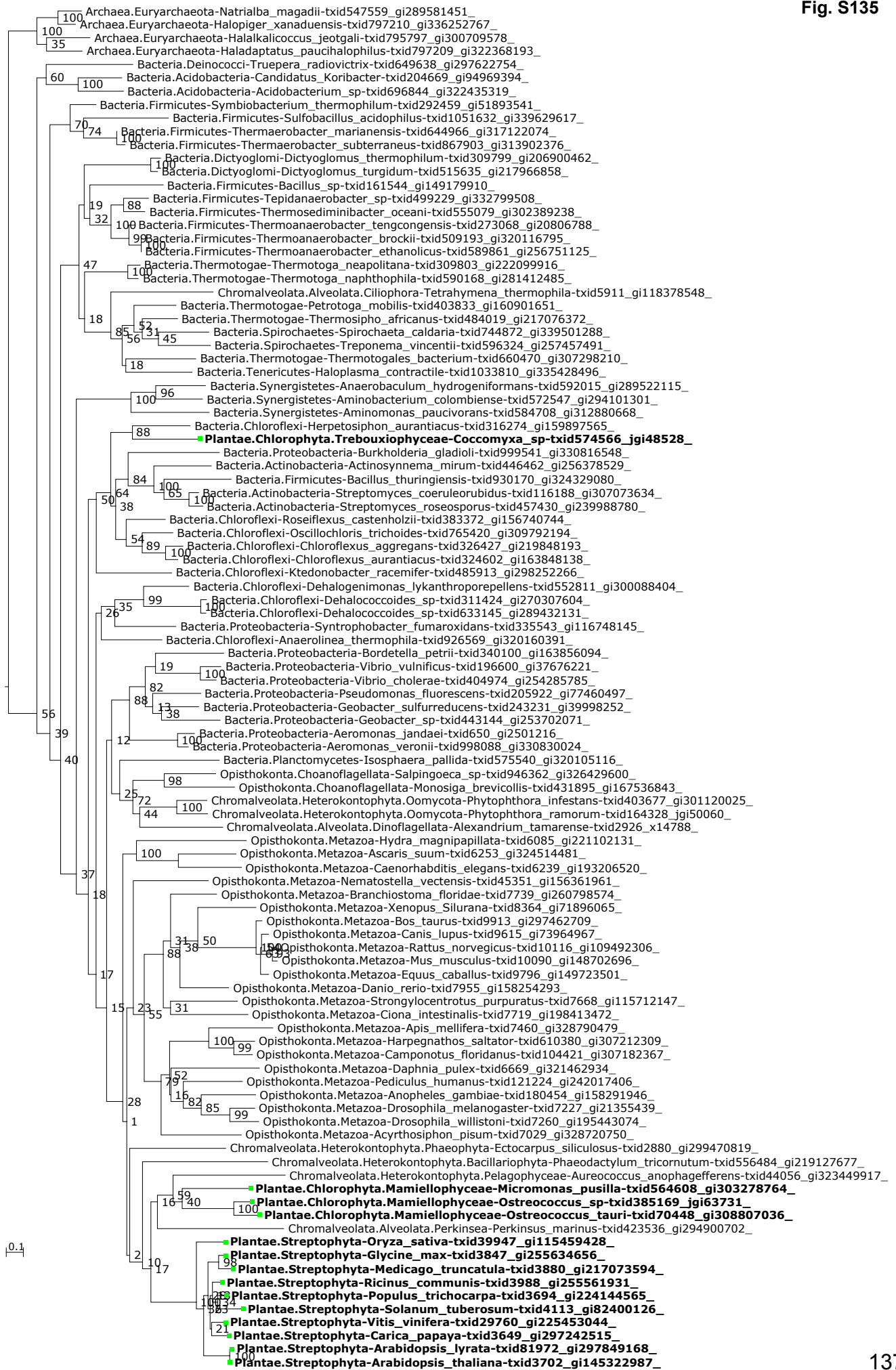


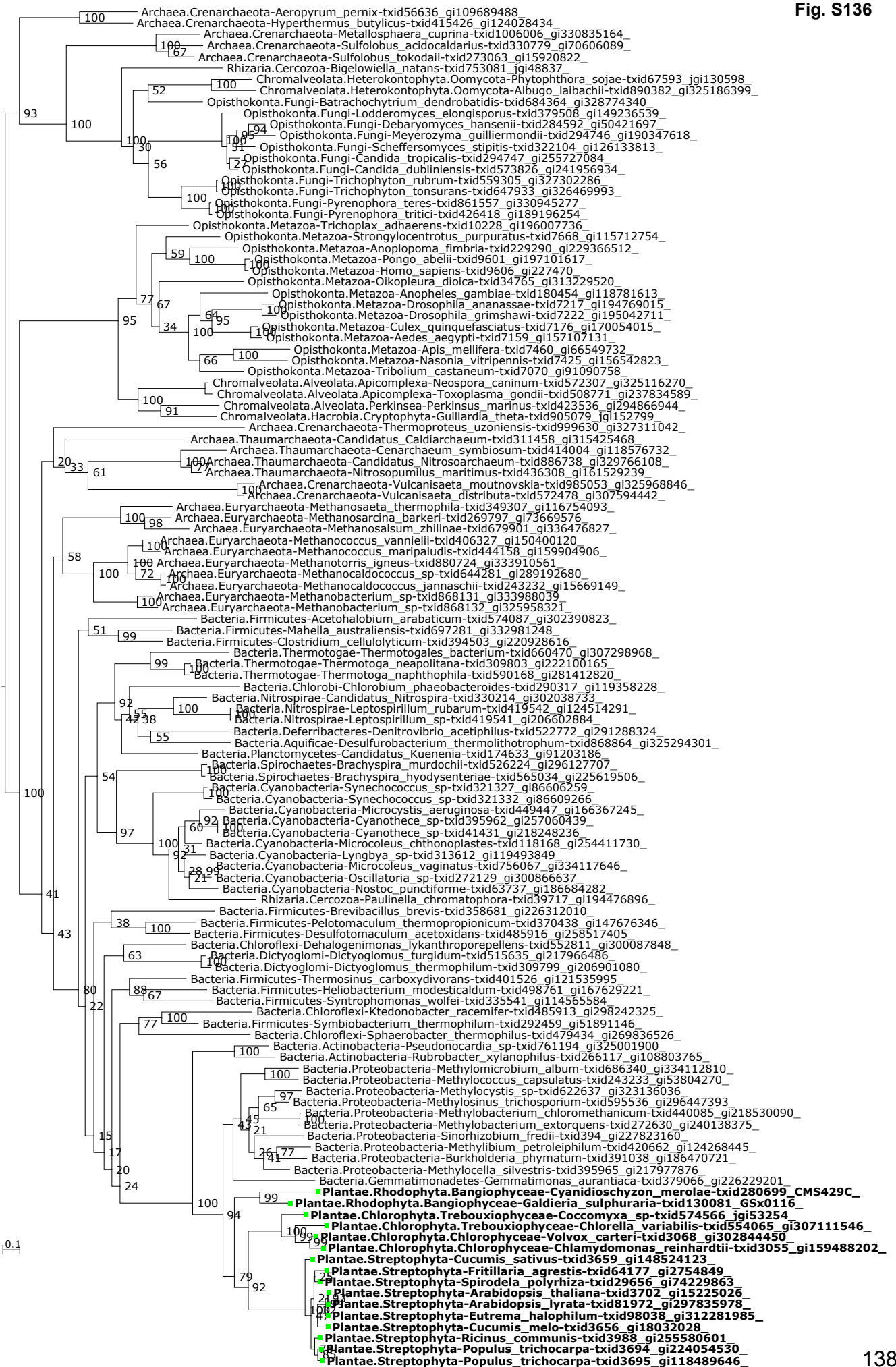




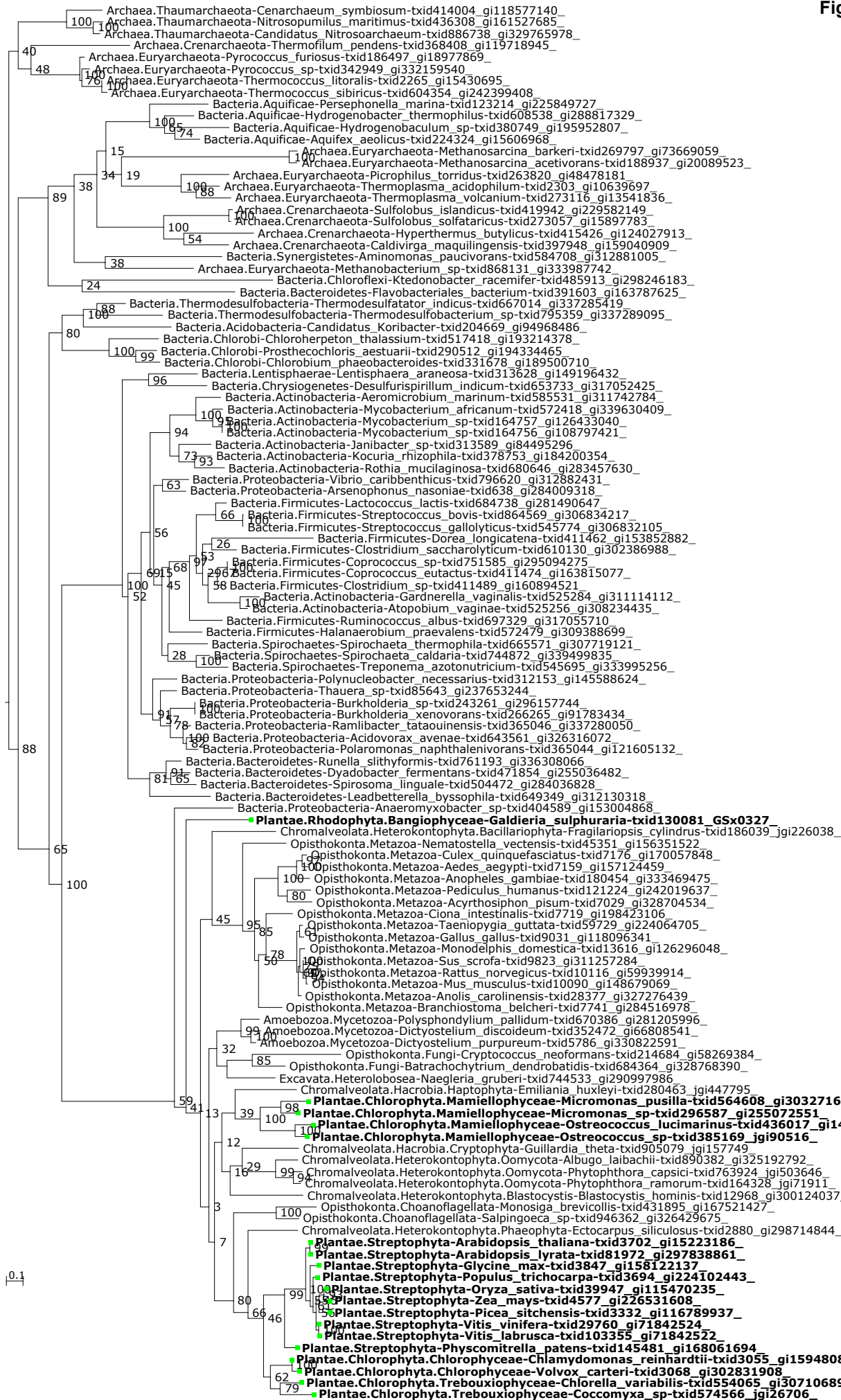




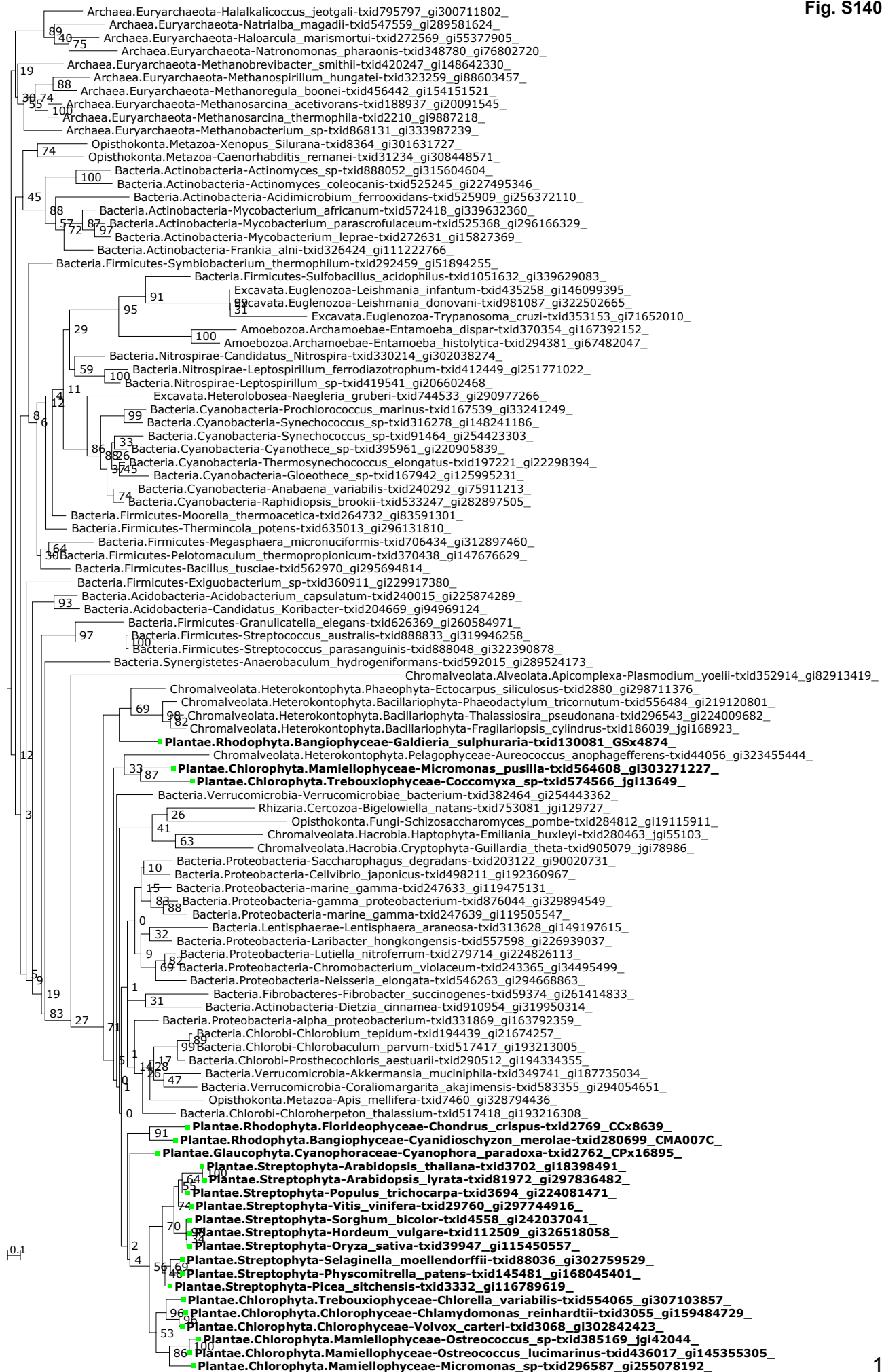


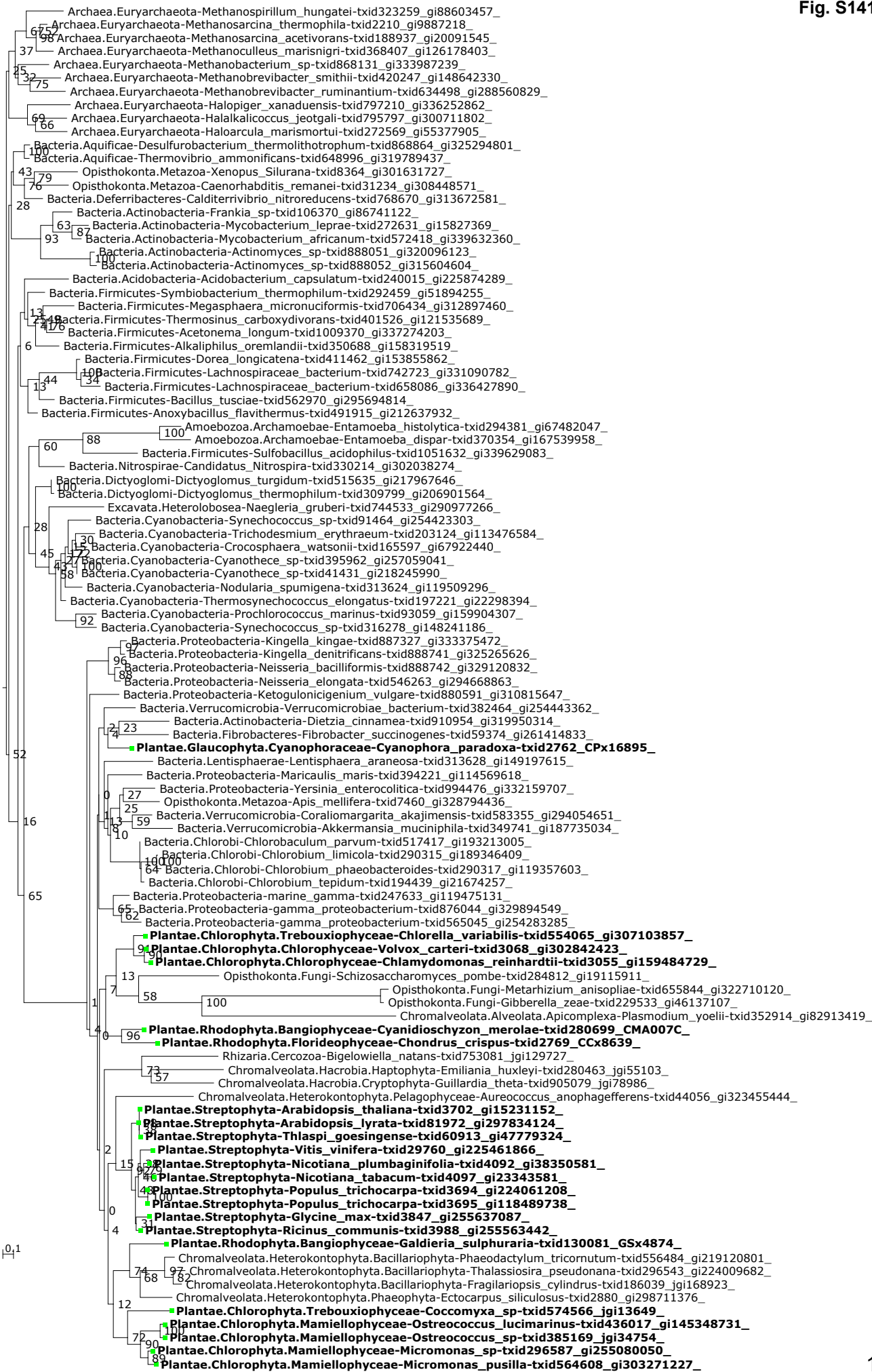


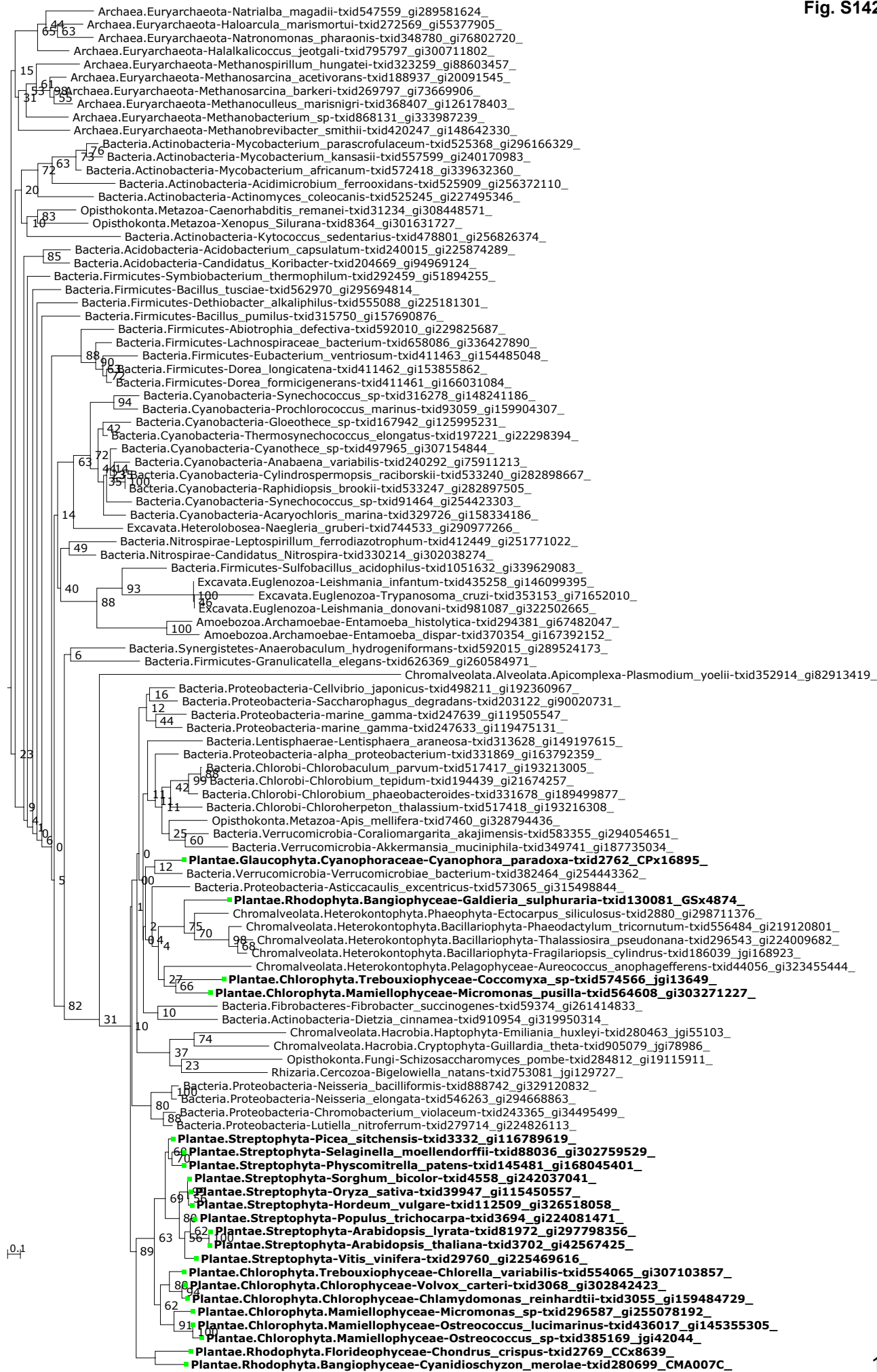




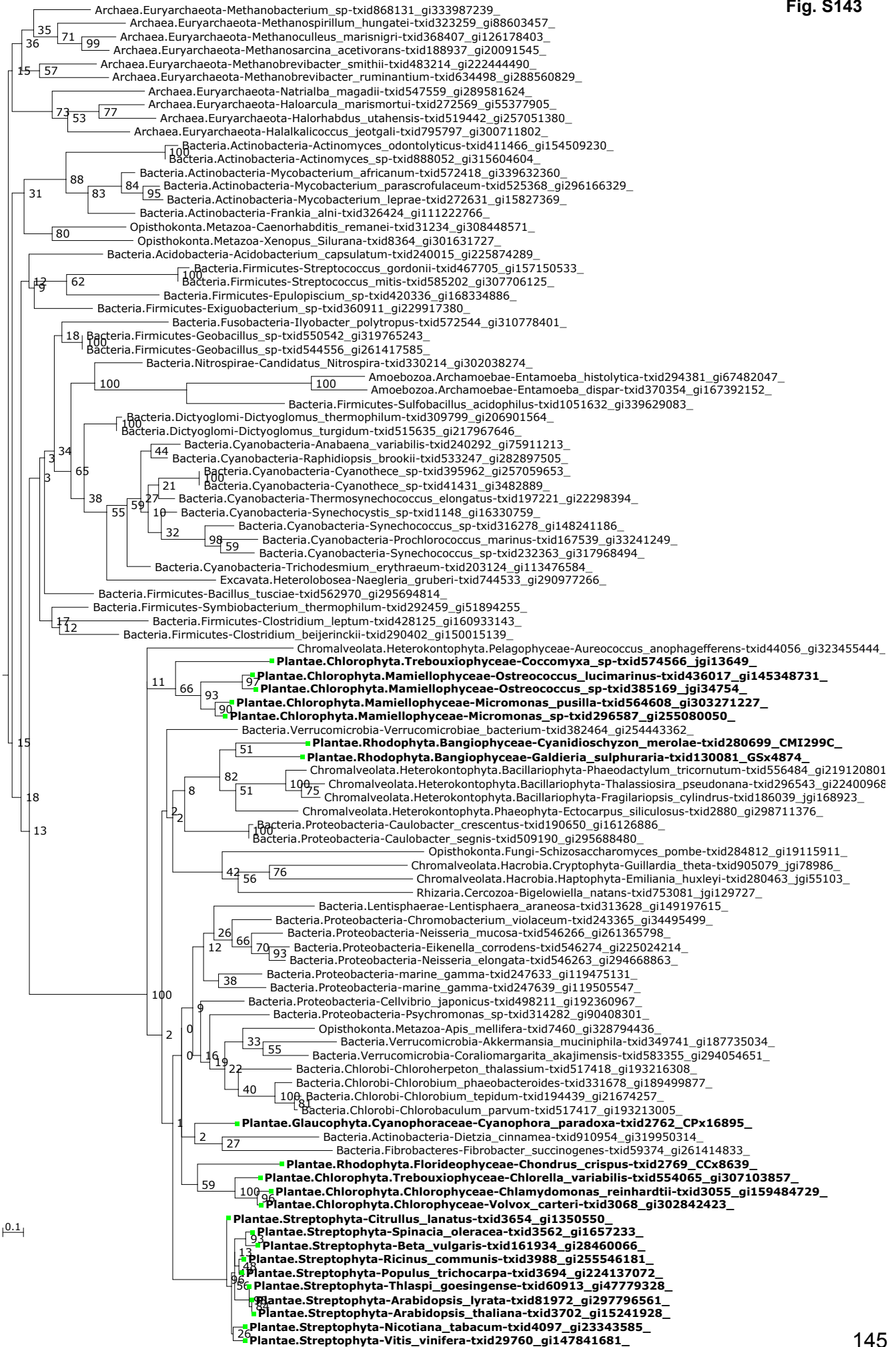


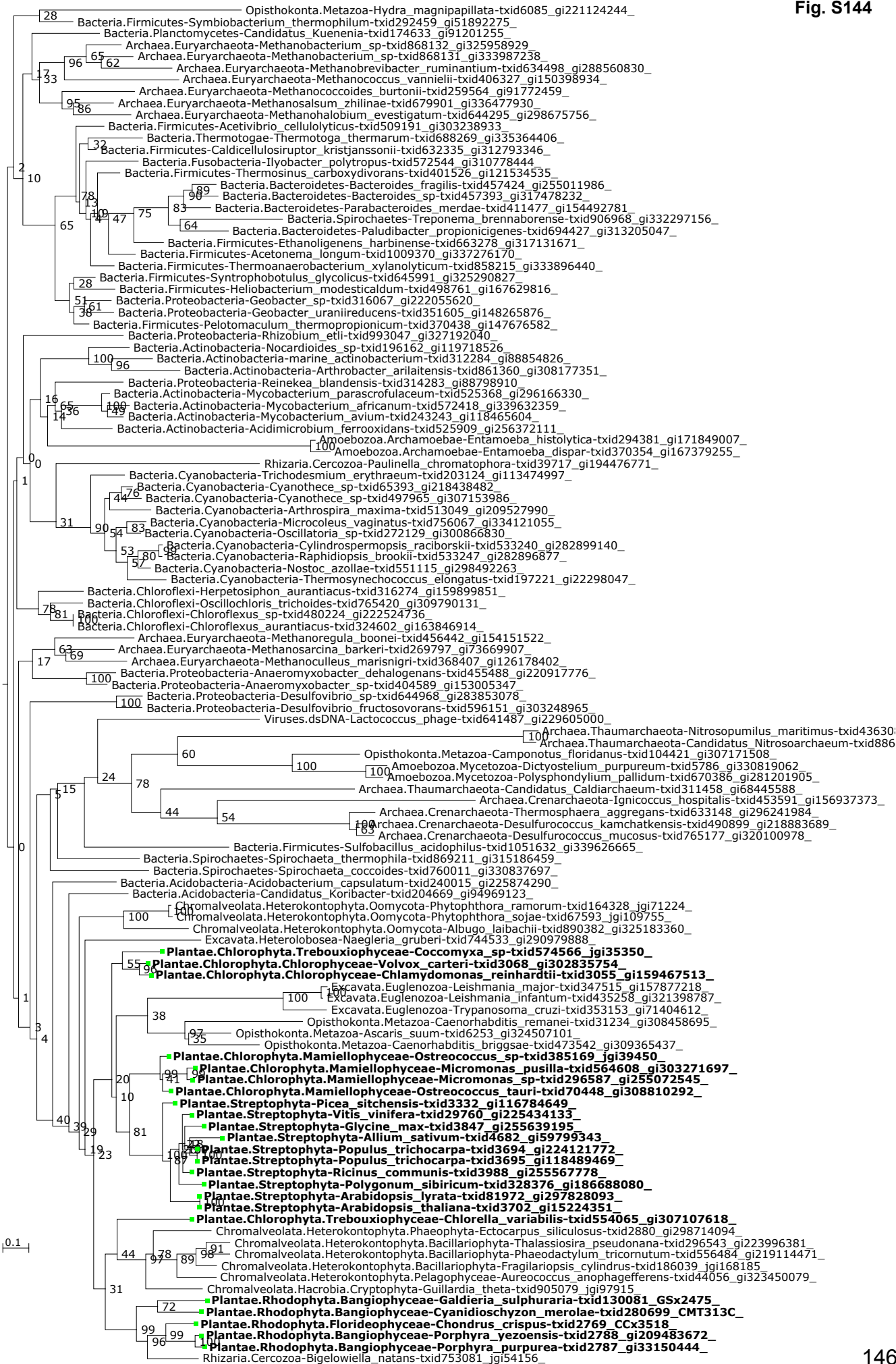


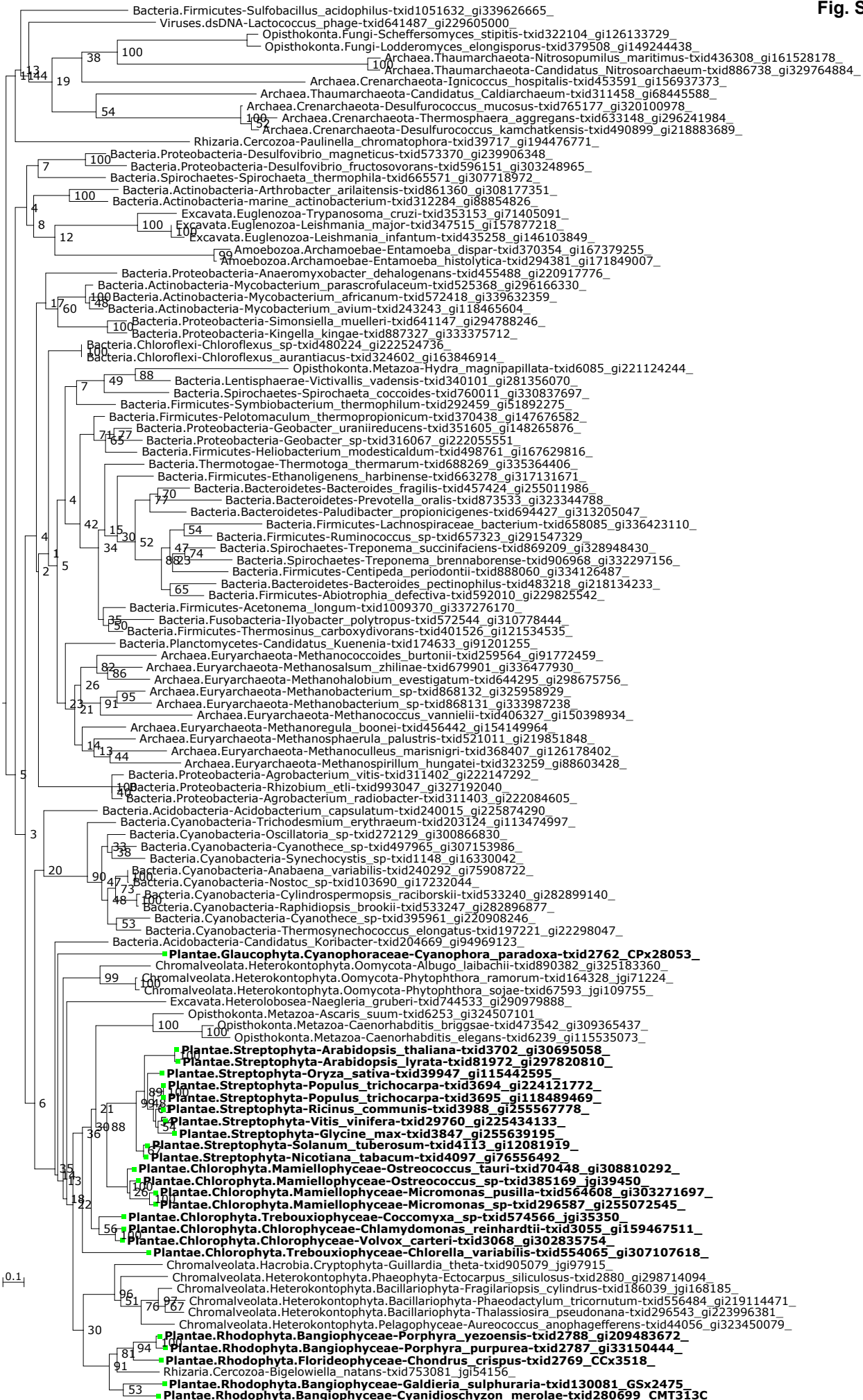




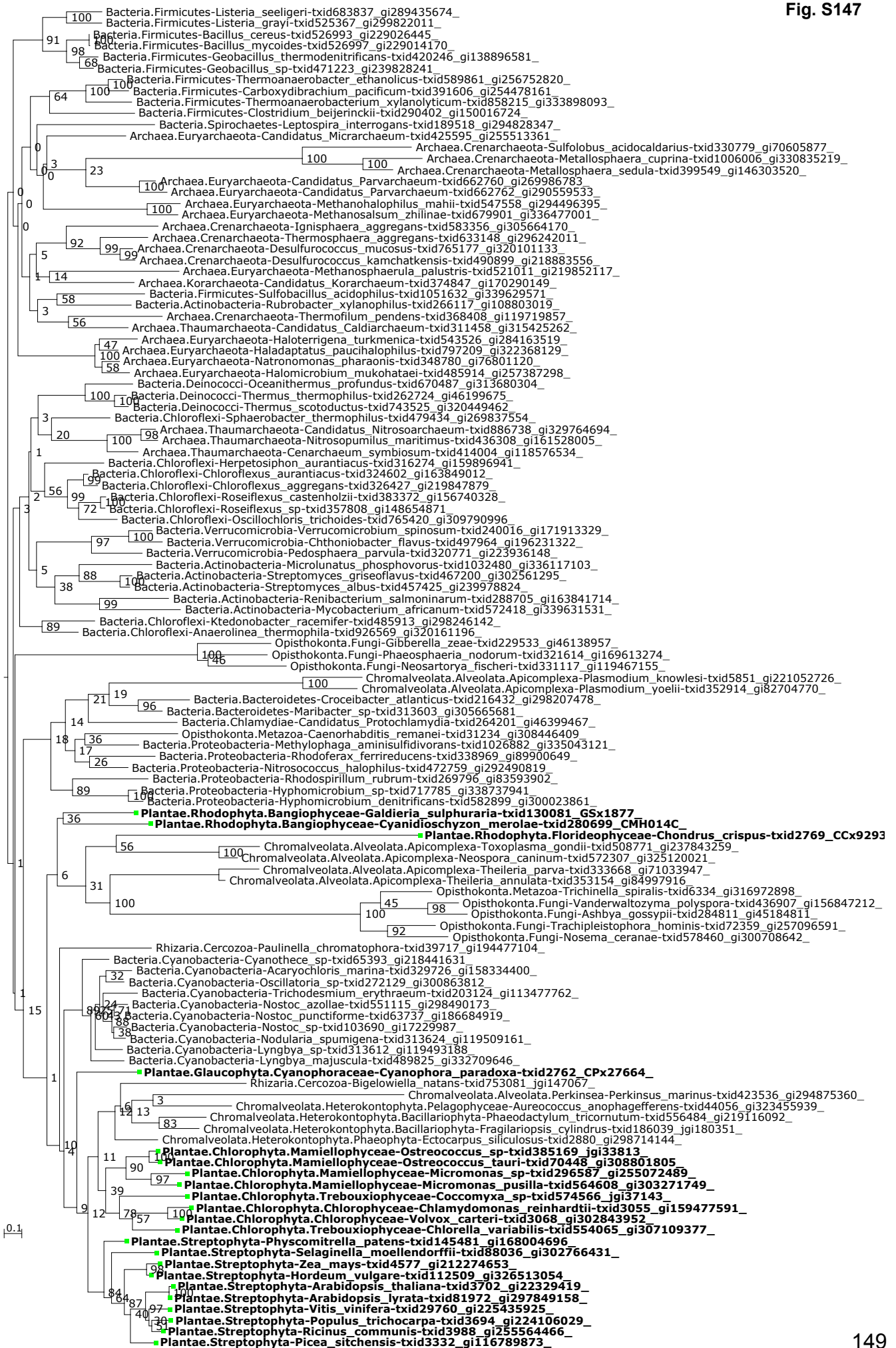




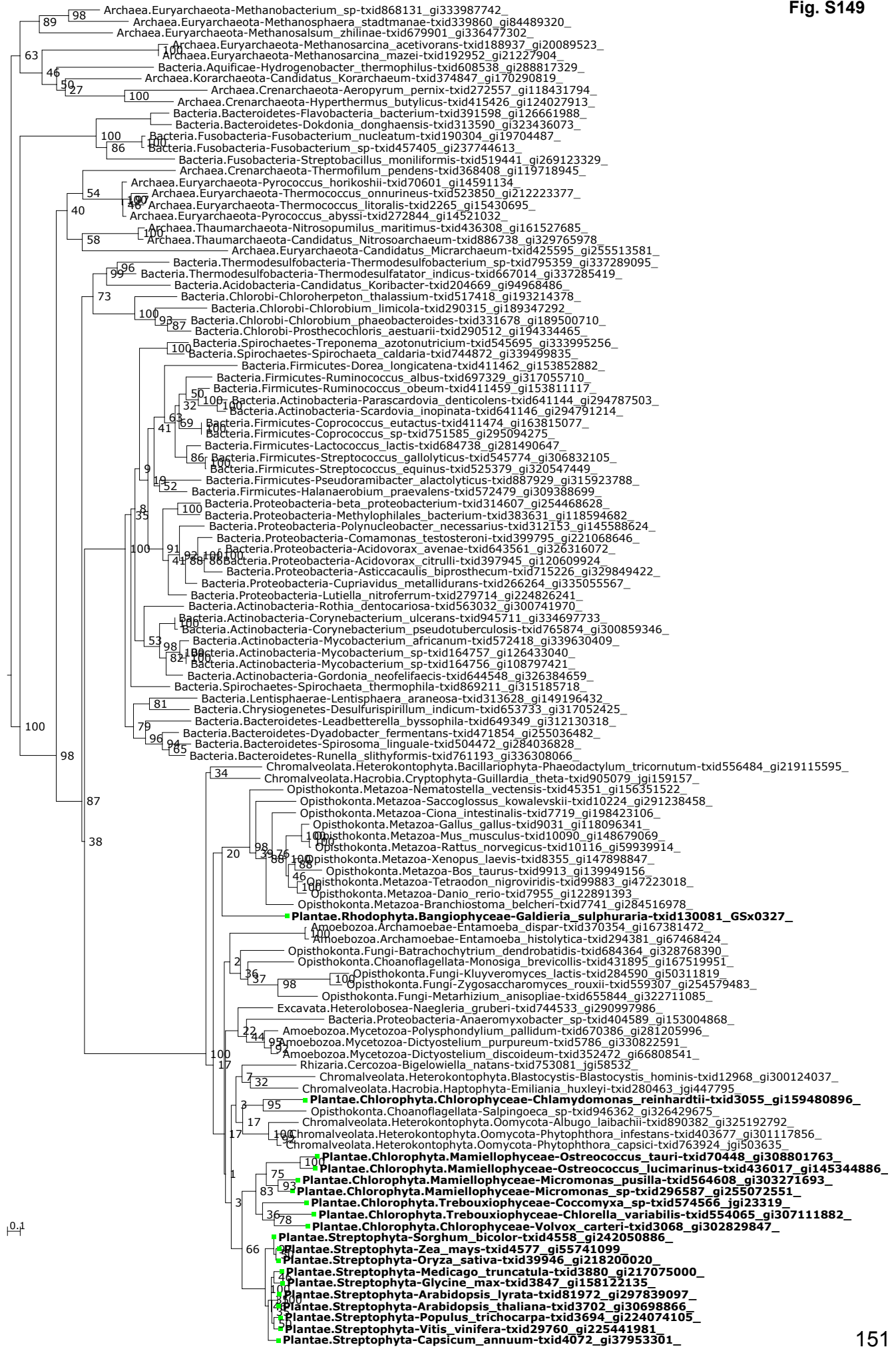


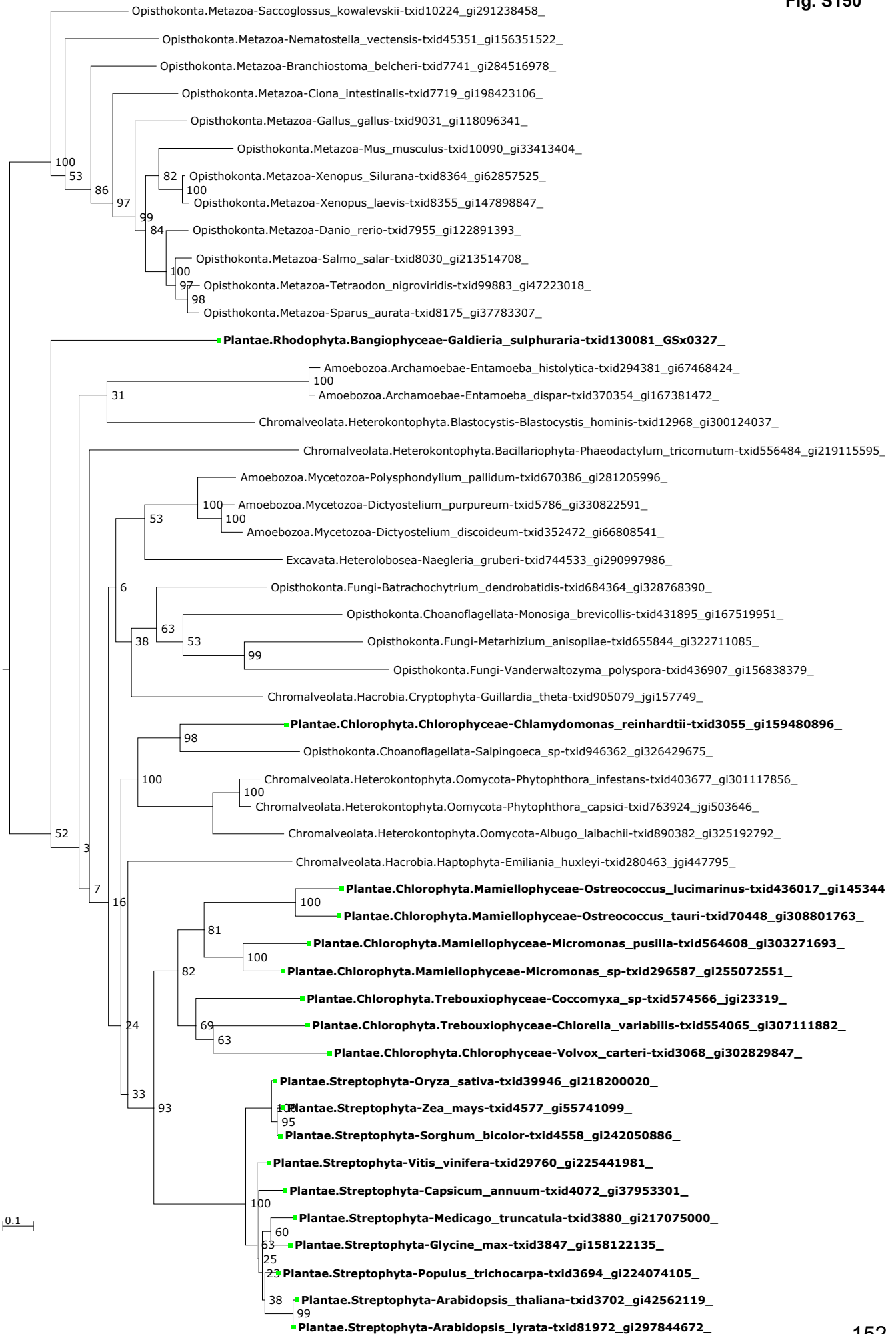




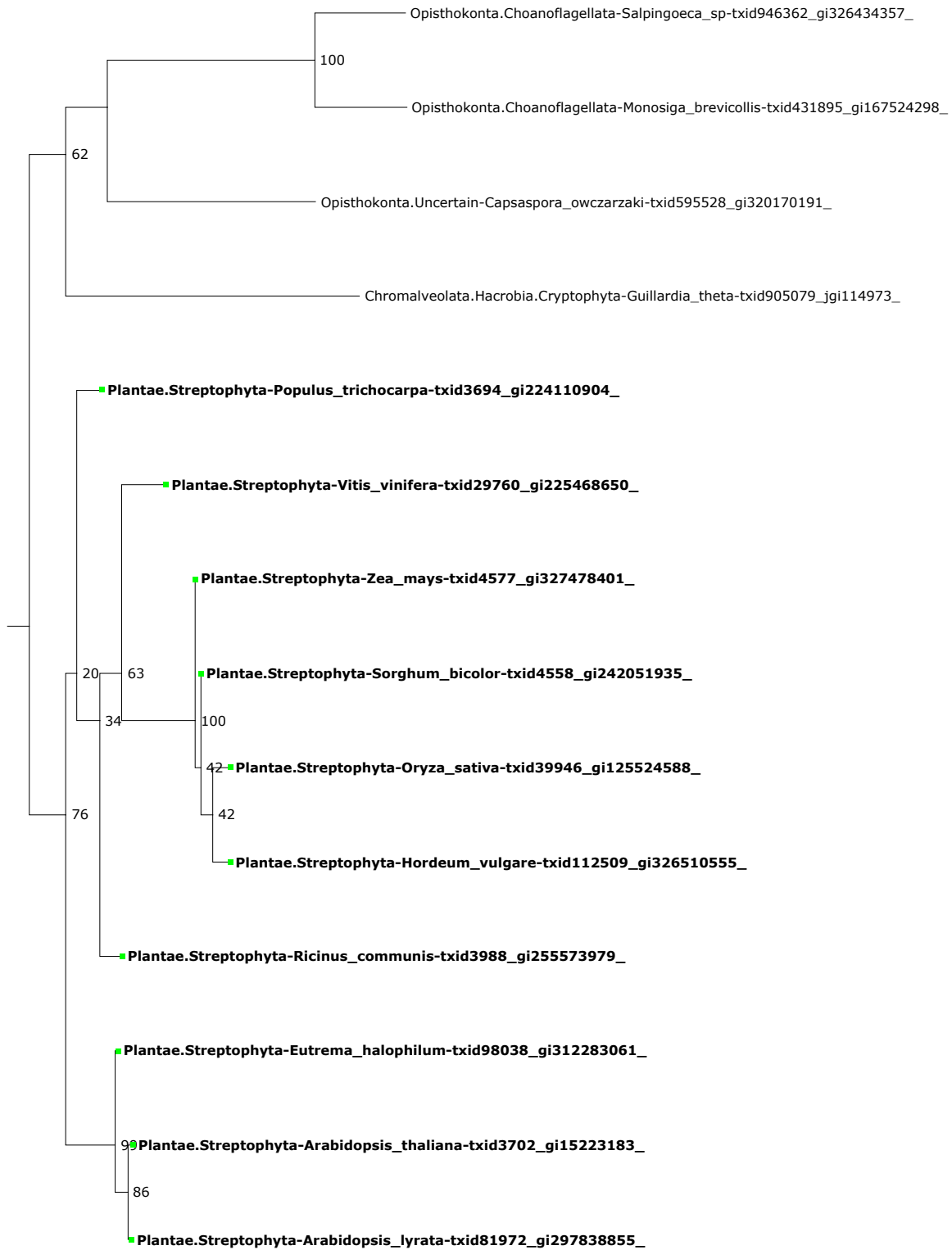




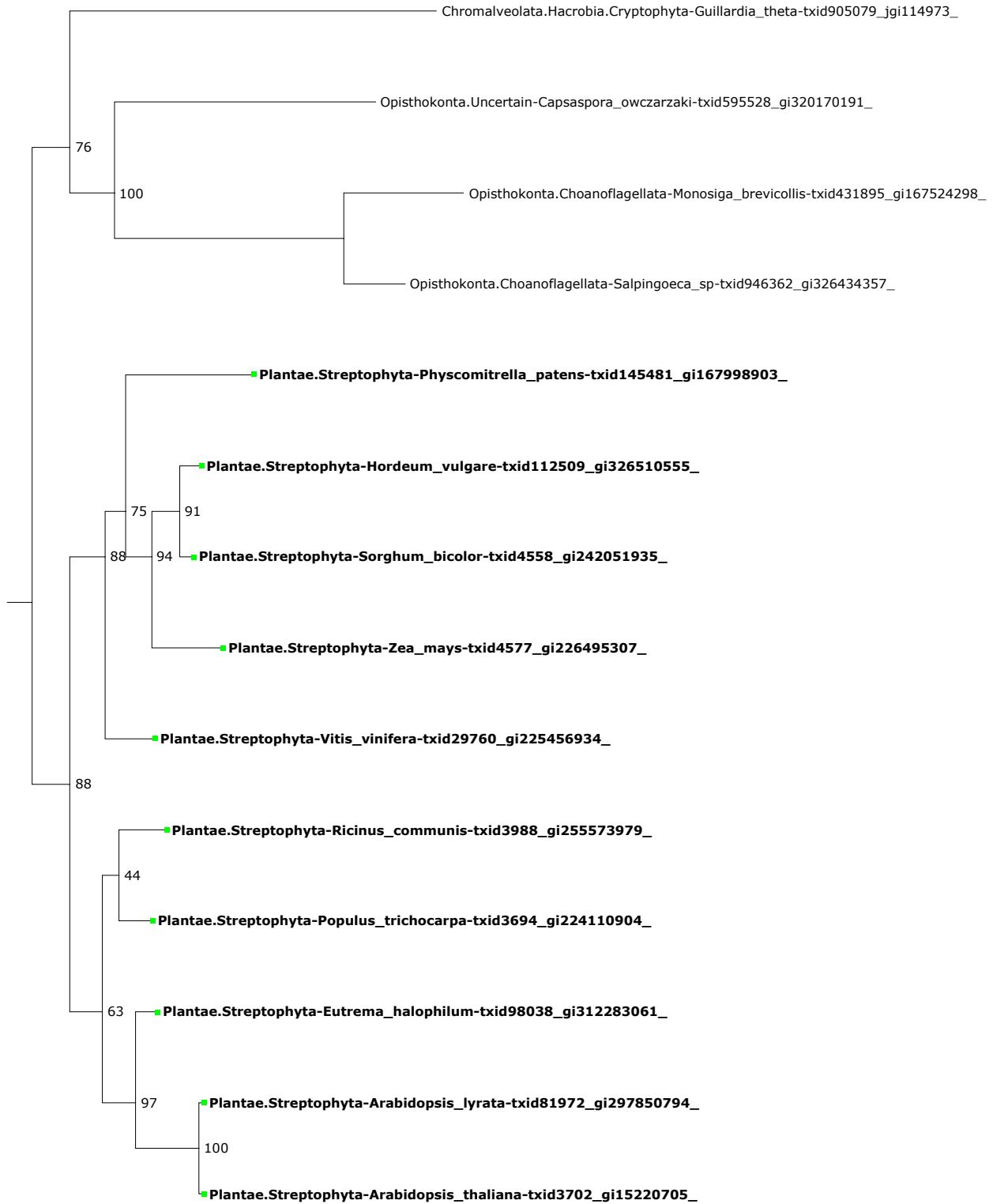








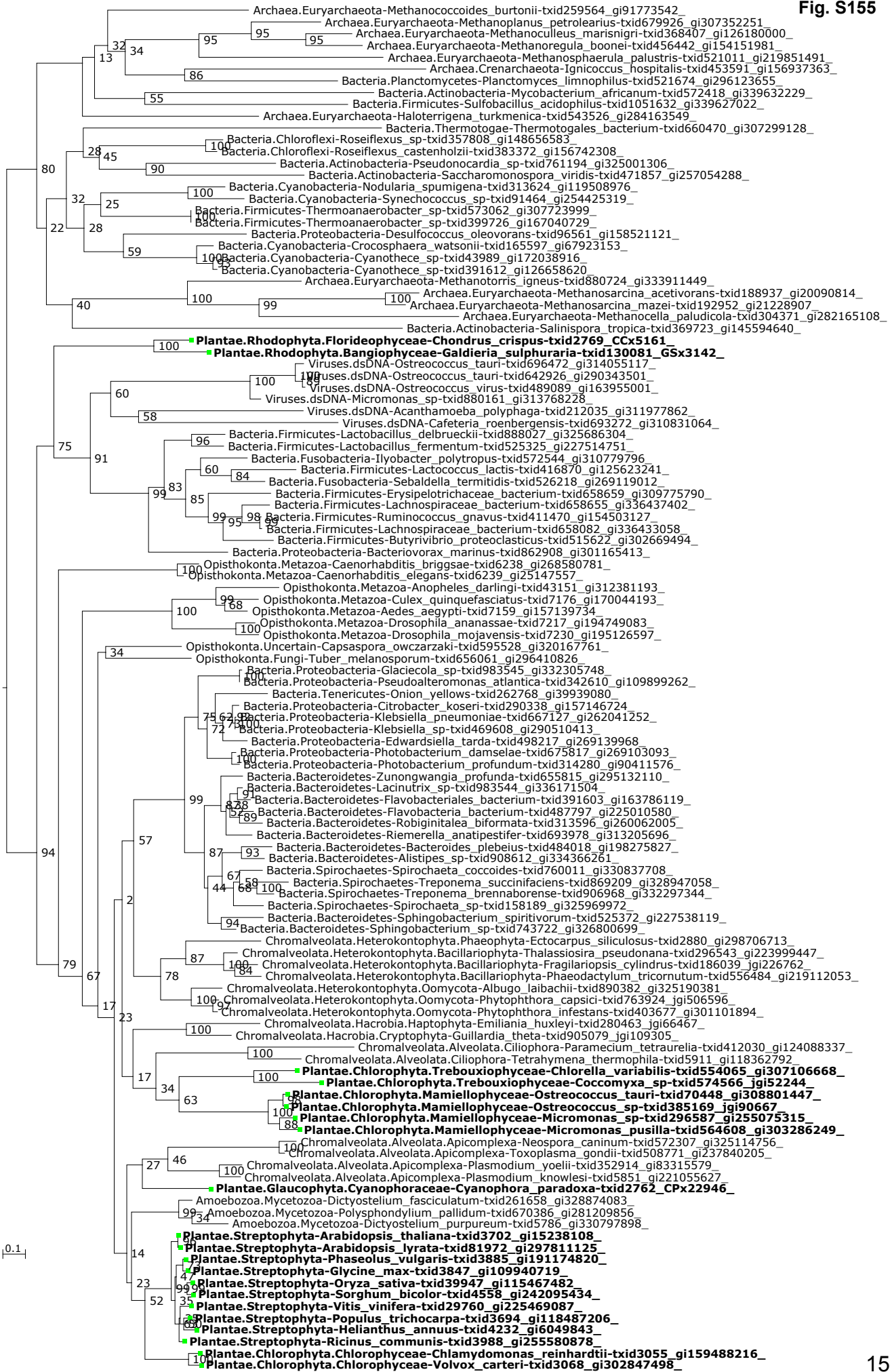
0.1

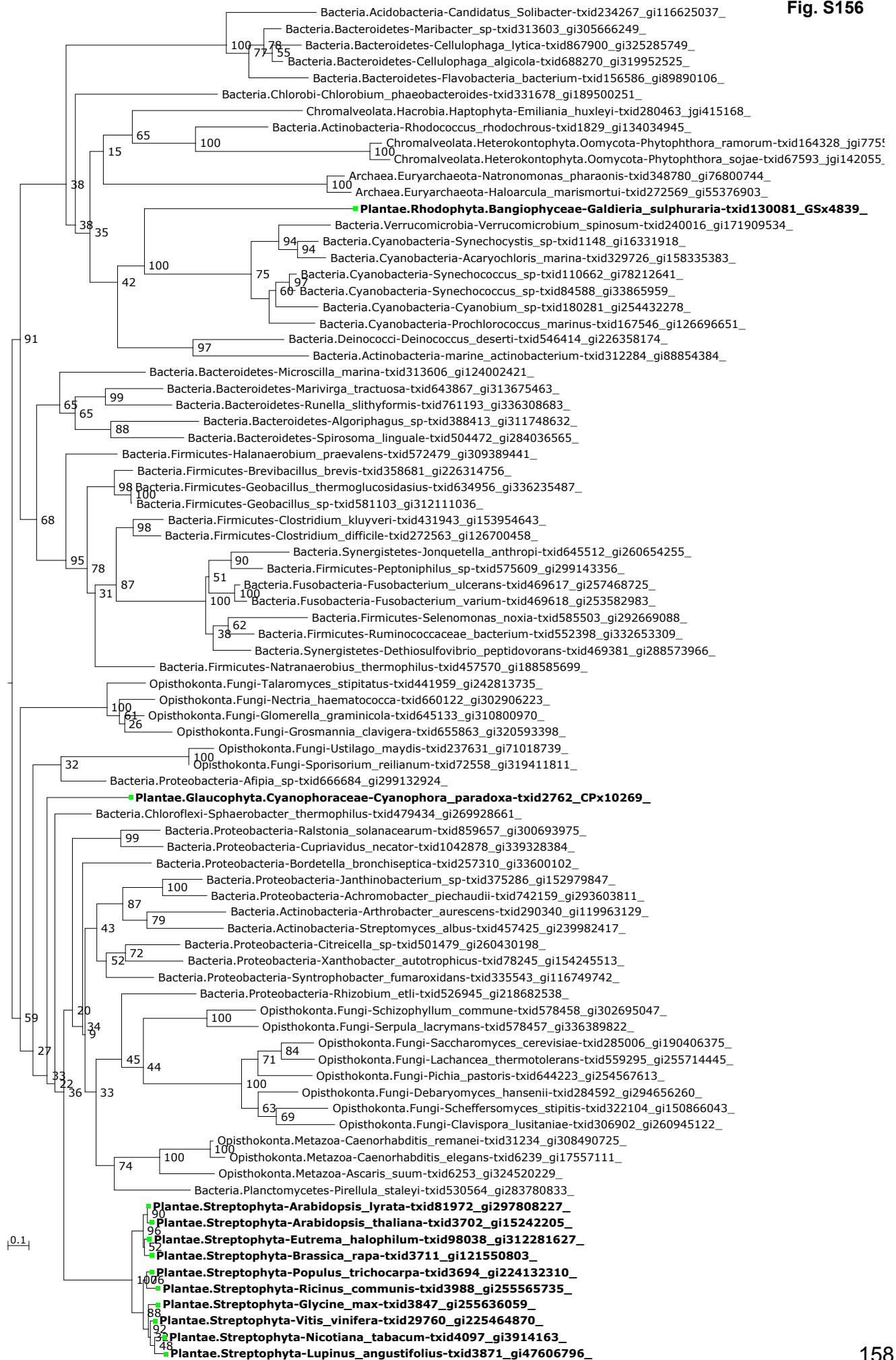


0.1









**Supplemental Table S1. *Arabidopsis thaliana* proteins involved in amino acid biosynthesis**

Pathway	Figure	Enzyme/Protein Name	EC N	Pt loc	Ref	PI Bra	CB ORI	CB HOM
<b>Membrane transport</b>		<i>Arabidopsis</i> gene number						
		<b>Nitrite transporter</b>						
	S1	At1g68570.1		○	(1)			
<b>Nitrite reduction</b>								
		<b>Nitrite reductase</b>						
	S2	At2g15620.1	EC 1.6.6.1	○	(2)	○	○	○
<b>Glutamine</b>								
		<b>Glutamine synthetase</b>						
	S3	At5g35630.1	EC 6.3.1.2	○	(3, 4)	○		
	S4	At5g37600.1			(5)	○		
	S5	At1g66200.2			(5)	○		
	S6	AT3G17820.1			(5)	○		
	S7	At5g16570.1			(5)	○		
	S8	At1g48470.1			(5)	○		
<b>Glutamate</b>								
		<b>Glutamate synthase Ferredoxin-dependent (Fd-GOGAT)</b>						
	S9	At5g04140.2	EC 1.4.7.1	○	(6, 7)	○	○	○
	S10	At2g41220.1		○	(6, 7)	○	○	○
		<b>Glutamate synthase NADH-dependent (GOGAT)</b>						
	S11	At5g53460.1	EC 1.4.1.14	○	(6, 7)			○
<b>Tryptophan</b>								
		<b>Anthranilate synthase alpha subunit</b>						
	Fig. 3A	At5g05730.1	EC 4.1.3.27	○	(8)	○		○
	S12	At2g29690.1		○	(8)	○		○
	S13	At3g55870.1				○		○
		<b>Anthranilate synthase beta subunit</b>						

<b>Fig. 3B</b>	At1g25220.1		○	(9)	○	○	○
S14	At1g24807.1		○	(10)	○	○	○
S15	At1g25155.1				○	○	○
S16	At5g57890.1				○	○	○
S17	At1g24909.1				○	○	○
S18	At1g25083.1				○	○	○
	<b>Anthranilate phosphoribosyl transferase</b>	EC 2.4.2.18					
S25	At5g17990.1		○	(8)	○		
	<b>Phosphoribosyl anthranilate isomerase</b>	EC 5.3.1.24					
S26	At5g05590.2		○	(8)	○		○
S27	At1g29410.1		○	(8)	○		
S28	At1g07780.3		○	(8)	○		
	<b>Indole-3-glycerol-phosphate synthase</b>	EC 4.1.1.48					
S19	At2g04400.1		○	(10, 11)		○	○
S20	At5g48220.1		○	(10-12)		○	○
	<b>Tryptophan synthase alpha subunit</b>	EC 4.2.1.20					
S21	At3g54640.1		○	(8)	○	○	○
S22	At4g02610.1		○	(8)	○	○	○
	<b>Tryptophan synthase beta subunit</b>						
S23	At5g54810.1		○	(8)	○	○	○
S24	At4g27070.1		○	(8)	○	○	○

---

**Phenylalanine and Tyrosine**

---

	<b>Chorismate mutase</b>	EC 5.4.99.5					
S29	At1g69370.1		○	(13, 14)			
S30	At3g29200.1			(13, 14)			
S31	At5g10870.1			(14)			
	<b>Glutamate/aspartate-prephenate aminotransferase</b>	EC 2.6.1.79					
S32	At2g22250.1		○	(15)			
	<b>Arogenate dehydratase/Prephenate Dehydratase</b>	EC 4.2.1.91					
S33	At3g44720.1		○	(16)	○		
S34	At1g11790.2		○	(16)	○		



S35	At3g07630.2		○	(16)	○		
S36	At2g27820.1		○	(16)	○		
S37	At5g22630.1		○	(16)	○		
S38	At1g08250.1		○	(16)	○		
	<b>Arogenate dehydrogenase</b>	EC 1.3.1.43					
S39	At5g34930.1		○	(16)	○		
S40	At1g15710.1		○	(16)	○		

---

### Histidine

---

	<b>ATP phosphoribosyltransferase</b>	EC 2.4.2.17					
S41	At1g58080.1		○	(17, 18)	○	?	○
S42	At1g09795.1		○	(18)	○	?	○
	<b>Phosphoribosyl ATP pyrophosphohydrolase</b>	EC 3.6.1.31					
	<b>Phosphoribosyl-AMP cyclohydrolase</b>	EC 3.5.4.19					
S43	At1g31860.1		○	(18, 19)			○
	<b>BBMII isomerase</b>	EC 5.3.1.16					
S44	At2g36230.1		○	(18, 20)	○		
	<b>Imidazoleglycerol-phosphate synthase</b>	EC 2.4.2.-					
S45	At4g26900.1		○	(10, 18, 20)			
	<b>Imidazoleglycerol-phosphate dehydratase</b>	EC 4.2.1.19					
S46	At3g22425.1		○	(18, 21)	○	○	○
<b>Fig. 2</b>	At4g14910.1		○	(18)	○	○	○
	<b>Histidinol-phosphate aminotransferase</b>	EC 2.6.1.9					
S48	At5g10330.1		○	(18)	○		○
<b>Fig. 4</b>	At1g71920.1		○	(18)	○		○
	<b>Histidinol phosphate phosphatase</b>	EC 3.1.3.15					
S50	At4g39120.1		○	(18)	○	?	○
	<b>Histidinol dehydrogenase</b>	EC 1.1.1.23					
S51	At5g63890.1		○	(18)			○

---

### Aspartate

---

	<b>Aspartate aminotransferase</b>	EC 2.6.1.1					
S52	At4g31990.1		○	(22-24)			
S53	At2g30970.1			(24)			
S54	At5g11520.1			(24)	○		

S55	At5g19550.1			(25)	○
S56	At1g62800.1				○
S57	At5g51690.1				
S58	At1g62960.1				

---

**Aspartate-derived**


---

	<b>Aspartate kinase</b>	EC 2.7.2.4			
S59	At5g13280.1		○	(26, 27)	○
S60	At5g14060.1			(26, 27)	○
S61	At3g02020.1			(26, 27)	○
	<b>Aspartate kinase/homoserine dehydrogenase</b>				
S62	At1g31230.1		○	(26-28)	
S63	At4g19710.1		○	(26-28)	
	<b>Aspartate semialdehyde dehydrogenase</b>	EC 1.2.1.11			
S64	At1g14810.1		○	(29)	? ○
	<b>Homoserine dehydrogenase</b>	EC 1.1.1.3			
S65	At5g21060.1			(30)	○
	<b>Homoserine kinase</b>	EC 2.7.1.39			
S66	At2g17265.1		○	(27, 31)	○

---

**Lysine**


---

	<b>Dihydrodipicolinate synthase</b>	EC 4.2.1.52			
S67	At3g60880.1		○	(32)	
S68	At2g45440.1		○	(31, 32)	
	<b>Dihydrodipicolinate reductase</b>	EC 1.3.1.26			
S69	At5g52100.1		○	(33)	○ ○
S70	At3g59890.2		○	(33, 34)	
S71	At2g44040.1		○	(33, 34)	○
	<b>L-diaminopimelate aminotransferase</b>	EC 2.6.1.83			
S72	At4g33680.1		○	(33, 35)	○ ○
	<b>Diaminopimelate epimerase</b>	EC 5.1.1.7			
S73	At3g53580.1		○	(33)	○ ○ ○
	<b>Diaminopimelate decarboxylase</b>	EC 4.1.1.20			
S74	At3g14390.1		○	(33, 34)	○
S75	At5g11880.1		○	(33, 34)	○

---

**Threonine**

	<b>Threonine synthase</b>	EC 4.2.99.2					
S76	At4g29840.1		○	(27, 36)	○	?	○
S77	At1g72810.1			-	○	?	○

**Methionine**

	from Cysteine + P-Homoserine (Aspartate)						
	<b>Cystathionine gamma-synthase</b>	EC 4.2.99.9					
S78	At3g01120.1		○	(37)	○		
	<b>Cystathionine beta-lyase</b>	EC 4.4.1.8					
S79	At3g57050.1		○	(37)			
	<b>Methionine synthase</b>	EC 2.1.1.14					
S80	At5g20980.1		○	(38)			
S81	At3g03780.1			(38)			
S82	At5g17920.1			(38)			

**Isoleucine**

	<b>Threonine deaminase (Threonine dehydratase)</b>	EC 4.2.1.16					
S83	At3g10050.1		○	(39, 40)		?	○
	<b>Acetolactate synthase large subunit</b>	EC 4.1.3.18					
<b>Fig. 5A</b>	At3g48560.1		○	(10, 41)	○		○
	<b>Acetolactate synthase small subunit</b>						
<b>Fig. 5B</b>	At2g31810		○	(10, 41)	○	○	○
S84	At5g16290.1					○	○
	<b>Ketolacid Reductoisomerase</b>	EC 1.1.1.86					
S85	At3g58610		○	(10, 41)	○		○
	<b>Dihydroxyacid Dehydratase</b>	EC 4.2.1.9					
S86	At3g23940.1		○	(10)			○
	<b>Branched-chain amino acid transaminases</b>	EC 2.6.1.42					
S87	At1g10070.1		○	(41)			○
S88	At3g49680.1		○	(10)			○
S89	At5g65780.1		○	(10)			○
	<b>methionine <math>\gamma</math>-lyase</b>	EC 4.4.1.11					
S90	At1g64660.1			(42)			

**Valine**

Uses the enzymes TDH1, ALS, KARI of  
isoleucine biosynthesis

Uses Branched-chain amino acid  
transaminases of isoleucine biosynthesis

**Leucine**

Uses the enzymes TDH1, ALS, KARI of  
isoleucine biosynthesis

Uses Branched-chain amino acid  
transaminases of isoleucine biosynthesis

**Isopropylmalate synthase**

EC 4.1.3.12

S91	At1g18500.1		○	(43)			○
-----	-------------	--	---	------	--	--	---

S92	At1g74040.1		○	(43)			○
-----	-------------	--	---	------	--	--	---

**Isopropylmalate synthase  
(Methylthioalkylmalate synthase)**

S93	At5g23020.1		○	(44)			○
-----	-------------	--	---	------	--	--	---

S94	At5g23010.1		○	(10, 45)			○
-----	-------------	--	---	----------	--	--	---

**Isopropylmalate isomerase**

EC 4.2.1.33

S95	At4g13430.1		○	(46, 47)	○		○
-----	-------------	--	---	----------	---	--	---

S96	At2g43090.1		○	(46, 47)	○		○
-----	-------------	--	---	----------	---	--	---

S97	At2g43100.1		○	(46, 47)	○		○
-----	-------------	--	---	----------	---	--	---

S98	At3g58990.1		○	(46, 47)	○		○
-----	-------------	--	---	----------	---	--	---

**Isopropylmalate dehydrogenase**

EC 1.1.1.85

S99	At1g80555.1		○		○	○	○
-----	-------------	--	---	--	---	---	---

S100	At5g14200.1		○	(12, 48)	○	○	○
------	-------------	--	---	----------	---	---	---

S101	At1g80560.1		○	(12, 48)	○	○	○
------	-------------	--	---	----------	---	---	---

S102	At1g31180.1		○	(12, 48)	○	○	○
------	-------------	--	---	----------	---	---	---

**Arginine***Ornithine pathway***N-Acetylglutamate synthase**

EC 2.3.1.1

S103	At4g37670.2		○	(49, 50)	○		
------	-------------	--	---	----------	---	--	--

S104	At2g22910.1			(49, 50)	○		
------	-------------	--	--	----------	---	--	--

**N-Acetylglutamate kinase**

EC 2.7.2.8

S105	At3g57560.1		○	(50, 51)	○	○	○
------	-------------	--	---	----------	---	---	---

S106	<b>N-Acetylglutamate-5-P reductase</b> At2g19940.2	EC 1.2.1.38	○	(10, 52, 53)	○		○
S107	<b>2-Acetylornithine aminotransferase</b> At1g80600	EC 2.6.1.11	○	(10, 12, 49)			○
S108	<b>N2-Acetylornithine:glutamate acetyltransferase</b> At2g37500.1	EC 2.3.1.35	○	(10, 12, 49)	○	○	○
S109	<b>N2-Acetylornithine deacetylase</b> At4g17830.1 <i>Arginine pathway</i>	EC 3.5.1.16		(50, 54)			
S110	<b>Carbamoyl-P synthetase Large subunit</b> At1g29900.1	EC 6.3.5.5	○	(10, 12, 49)	○	○	○
S111	<b>Carbamoyl-P synthetase Small subunit</b> At3g27740.1		○	(10, 12, 49)			○
S112	<b>Ornithine carbamoyltransferase</b> At1g75330.1	EC 2.1.3.3	○	(10, 12, 49)	○		○
S113	<b>Argininosuccinate synthase</b> At4g24830.1	EC 6.3.4.5	○	(53, 55)			
S114	<b>Argininosuccinate lyase</b> At5g10920.1	EC 4.3.2.1	○	(10, 12)			○
S115	<b>P II nitrogen sensing protein GLB I</b> At4g01900.1		○	(56)	○	○	○
<b>Proline</b>							
	from Glutamate						
S116	<b>Pyrroline-5-carboxylate synthase</b> At2g39800.1	EC 1.2.1.41		(57)			○
S117	At3g55610.1		○	(57)			○
S118	<b>Pyrroline-5-carboxylate reductase</b> At5g14800.1 from Ornithine	EC 1.5.1.2		(58, 59)			○
S119	<b>Ornithine-<math>\delta</math>-aminotransferase</b> At5g46180.1	EC 2.6.1.13		(60, 61)			○
<b>Serine</b>							
	<b>3-phosphoglycerate dehydrogenase</b>	EC 1.1.1.95					

S120	At1g17745.1		○	(62, 63)	○	○
S121	At1g72190.1					
S122	At1g68010.1		○	(10)	○	○
S123	At3g19480.1			(10)	○	○
	<b>Phosphoserine aminotransferase</b>	EC 2.6.1.52				
S124	At4g35630.1		○	(63)		
S125	At2g17630.1		○	(63)		
	<b>3-phosphoserine phosphatase</b>	EC 3.1.3.3				
S126	At1g18640.1		○	(63)		

---

**Glycine**

---

	<i>from Serine</i>					
	<b>Serine(glycine) hydroxymethyltransferase</b>	EC 2.1.2.1				
S127	At4g32520.1		○	(64)	○	○
S128	At4g37930.1			(65)	○	○
S129	At5g26780.3				○	○
S130	At4g13890.1				○	○
S131	At4g13930.1				○	○
S132	At1g22020.1				○	○
S133	At1g36370.1		○		○	○
	<i>from Threonine</i>					
	<b>Threonine aldolase</b>	EC 4.1.2.5				
S134	At3g04520.1			(66)		
S135	At1g08630.1			(66)		
	<i>from Glyoxylate</i>					
	<b>Serine(alanine):glyoxylate aminotransferase</b>	EC 2.6.1.45				
<b>Fig. 6</b>	At2g13360.2		○	(53, 55)	○	○
	<b>Glutamate:glyoxylate (alanine) aminotransferase</b>	EC 2.6.1.4				
S137	At1g23310.1		○	(55)		
S138	At1g70580.1		○	(55)		

---

**Cysteine**

---

	<b>Serine O-acetyltransferase</b>	EC 2.3.1.30				
S139	At1g55920.1		○	(67, 68)		○

S140	At2g17640.1		(67, 69)	○		○
S141	At3g13110.1		(67, 68)			○
S142	At4g35640.1		(67, 69)	○		○
S143	At5g56760.1		(67, 68)	○		○
	<b>Cysteine synthase/O-acetylserine thiol lyase</b>	EC 4.2.99.8				
S144	At2g43750.1		(70)	○		○
S145	At3g59760.1		(70)	○	?	○
S146	At4g14880.1		(71)	○	?	○

---

### Alanine

---

	From cysteine					
	<b>Cysteine desulfurase</b>	EC 2.8.1.7				
S147	At1g08490.1		(72, 73)	○	○	○
S148	At5g65720.1		(74)			○
	<b>Alanine aminotransferase</b>	EC 2.6.1.2				
S149	At1g72330.1		(75)			
S150	At1g17290.1		(75, 76)			
	<b>Alanine (tryptophan) aminotransferase</b>					
S151	At1g70560.1		(77)			
S152	At1g23320.1		(77)			

---

### Asparagine

---

	From Aspartate					
	<b>Asparagine synthetase (glutamine dependent)</b>	EC 6.3.5.4				
S153	At3g47340.2		(78)			○
	<b>Asparagine synthetase</b>					
S154	At5g65010.1		(78)			○
S155	At5g10240.1		(78)			○
	cyanide detoxification					
	<b>β-cyano-L-alanine hydratase/nitrilase</b>	EC 3.5.5.4				
S156	At5g22300.1		(79)			○

---

**EC N:** Enzyme commission numbers

**Pt loc:** Plastid localized enzymes (see Methods)

**Ref:** Reference to experimental data regarding sub-cellular localization of the corresponding protein

**PI Bra:** At least two members of the three different lineages of the Plantae supergroup (sensu Cavalier-Smith, 2981) branch in the same clade.

**CB ORI:** the query protein is of putative cyanobacterial origin

**CB HOM:** cyanobacterial homologous sequences are present in the ML tree when a non-cyanobacterial origin is presumed in the Plantae homolog



## References

1. Sugiura, M., Georgescu, M.N. & Takahashi, M. A nitrite transporter associated with nitrite uptake by higher plant chloroplasts. *Plant Cell Physiol* **48**, 1022-1035 (2007).
2. Takahashi, M., Sasaki, Y., Ida, S. & Morikawa, H. Nitrite reductase gene enrichment improves assimilation of NO(2) in Arabidopsis. *Plant Physiol* **126**, 731-741 (2001).
3. O'Neal, D. & Joy, K.W. Localisation of glutamine synthetase in chloroplasts. *Nat New Biol* **246**, 61-62 (1973).
4. Tingey, S.V., Tsai, F.Y., Edwards, J.W., Walker, E.L. & Coruzzi, G.M. Chloroplast and cytosolic glutamine synthetase are encoded by homologous nuclear genes which are differentially expressed in vivo. *J Biol Chem* **263**, 9651-9657 (1988).
5. Bernard, S.M. & Habash, D.Z. The importance of cytosolic glutamine synthetase in nitrogen assimilation and recycling. *New Phytol* **182**, 608-620 (2009).
6. Botella, J.R., Verbelen, J.P. & Valpuesta, V. Immunocytolocalization of Ferredoxin-GOGAT in the Cells of Green Leaves and Cotyledons of *Lycopersicon esculentum*. *Plant Physiol* **87**, 255-257 (1988).
7. Suzuki, A. & Knaff, D.B. Glutamate synthase: structural, mechanistic and regulatory properties, and role in the amino acid metabolism. *Photosynth Res* **83**, 191-217 (2005).
8. Zhao, J. & Last, R.L. Immunological characterization and chloroplast localization of the tryptophan biosynthetic enzymes of the flowering plant *Arabidopsis thaliana*. *J Biol Chem* **270**, 6081-6087 (1995).
9. Niyogi, K.K., Last, R.L., Fink, G.R. & Keith, B. Suppressors of trp1 fluorescence identify a new arabidopsis gene, TRP4, encoding the anthranilate synthase beta subunit. *Plant Cell* **5**, 1011-1027 (1993).
10. Zybailov, B. *et al.* Sorting signals, N-terminal modifications and abundance of the chloroplast proteome. *PLoS One* **3**, e1994 (2008).
11. Li, J., Chen, S., Zhu, L. & Last, R.L. Isolation of cDNAs encoding the tryptophan pathway enzyme indole-3-glycerol phosphate synthase from *Arabidopsis thaliana*. *Plant Physiol* **108**, 877-878 (1995).
12. Peltier, J.B. *et al.* The oligomeric stromal proteome of *Arabidopsis thaliana* chloroplasts. *Mol Cell Proteomics* **5**, 114-133 (2006).
13. Eberhard, J. *et al.* Cytosolic and plastidic chorismate mutase isozymes from *Arabidopsis thaliana*: molecular characterization and enzymatic properties. *Plant J* **10**, 815-821 (1996).
14. Mobley, E.M., Kunkel, B.N. & Keith, B. Identification, characterization and comparative analysis of a novel chorismate mutase gene in *Arabidopsis thaliana*. *Gene* **240**, 115-123 (1999).
15. Graindorge, M. *et al.* Identification of a plant gene encoding glutamate/aspartate-prephenate aminotransferase: the last homeless enzyme of aromatic amino acids biosynthesis. *FEBS Lett* **584**, 4357-4360 (2010).
16. Rippert, P., Puyaubert, J., Grisolle, D., Derrier, L. & Matringe, M. Tyrosine and phenylalanine are synthesized within the plastids in *Arabidopsis*. *Plant Physiol* **149**, 1251-1260 (2009).
17. Ohta, D. *et al.* Molecular cloning and characterization of ATP-phosphoribosyl transferase from *Arabidopsis*, a key enzyme in the histidine biosynthetic pathway. *Plant Physiol* **122**, 907-914 (2000).
18. Stepansky, A. & Leustek, T. Histidine biosynthesis in plants. *Amino Acids* **30**, 127-142 (2006).
19. Fujimori, K. & Ohta, D. Isolation and characterization of a histidine biosynthetic gene in *Arabidopsis* encoding a polypeptide with two separate domains for phosphoribosyl-ATP pyrophosphohydrolase and phosphoribosyl-AMP cyclohydrolase. *Plant Physiol* **118**, 275-283 (1998).
20. Muralla, R., Sweeney, C., Stepansky, A., Leustek, T. & Meinke, D. Genetic dissection of histidine biosynthesis in *Arabidopsis*. *Plant Physiol* **144**, 890-903 (2007).
21. Tada, S. *et al.* Insect cell expression of recombinant imidazoleglycerolphosphate dehydratase of *Arabidopsis* and wheat and inhibition by triazole herbicides. *Plant Physiol* **109**, 153-159 (1995).
22. Kwok, E.Y. & Hanson, M.R. GFP-labelled Rubisco and aspartate aminotransferase are present in plastid stromules and traffic between plastids. *J Exp Bot* **55**, 595-604 (2004).
23. Wilkie, S.E., Lambert, R. & Warren, M.J. Chloroplastic aspartate aminotransferase from *Arabidopsis thaliana*: an examination of the relationship between the structure of the gene and the spatial structure of the protein. *Biochem J* **319** ( Pt 3, 969-976 (1996).

24. Wilkie, S.E. & Warren, M.J. Recombinant expression, purification, and characterization of three isoenzymes of aspartate aminotransferase from *Arabidopsis thaliana*. *Protein Expr Purif* **12**, 381-389 (1998).
25. Schultz, C.J., Hsu, M., Miesak, B. & Coruzzi, G.M. *Arabidopsis* mutants define an in vivo role for isoenzymes of aspartate aminotransferase in plant nitrogen assimilation. *Genetics* **149**, 491-499 (1998).
26. Curien, G., Laurencin, M., Robert-Genthon, M. & Dumas, R. Allosteric monofunctional aspartate kinases from *Arabidopsis*. *FEBS J* **274**, 164-176 (2007).
27. Wallsgrave, R.M., Lea, P.J. & Mifflin, B.J. Intracellular localization of aspartate kinase and the enzymes of threonine and methionine biosynthesis in green leaves. *Plant Physiol* **71**, 780-784 (1983).
28. Ghislain, M., Frankard, V., Vandenbossche, D., Matthews, B.F. & Jacobs, M. Molecular analysis of the aspartate kinase-homoserine dehydrogenase gene from *Arabidopsis thaliana*. *Plant Mol Biol* **24**, 835-851 (1994).
29. Paris, S., Wessel, P.M. & Dumas, R. Overproduction, purification, and characterization of recombinant aspartate semialdehyde dehydrogenase from *Arabidopsis thaliana*. *Protein Expr Purif* **24**, 99-104 (2002).
30. Schroeder, A.C. *et al.* Threonine-insensitive homoserine dehydrogenase from soybean: genomic organization, kinetic mechanism, and in vivo activity. *J Biol Chem* **285**, 827-834 (2010).
31. Sarrobert, C. *et al.* Identification of an *Arabidopsis thaliana* mutant accumulating threonine resulting from mutation in a new dihydrodipicolinate synthase gene. *Plant J* **24**, 357-367 (2000).
32. Vauterin, M. & Jacobs, M. Isolation of a poplar and an *Arabidopsis thaliana* dihydrodipicolinate synthase cDNA clone. *Plant Mol Biol* **25**, 545-550 (1994).
33. Hudson, A.O. *et al.* Biosynthesis of lysine in plants: evidence for a variant of the known bacterial pathways. *Biochim Biophys Acta* **1721**, 27-36 (2005).
34. Azevedo, R.A., Arruda, P., Turner, W.L. & Lea, P.J. The biosynthesis and metabolism of the aspartate derived amino acids in higher plants. *Phytochemistry* **46**, 395-419 (1997).
35. Hudson, A.O., Singh, B.K., Leustek, T. & Gilvarg, C. An LL-diaminopimelate aminotransferase defines a novel variant of the lysine biosynthesis pathway in plants. *Plant Physiol* **140**, 292-301 (2006).
36. Laber, B. *et al.* Characterization of recombinant *Arabidopsis thaliana* threonine synthase. *Eur J Biochem* **263**, 212-221 (1999).
37. Ravanel, S., Gakiere, B., Job, D. & Douce, R. The specific features of methionine biosynthesis and metabolism in plants. *Proc Natl Acad Sci U S A* **95**, 7805-7812 (1998).
38. Ravanel, S. *et al.* Methionine metabolism in plants: chloroplasts are autonomous for de novo methionine synthesis and can import S-adenosylmethionine from the cytosol. *J Biol Chem* **279**, 22548-22557 (2004).
39. Mifflin, B.J. The location of nitrite reductase and other enzymes related to amino Acid biosynthesis in the plastids of root and leaves. *Plant Physiol* **54**, 550-555 (1974).
40. Schulze-Siebert, D., Heineke, D., Scharf, H. & Schultz, G. Pyruvate-Derived Amino Acids in Spinach Chloroplasts : Synthesis and Regulation during Photosynthetic Carbon Metabolism. *Plant Physiol* **76**, 465-471 (1984).
41. Binder, S. Branched-Chain Amino Acid Metabolism in *Arabidopsis thaliana*. *The Arabidopsis Book* **null**, 1-14 (2008).
42. Joshi, V. & Jander, G. *Arabidopsis* methionine gamma-lyase is regulated according to isoleucine biosynthesis needs but plays a subordinate role to threonine deaminase. *Plant physiology* **151**, 367-78 (2009).
43. de Kraker, J.W., Luck, K., Textor, S., Tokuhisa, J.G. & Gershenzon, J. Two *Arabidopsis* genes (IPMS1 and IPMS2) encode isopropylmalate synthase, the branchpoint step in the biosynthesis of leucine. *Plant Physiol* **143**, 970-986 (2007).
44. Textor, S., de Kraker, J.W., Hause, B., Gershenzon, J. & Tokuhisa, J.G. MAM3 catalyzes the formation of all aliphatic glucosinolate chain lengths in *Arabidopsis*. *Plant Physiol* **144**, 60-71 (2007).
45. Field, B. *et al.* Glucosinolate and amino acid biosynthesis in *Arabidopsis*. *Plant Physiol* **135**, 828-839 (2004).
46. Knill, T., Reichelt, M., Paetz, C., Gershenzon, J. & Binder, S. *Arabidopsis thaliana* encodes a bacterial-type heterodimeric isopropylmalate isomerase involved in both Leu biosynthesis and the Met chain elongation pathway of glucosinolate formation. *Plant Mol Biol* **71**, 227-239 (2009).
47. He, Y., Chen, B., Pang, Q., Strul, J.M. & Chen, S. Functional specification of *Arabidopsis* isopropylmalate isomerases in glucosinolate and leucine biosynthesis. *Plant Cell Physiol* **51**, 1480-1487 (2010).

48. Nozawa, A., Takano, J., Miwa, K., Nakagawa, Y. & Fujiwara, T. Cloning of cDNAs encoding isopropylmalate dehydrogenase from *Arabidopsis thaliana* and accumulation patterns of their transcripts. *Biosci Biotechnol Biochem* **69**, 806-810 (2005).
49. Jain, J.C., Shargool, P.D. & Chung, S. Compartmentation studies on enzymes of ornithine biosynthesis in plant cells. *Plant Science* **51**, 17-20 (1987).
50. Slocum, R.D. Genes, enzymes and regulation of arginine biosynthesis in plants. *Plant Physiol Biochem* **43**, 729-745 (2005).
51. Sugiyama, K., Hayakawa, T., Kudo, T., Ito, T. & Yamaya, T. Interaction of N-acetylglutamate kinase with a PII-like protein in rice. *Plant Cell Physiol* **45**, 1768-1778 (2004).
52. Mitra, S.K., Gantt, J.A., Ruby, J.F., Clouse, S.D. & Goshe, M.B. Membrane proteomic analysis of *Arabidopsis thaliana* using alternative solubilization techniques. *J Proteome Res* **6**, 1933-1950 (2007).
53. Rutschow, H., Ytterberg, A.J., Friso, G., Nilsson, R. & van Wijk, K.J. Quantitative proteomics of a chloroplast SRP54 sorting mutant and its genetic interactions with CLPC1 in *Arabidopsis*. *Plant Physiol* **148**, 156-175 (2008).
54. Yokota, A. *et al.* Citrulline and DRIP-1 protein (ArgE homologue) in drought tolerance of wild watermelon. *Ann Bot* **89 Spec No.**, 825-832 (2002).
55. Kleffmann, T. *et al.* The *Arabidopsis thaliana* chloroplast proteome reveals pathway abundance and novel protein functions. *Curr Biol* **14**, 354-362 (2004).
56. Hsieh, M.H., Lam, H.M., van de Loo, F.J. & Coruzzi, G. A PII-like protein in *Arabidopsis*: putative role in nitrogen sensing. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 13965-70 (1998).
57. Szekely, G. *et al.* Duplicated P5CS genes of *Arabidopsis* play distinct roles in stress regulation and developmental control of proline biosynthesis. *Plant J* **53**, 11-28 (2008).
58. Rayapati, P.J., Stewart, C.R. & Hack, E. Pyrroline-5-Carboxylate Reductase Is in Pea (*Pisum sativum* L.) Leaf Chloroplasts. *Plant Physiol* **91**, 581-586 (1989).
59. Szoke, A., Miao, G.H., Hong, Z. & Verma, D.P. Subcellular location of delta-pyrroline-5-carboxylate reductase in root/nodule and leaf of soybean. *Plant Physiol* **99**, 1642-1649 (1992).
60. Funck, D., Stadelhofer, B. & Koch, W. Ornithine-delta-aminotransferase is essential for arginine catabolism but not for proline biosynthesis. *BMC Plant Biol* **8**, 40 (2008).
61. Roosens, N.H., Thu, T.T., Iskandar, H.M. & Jacobs, M. Isolation of the ornithine-delta-aminotransferase cDNA and effect of salt stress on its expression in *Arabidopsis thaliana*. *Plant Physiol* **117**, 263-271 (1998).
62. Ho, C.L., Noji, M. & Saito, K. Plastidic pathway of serine biosynthesis. Molecular cloning and expression of 3-phosphoserine phosphatase from *Arabidopsis thaliana*. *J Biol Chem* **274**, 11007-11012 (1999).
63. Ho, C.L. & Saito, K. Molecular biology of the plastidic phosphorylated serine biosynthetic pathway in *Arabidopsis thaliana*. *Amino Acids* **20**, 243-259 (2001).
64. Zhang, Y., Sun, K., Sandoval, F.J., Santiago, K. & Roje, S. One-carbon metabolism in plants: characterization of a plastid serine hydroxymethyltransferase. *Biochem J* **430**, 97-105 (2010).
65. Jamai, A., Salome, P.A., Schilling, S.H., Weber, A.P. & McClung, C.R. *Arabidopsis* photorespiratory serine hydroxymethyltransferase activity requires the mitochondrial accumulation of ferredoxin-dependent glutamate synthase. *Plant Cell* **21**, 595-606 (2009).
66. Joshi, V., Laubengayer, K.M., Schauer, N., Fernie, A.R. & Jander, G. Two *Arabidopsis* threonine aldolases are nonredundant and compete with threonine deaminase for a common substrate pool. *The Plant cell* **18**, 3564-75 (2006).
67. Howarth, J.R. *et al.* The serine acetyltransferase gene family in *Arabidopsis thaliana* and the regulation of its expression by cadmium. *Plant Mol Biol* **51**, 589-598 (2003).
68. Noji, M., Inoue, K., Kimura, N., Gouda, A. & Saito, K. Isoform-dependent differences in feedback regulation and subcellular localization of serine acetyltransferase involved in cysteine biosynthesis from *Arabidopsis thaliana*. *J Biol Chem* **273**, 32739-32745 (1998).
69. Kawashima, C.G., Berkowitz, O., Hell, R., Noji, M. & Saito, K. Characterization and expression analysis of a serine acetyltransferase gene family involved in a key step of the sulfur assimilation pathway in *Arabidopsis*. *Plant Physiol* **137**, 220-230 (2005).
70. Hesse, H., Lipke, J., Altmann, T. & Hofgen, R. Molecular cloning and expression analyses of mitochondrial and plastidic isoforms of cysteine synthase (O-acetylserine(thiol)lyase) from *Arabidopsis thaliana*. *Amino Acids* **16**, 113-131 (1999).
71. Barroso, C., Vega, J.M. & Gotor, C. A new member of the cytosolic O-acetylserine(thiol)lyase gene family in *Arabidopsis thaliana*. *FEBS Lett* **363**, 1-5 (1995).

72. Leon, S., Touraine, B., Briat, J.F. & Lobreaux, S. The AtNFS2 gene from *Arabidopsis thaliana* encodes a NifS-like plastidial cysteine desulphurase. *Biochem J* **366**, 557-564 (2002).
73. Van Hoewyk, D. *et al.* Chloroplast iron-sulfur cluster protein maturation requires the essential cysteine desulfurase CpNifS. *Proc Natl Acad Sci U S A* **104**, 5686-5691 (2007).
74. Frazzon, A.P. *et al.* Functional analysis of *Arabidopsis* genes involved in mitochondrial iron-sulfur cluster assembly. *Plant Mol Biol* **64**, 225-240 (2007).
75. Liepman, A.H. & Olsen, L.J. Alanine aminotransferase homologs catalyze the glutamate:glyoxylate aminotransferase reaction in peroxisomes of *Arabidopsis*. *Plant Physiol* **131**, 215-227 (2003).
76. Igarashi, D. *et al.* Identification of photorespiratory glutamate:glyoxylate aminotransferase (GGAT) gene in *Arabidopsis*. *Plant J* **33**, 975-987 (2003).
77. Stepanova, A.N. *et al.* TAA1-mediated auxin biosynthesis is essential for hormone crosstalk and plant development. *Cell* **133**, 177-191 (2008).
78. Lam, H.M., Hsieh, M.H. & Coruzzi, G. Reciprocal regulation of distinct asparagine synthetase genes by light and metabolites in *Arabidopsis thaliana*. *Plant J* **16**, 345-353 (1998).
79. Piotrowski, M., Schonfelder, S. & Weiler, E.W. The *Arabidopsis thaliana* isogene NIT4 and its orthologs in tobacco encode beta-cyano-L-alanine hydratase/nitrilase. *J Biol Chem* **276**, 2616-2621 (2001).

**Supplemental Table S2. Protein-coding genes from Plantae plastid genomes of putative non- cyanobacterial origin**

Taxa	<i>Cyanophora paradoxa</i>	<i>Arabidopsis thaliana</i>	<i>Chlamydomonas reinhardtii</i>	<i>Mesostigma viride</i>	<i>Pyropia yezoensis</i>	<i>Cyanidisochoyzon merolae</i>	<i>Cyanidium caldarium</i>
GenBank accession	NC 001675	NC 000932	NC 005353	NC 005353	NC 007932	NC 004799	NC 001840
Protein-coding genes	149	85	69	105	209	207	203
11467418	126022795		195976152	11466449	90994553	30468108	126165876
11467417	7525097		41179071	11466448	90994541	30468247	11465407
11467409	7525092		41179064	11466447	90994535	30468166	11465514
11467404	7525076		41179058	11466445	90994512	30468102	11465461
11467394	7525075		41179042	11466438	90994505	30468200	11465496
11467388	7525066		41179035	11466434	90994485	30468103	11465440
11467353	7525056		41179008	11466422	90994472	30468226	11465403
11467313				11466421	90994461	30468219	11465578
11467289				11466408	90994424	30468133	11465480
11467286				11466359	90994419	30468105	11465409
					90994418	30468104	11465581
					90994417	30468190	11465404
					90994414	30468169	11465481
					90994404	30468165	11465586
						30468147	11465437
						30468090	11465551
						30468097	11465424
						30468106	11465410
						30468123	11465441
						30468228	11465447
						30468191	11465442
						30468155	11465506
						30468187	11465429
						30468204	11465519
						30468071	11465435
						30468197	11465448
						30468180	11465555
						30468170	11465523
						30468192	11465502
						30468205	11465503
						30468202	11465548
						30468201	11465504
						30468144	11465501
						30468059	11465505
						30468156	11465500
						30468251	11465556
						30468196	11465462
						30468074	11465495
						30468110	11465478
						30468063	11465475
						30468129	11465473
						30468185	11465470
						30468049	11465465
						30468055	

30468137  
30468140  
30468141  
30468060  
30468142  
30468250  
30468138  
30468066  
30468053  
30468058  
30468253  
30468139  
30468107  
30468230

---