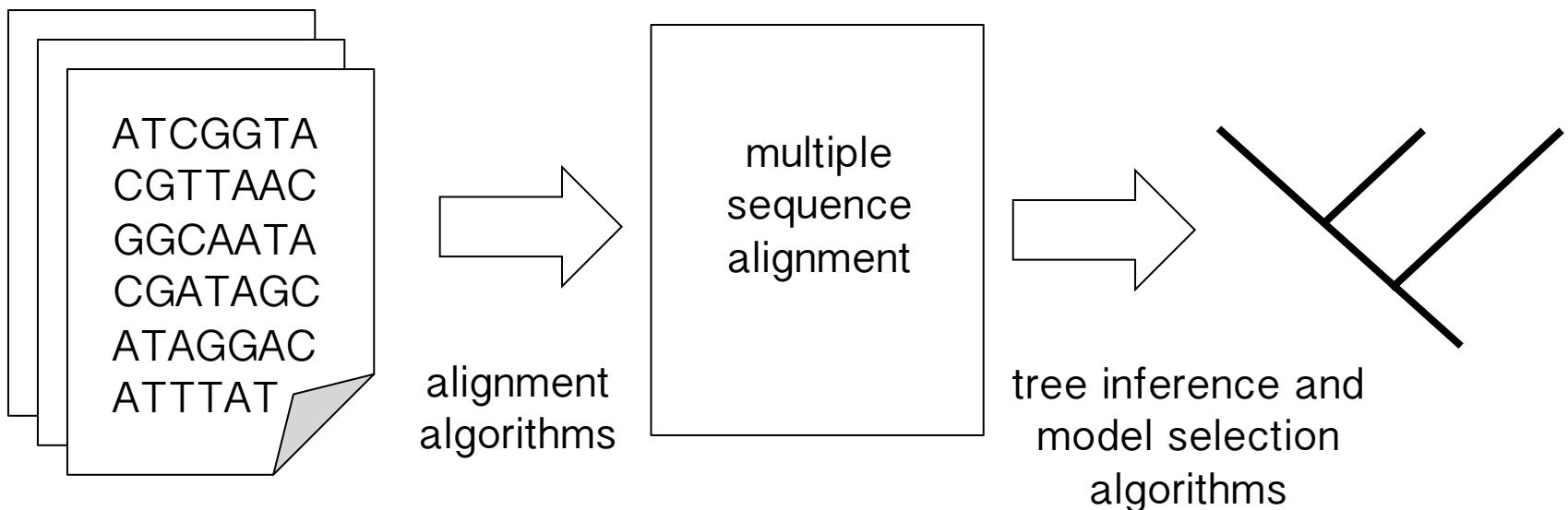
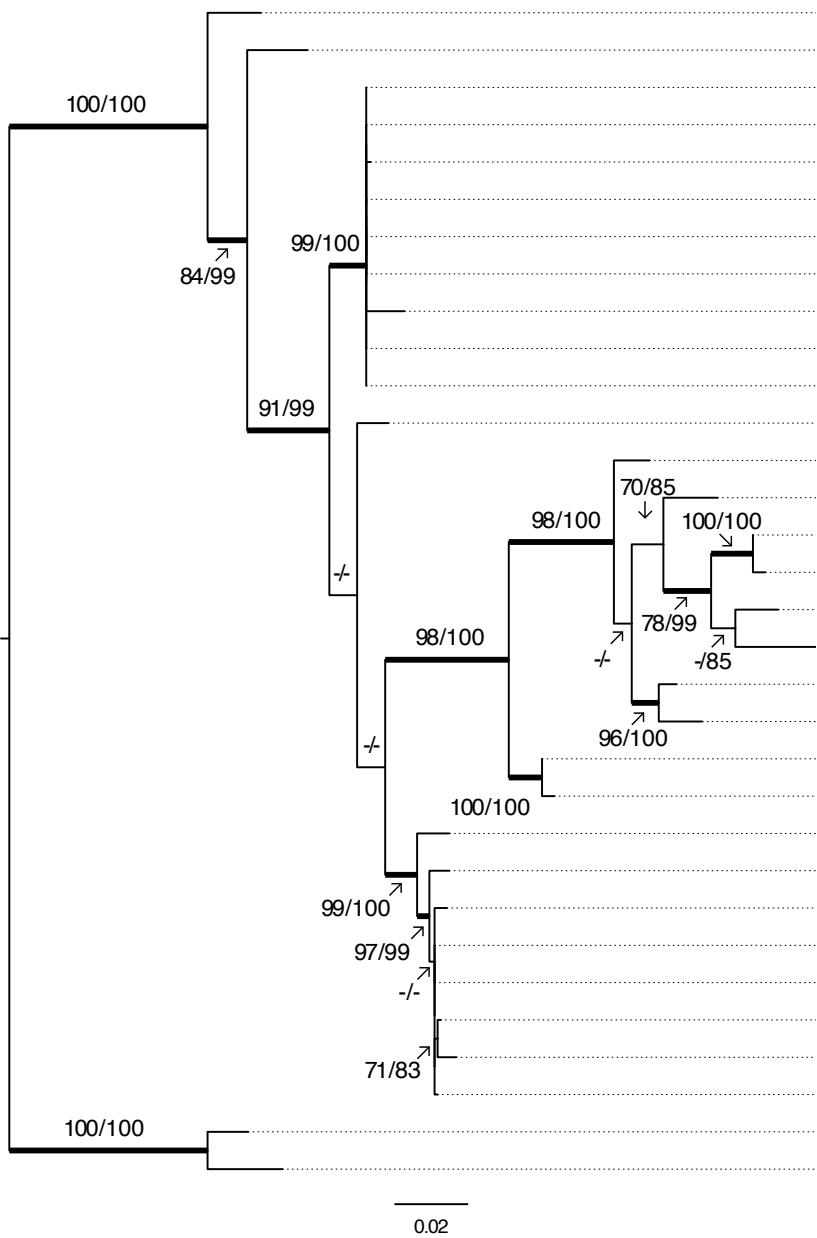


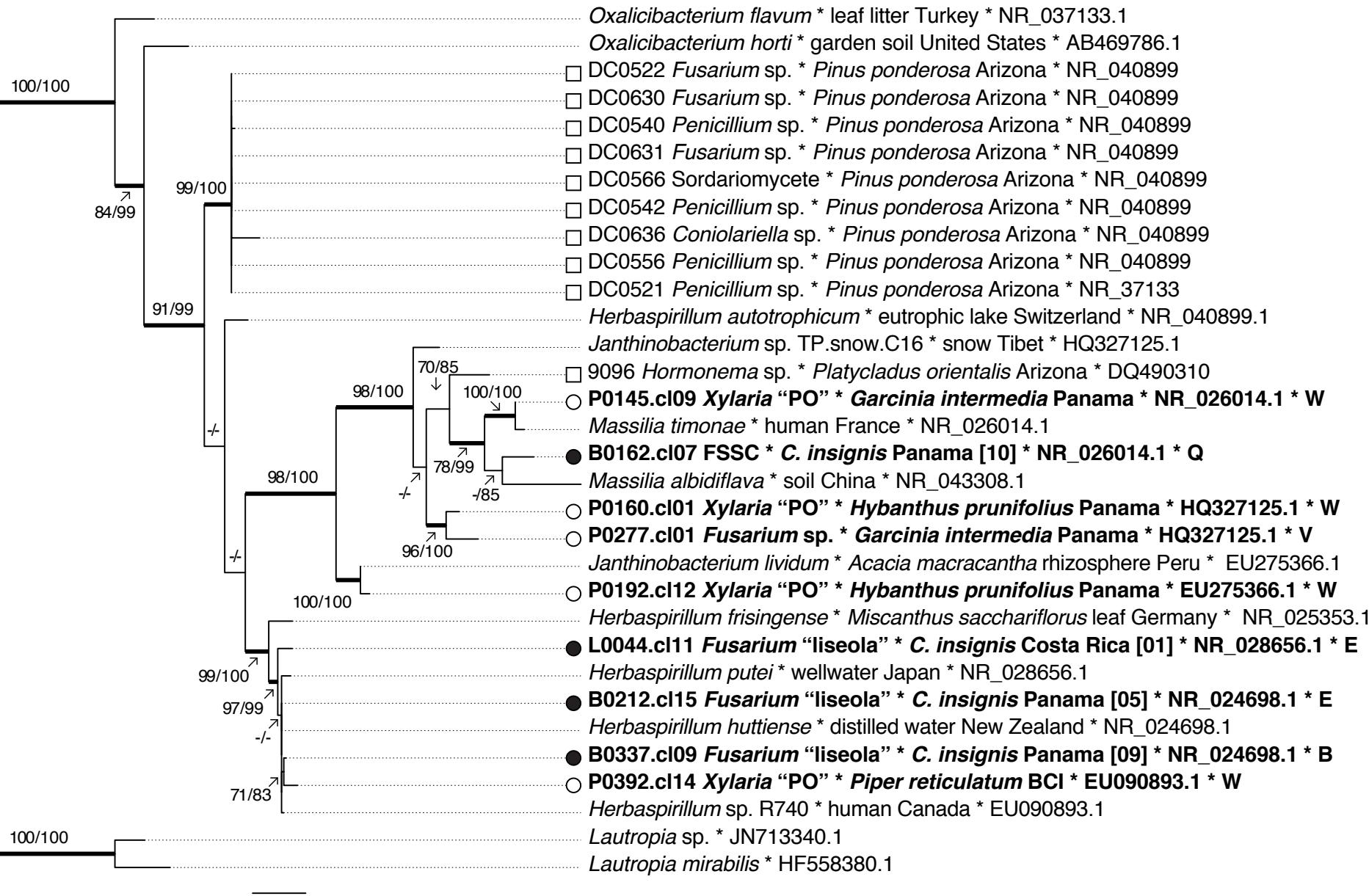


traditional workflow

e.g., ITS region







selecting a group for study

- usually based on preliminary identification of focal isolates using sequence homology

selecting a group for study

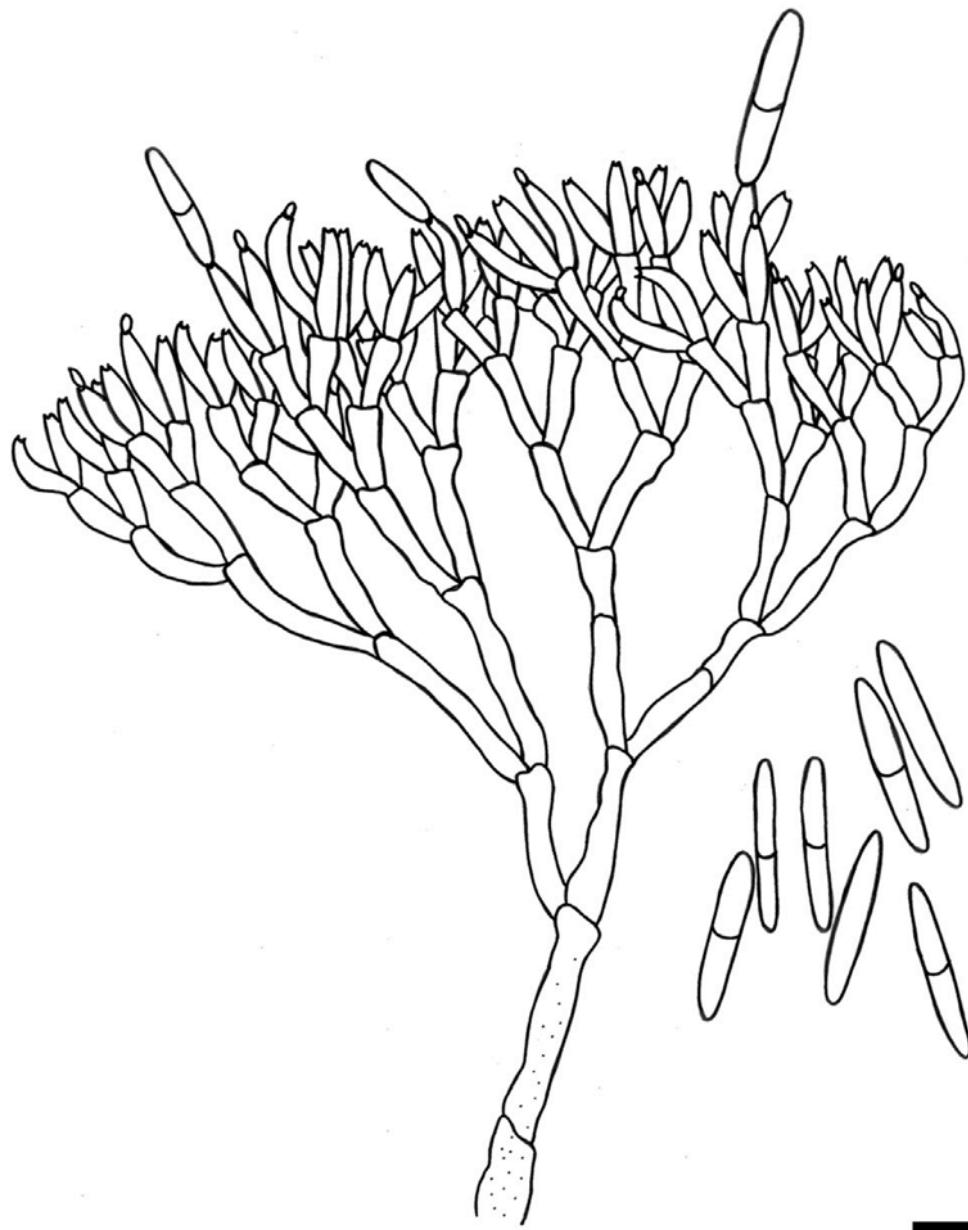
- usually based on preliminary identification of focal isolates using sequence homology
- host plants

selecting a group for study

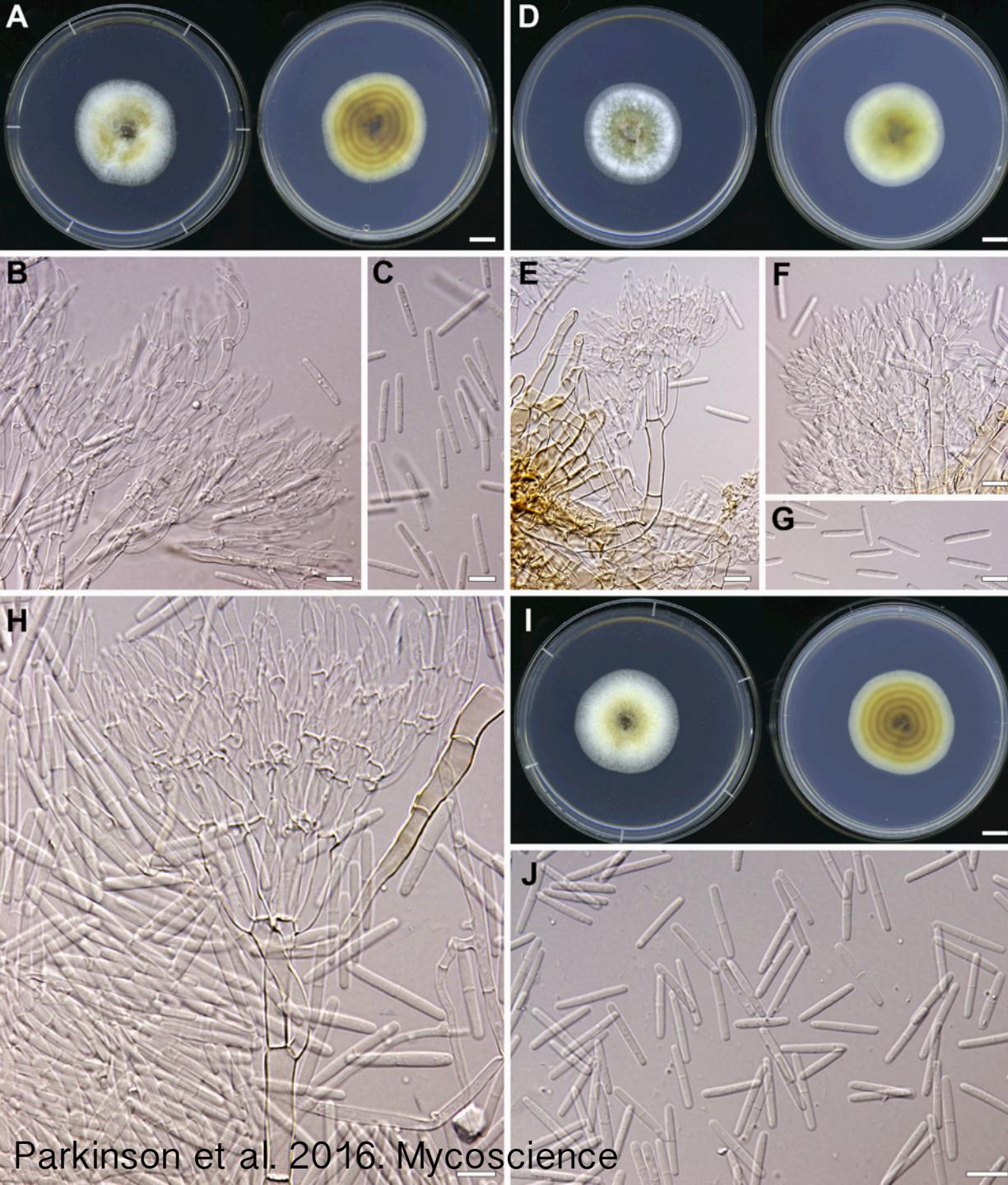
- usually based on preliminary identification of focal isolates using sequence homology
- host plants
- as newer methods and additional data become available

selecting a group for study

- **example:** unidentified isolate from Panama seed project: top BLAST matches for *TEF* and *RPB1* – *Gliocladiopsis pseudotenuis*



10 μm



gather data

- now usually sequence data
- also morphological, chemical, behavioral, and other types of traits

gather data

- **example:** place unknown isolate within *Gliocladiopsis* by sampling other described species

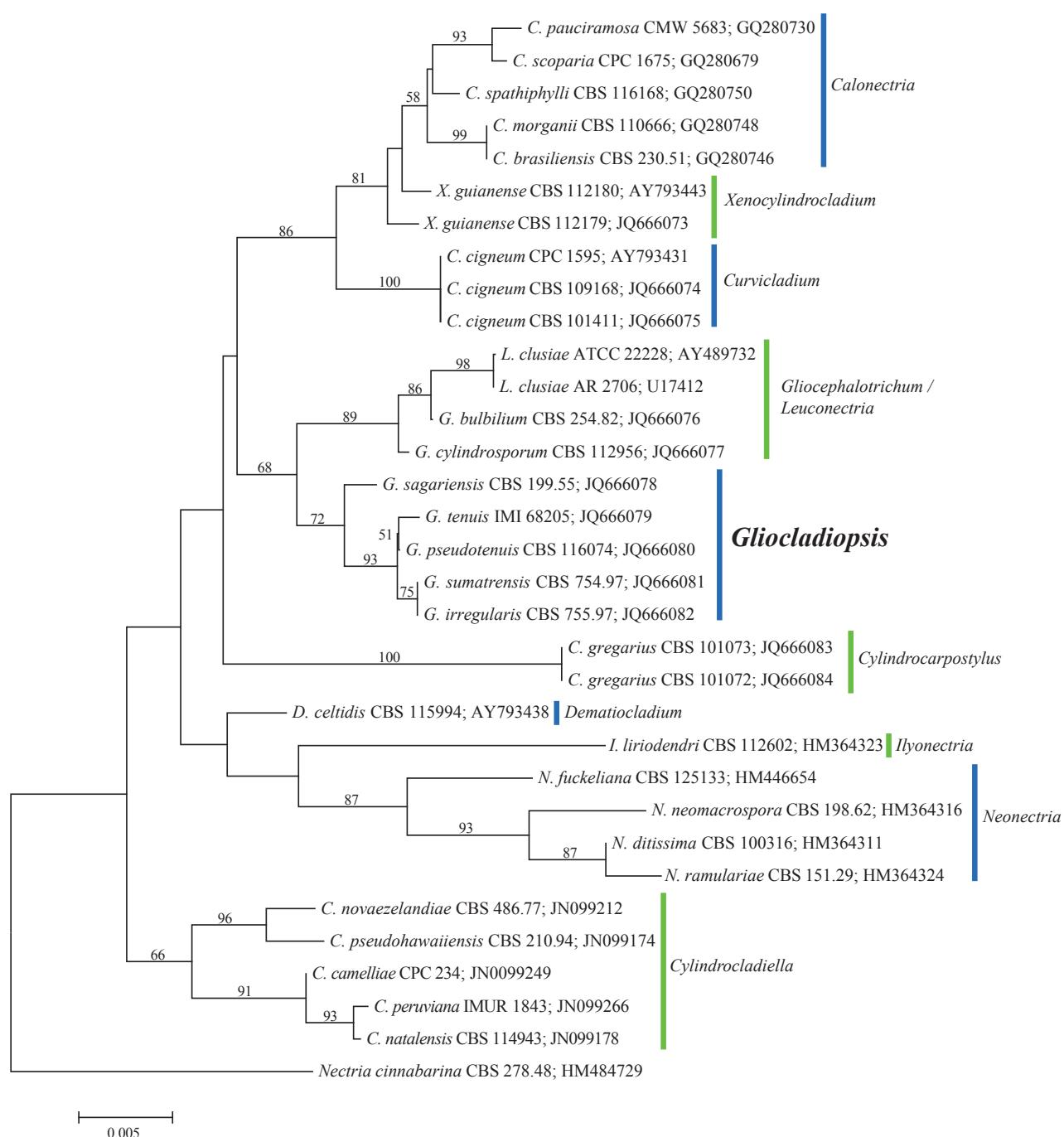
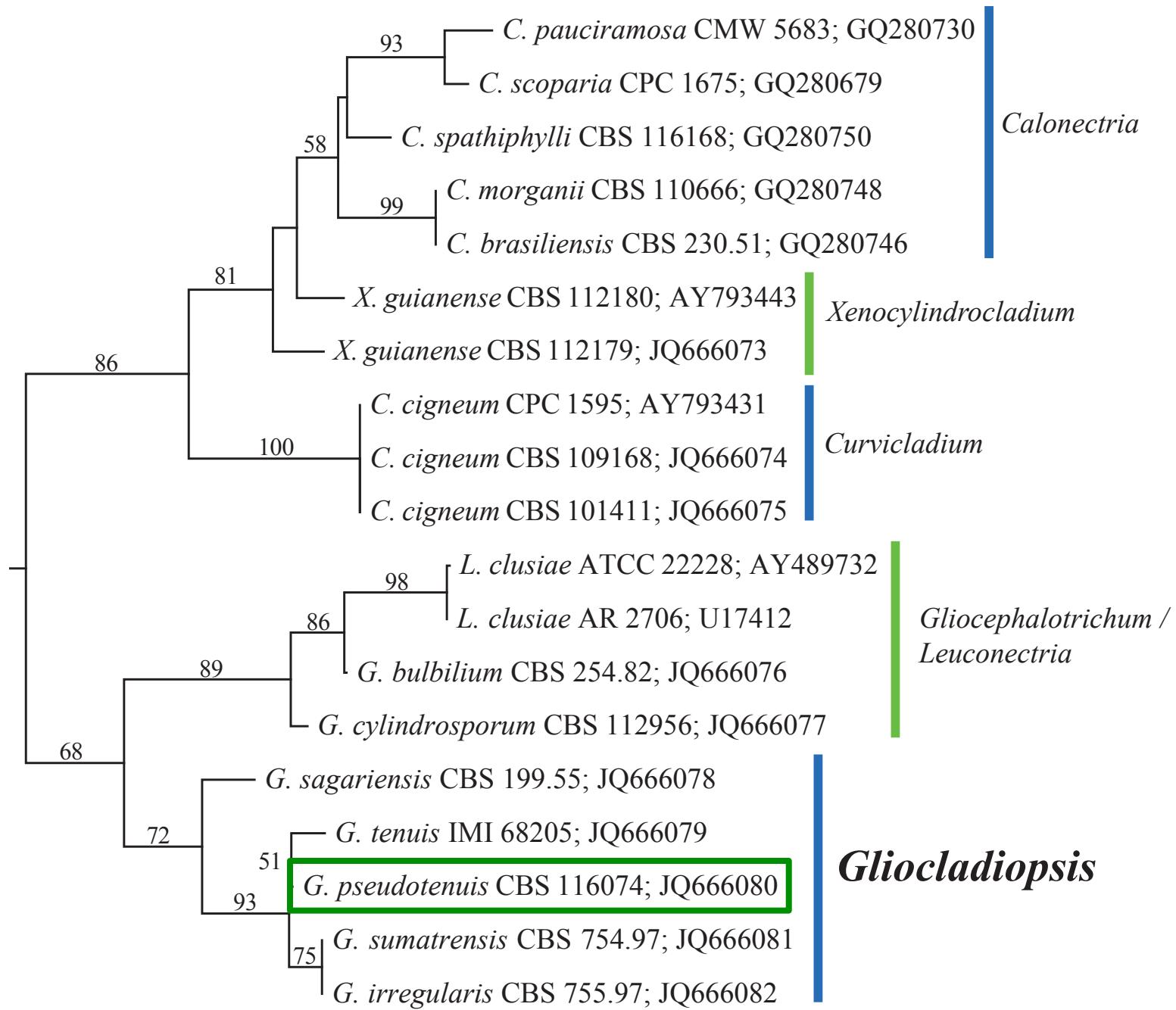


Fig. 1 Neighbour-Joining tree (Kimura-2-parameter) using only the partial LSU sequence alignment with bootstrap values after 1 000 repetitions.



gather data

- **example:** place unknown isolate within *Gliocladiopsis* by sampling other described species

gather data

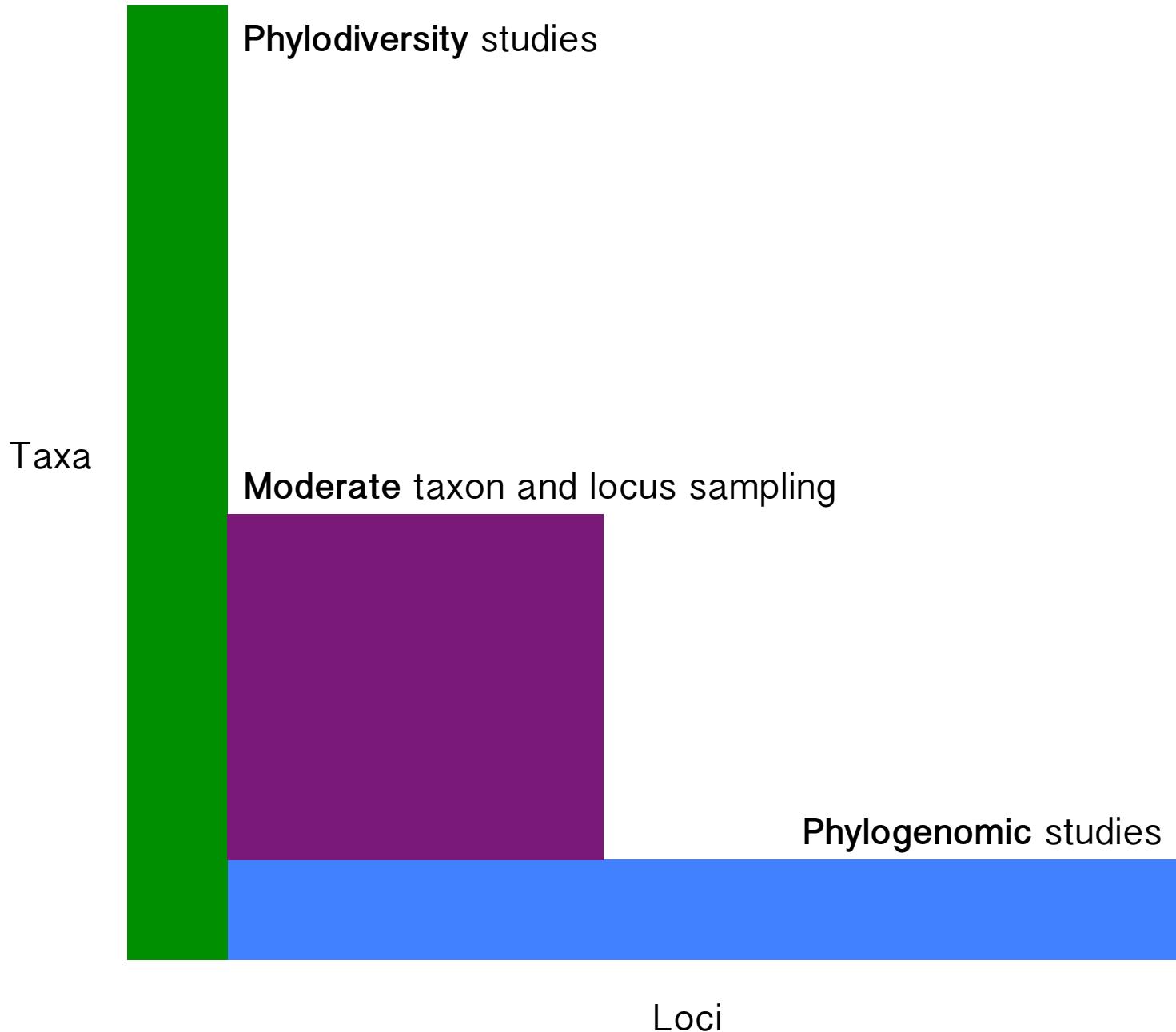
- **example:** place unknown isolate within *Gliocladiopsis* by sampling other described species
- obtain sequences from known groups and/ or type material if possible

gather data

- organize sequences into a FASTA file
 - glio_its.fas

gather data

- organize sequences into a FASTA file
 - glio_its.fas
- use sequence editing programs to view and edit sequences in the form of a matrix
 - examples: Mesquite, Geneious, BLAST2GO



align data

- sequence alignment programs
 - examples: MUSCLE, Clustal, MAFFT, Opal

align data

- sequence alignment programs
 - examples: MUSCLE, Clustal, MAFFT, Opal
 - parameters set penalties for inserting and extending gaps

align data

- **example:** use MUSCLE via CIPRES to align the ITS sequences for our group

glio_its.fas

glio_its_aligned.nex viewed in Mesquite

glio_its_aligned.nex viewed in Mesquite

manually trim ends and verify alignment

choose inference method

choose inference method

- non-model-based:
 - distance methods: neighbor joining

choose inference method

- non-model-based:
 - distance methods: neighbor joining
 - parsimony

choose inference method

- non-model-based:
 - distance methods: neighbor joining
 - parsimony
- model-based:
 - likelihood

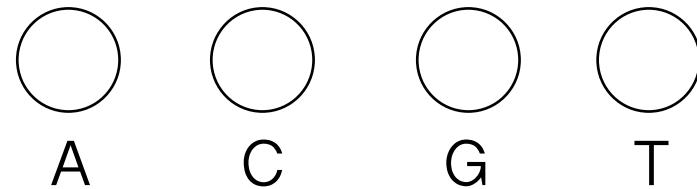
choose inference method

- non-model-based:
 - distance methods: neighbor joining
 - parsimony
- model-based:
 - likelihood
 - Bayesian

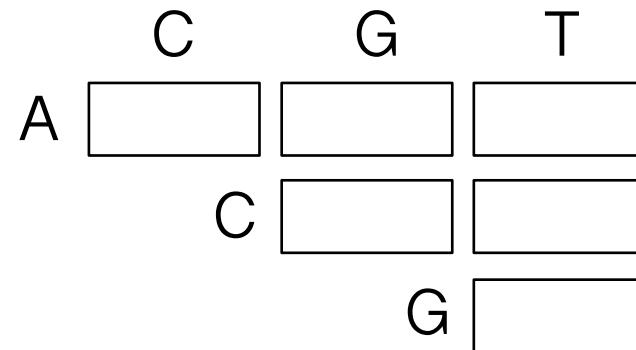
choose models of sequence evolution

- many models
- consider base frequencies and transition and transversion rates

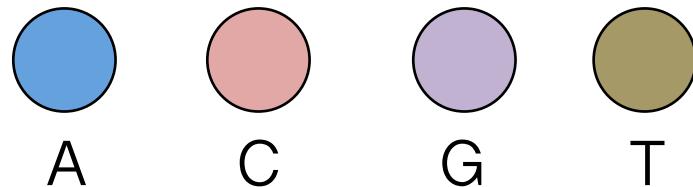
Base frequencies:



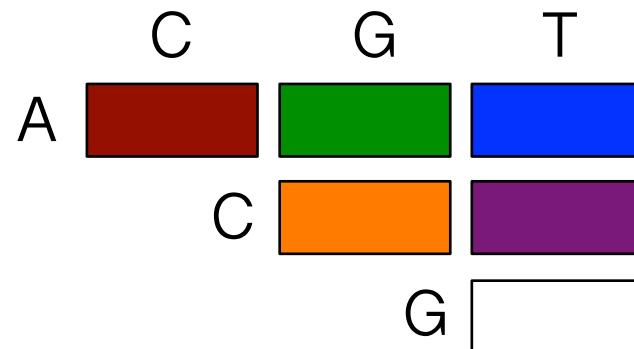
Substitution rates:



Base frequencies:



Substitution rates:



General Time Reversible (GTR): all different

choose models of sequence evolution

- traditional approach is to assign one model to each gene, or to each codon position for amino acids (jModelTest through CIPRES)

choose models of sequence evolution

- traditional approach is to assign one model to each gene, or to each codon position for amino acids (jModelTest through CIPRES)
- moving towards assigning partitions within genes such that multiple models can be used (PartitionFinder)

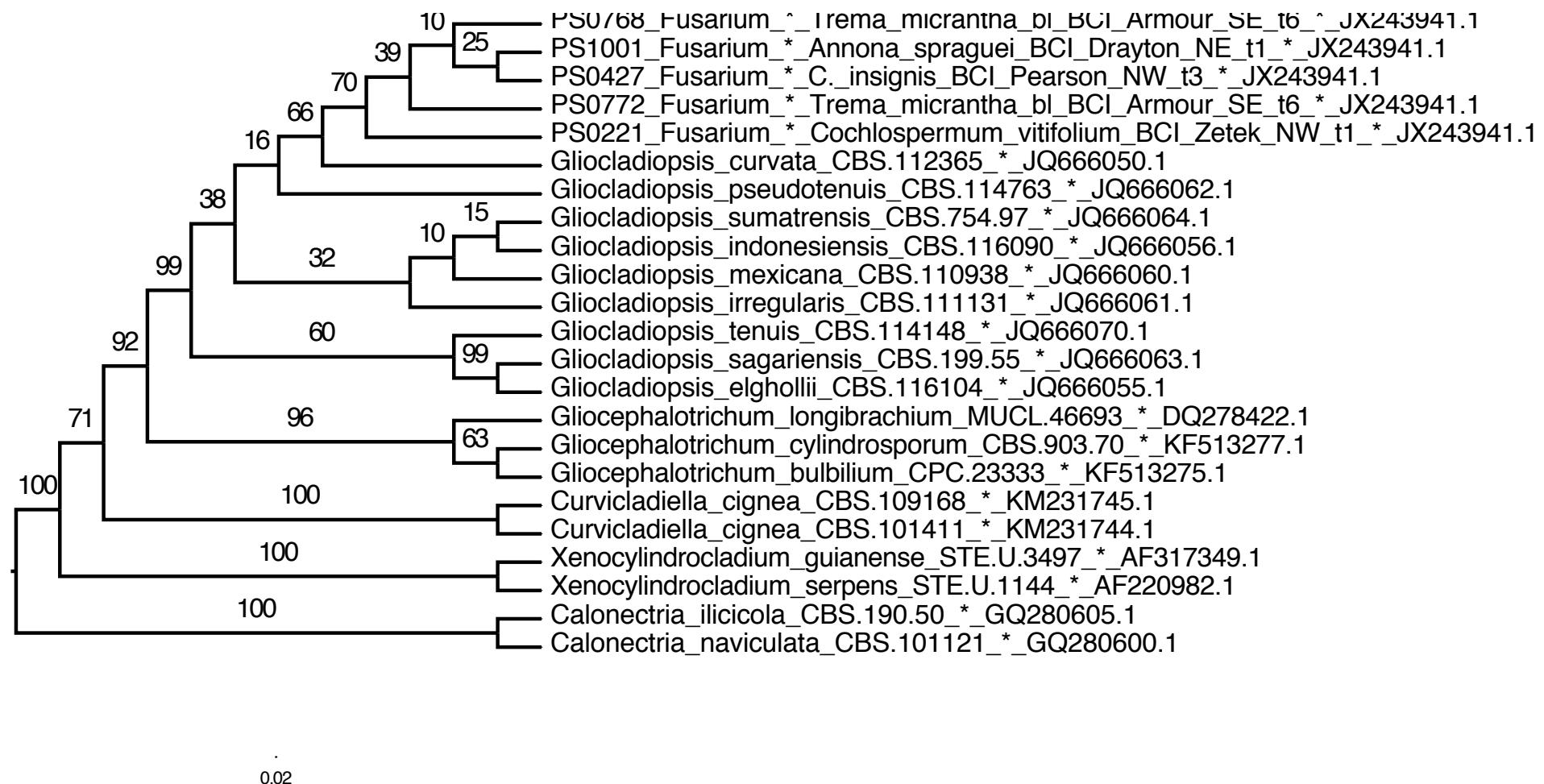
choose models of sequence evolution

- algorithms use alignment to produce AIC and BIC scores for each model

choose models of sequence evolution

- **example:** use maximum likelihood with the GTR+GAMMA+I model (RAxML in CIPRES) to infer phylogeny of focal group

glio_its_MLinvar viewed in FigTree



bootstrap support



bootstrap support

bootstrapping is a method that resamples a single dataset with replacement in order to provide a measure of confidence in the original data.

bootstrap support

original data

Taxon \ Character
1 Alternaria burnii * Dothideomycetes * KR604838.1
2 Morchella sp. * JN081547.1
3 Morchella sp. * JN081549.1
4 3.18.Dannie_Mozzinga.Citrus_aurantium.Site_4
5 Fusarium equiseti * Sordariomycetes * KR094457.1
6 Chaetomium madrasense * Sordariomycetes * JN209900.1
7 6_14.Andres_Santa_Cruz.Citrus_sinensis.Site_5
8 Retroconis fusiformis * Unclassified hymenomycete * EU40239.1
9 Fungal sp. * KJ867413.1
10 4_8.Joshi_Franco.Citrofortunella_mitis.Site_1
11 Penicillium crustosum * Eurotiomycetes * HM469424.1
12 1_14.Iris_Zamudio.Citrus_aurantium.Site_2
13 Dothideomycetes sp. * Q0153254.1
14 3_15.Nereyda_Garcia.Citrus_x_paradisi.Site_4
15 Cladosporium cladosporioides * Dothideomycetes * AY251074.2
16 Cladosporium sp. * Dothideomycetes * JF449608.1
17 Cladosporium sp. * Dothideomycetes * KP143685.1
18 Fungal sp. * KJ543763.1
19 Aureobasidium sp. * Dothideomycetes * FN665417.1

resampled data

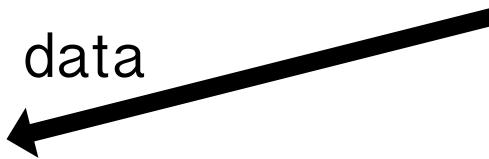
bootstrap support

Taxon \ Character	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1 Alternaria burnii * Dothideomycetes * KR604838.1	G	C	C	A	G	A	C	T	T	G	C	T	C	T	G	T	A	G	C
2 Morchella sp. * JN085147.1	A	C	C	A	G	T	C	G	T	A	G	G	C	A	T	T	G	C	C
3 Morchella sp. * JN085149.1	A	C	C	A	G	T	C	G	T	A	G	G	C	A	T	T	G	C	C
4 3_18.Dannie_Mozinga.Citrus_aurantium.Site_4	A	C	C	A	G	T	C	G	T	A	A	T	C	T	G	G	T	C	C
5 Fusarium equiseti * Sordariomycetes * KR094457.1	A	C	C	A	G	T	C	G	T	A	A	T	C	T	G	G	T	C	C
6 Chaetomium madrasense * Sordariomycetes * JN209900.1	A	C	C	A	G	T	C	G	T	A	A	T	C	T	G	G	T	C	C
7 6_14.Andres_Santa_Cruz.Citrus.sinensis.Site_5	A	C	C	A	G	T	C	G	T	A	A	T	C	T	G	G	T	C	C
8 Retroconis fusiformis * Unclassified hyphomycete * EU040239.1	A	C	C	A	G	T	C	G	T	A	A	T	C	T	G	G	T	C	C
9 Fungal sp. * KJ867413.1	A	C	C	A	G	T	C	G	T	A	A	T	C	T	G	G	T	C	C
10 4_8.Josh_Franco.Citrofortunella_mitis.Site_1	A	C	C	A	G	T	C	G	T	A	A	T	C	T	G	G	T	C	C
11 Penicillium crustosum * Eurotiomycetes * HM469424.1	A	C	C	A	G	T	C	G	T	A	A	T	C	T	G	G	T	C	C
12 1_14.Iris_Zamudio.Citrus_aurantium.Site_2	A	C	C	A	G	T	C	G	T	A	A	T	C	T	G	G	T	C	C
13 Dothideomycetes sp. * GQ153254.1	A	C	C	A	G	T	C	G	T	A	A	T	C	T	G	G	T	C	C
14 3_15.Nereyda_Garcia.Citrus_x_paradisi.Site_4	A	C	C	A	G	T	C	G	T	A	A	T	C	T	G	G	T	C	C
15 Cladosporium cladosporioides * Dothideomycetes * AY251074.2	A	C	C	A	G	T	C	G	T	A	A	T	C	T	G	G	T	C	C
16 Cladosporium sp. * Dothideomycetes * JF449608.1	A	C	C	A	G	T	C	G	T	A	A	T	C	T	G	G	T	C	C
17 Cladosporium sp. * Dothideomycetes * KP143685.1	A	C	C	A	G	T	C	G	T	A	A	T	C	T	G	G	T	C	C
18 Fungal sp. * KJ543763.1	A	T	C	A	G	T	C	G	T	A	A	T	C	T	G	G	T	C	C
19 Aureobasidium sp. * Dothideomycetes * FN665417.1	A	T	C	A	G	T	C	G	T	A	A	T	C	T	G	G	T	C	C

original data

Taxon \ Character	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1 Alternaria burnii * Dothideomycetes * KR604838.1	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
2 Morchella sp. * JN085147.1	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
3 Morchella sp. * JN085149.1	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
4 3_18.Dannie_Mozinga.Citrus_aurantium.Site_4	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
5 Fusarium equiseti * Sordariomycetes * KR094457.1	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
6 Chaetomium madrasense * Sordariomycetes * JN209900.1	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
7 6_14.Andres_Santa_Cruz.Citrus.sinensis.Site_5	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
8 Retroconis fusiformis * Unclassified hyphomycete * EU040239.1	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
9 Fungal sp. * KJ867413.1	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
10 4_8.Josh_Franco.Citrofortunella_mitis.Site_1	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
11 Penicillium crustosum * Eurotiomycetes * HM469424.1	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
12 1_14.Iris_Zamudio.Citrus_aurantium.Site_2	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
13 Dothideomycetes sp. * GQ153254.1	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
14 3_15.Nereyda_Garcia.Citrus_x_paradisi.Site_4	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
15 Cladosporium cladosporioides * Dothideomycetes * AY251074.2	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
16 Cladosporium sp. * Dothideomycetes * JF449608.1	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
17 Cladosporium sp. * Dothideomycetes * KP143685.1	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
18 Fungal sp. * KJ543763.1	C	T	C	A	G	T	C	G	T	A	A	T	C	T	G	G	T	C	C
19 Aureobasidium sp. * Dothideomycetes * FN665417.1	A	T	C	A	G	T	C	G	T	A	A	T	C	T	G	G	T	C	C

resampled data



bootstrap support

Taxon \ Character	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1 Alternaria burnii * Dothideomycetes * KR604838.1	C	C	C	A	G	A	C	T	T	C	A	C	C	G	G	C	A	T	C
2 Morchella sp. * JN085147.1	A	C	A	G	C	T	C	G	T	A	G	C	C	G	T	G	C	C	C
3 Morchella sp. * JN085149.1	A	C	A	G	C	T	C	G	T	A	G	C	C	G	T	G	C	C	C
4 3_18.Dannie_Mozinga.Citrus_aurantium.Site_4	A	C	A	G	C	T	G	G	T	A	A	C	T	G	G	T	C	C	C
5 Fusarium equiseti * Sordariomycetes * KR094457.1	A	C	A	G	C	T	G	G	T	A	A	C	T	G	G	T	C	C	C
6 Chaetomium madrasense * Sordariomycetes * JN209900.1	A	C	A	G	C	T	G	G	T	A	A	C	T	G	G	T	C	C	C
7 6_14.Andres_Santa_Cruz.Citrus.sinensis.Site_5	A	C	A	G	C	T	G	G	T	A	A	C	T	G	G	T	C	C	C
8 Retroconis fusiformis * Unclassified hyphomycete * EU040239.1	A	C	A	G	C	T	G	G	T	A	A	C	T	G	G	T	C	C	C
9 Fungal sp. * KJ867413.1	A	C	A	G	C	T	G	G	T	A	A	C	T	G	G	T	C	C	C
10 4_8.Josh_Franco.Citrofortunella_mitis.Site_1	A	C	A	G	C	T	G	G	T	A	A	C	T	G	G	T	C	C	C
11 Penicillium crustosum * Eurotiomycetes * HM469424.1	A	C	A	G	C	T	G	G	T	A	A	C	T	G	G	T	C	C	C
12 1_14.Iris_Zamudio.Citrus_aurantium.Site_2	A	C	A	G	C	T	G	G	T	A	A	C	T	G	G	T	C	C	C
13 Dothideomycetes sp. * GQ153254.1	A	C	A	G	C	T	G	G	T	A	A	C	T	G	G	T	C	C	C
14 3_15.Nereyda_Garcia.Citrus_x_paradisi.Site_4	A	C	A	G	C	T	G	G	T	A	A	C	T	G	G	T	C	C	C
15 Cladosporium cladosporioides * Dothideomycetes * AY251074.2	A	C	A	G	C	T	G	G	T	A	A	C	T	G	G	T	C	C	C
16 Cladosporium sp. * Dothideomycetes * JF449608.1	A	C	A	G	C	T	G	G	T	A	A	C	T	G	G	T	C	C	C
17 Cladosporium sp. * Dothideomycetes * KP143685.1	A	C	A	G	C	T	G	G	T	A	A	C	T	G	G	T	C	C	C
18 Fungal sp. * KJ543763.1	A	T	C	A	G	T	T	T	A	A	C	T	G	T	T	A	C	G	T
19 Aureobasidium sp. * Dothideomycetes * FN665417.1	A	T	C	A	G	T	T	T	A	A	C	T	G	T	T	A	C	G	T

original data ↗

Taxon \ Character	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1 Alternaria burnii * Dothideomycetes * KR604838.1	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
2 Morchella sp. * JN085147.1	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
3 Morchella sp. * JN085149.1	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
4 3_18.Dannie_Mozinga.Citrus_aurantium.Site_4	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
5 Fusarium equiseti * Sordariomycetes * KR094457.1	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
6 Chaetomium madrasense * Sordariomycetes * JN209900.1	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
7 6_14.Andres_Santa_Cruz.Citrus.sinensis.Site_5	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
8 Retroconis fusiformis * Unclassified hyphomycete * EU040239.1	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
9 Fungal sp. * KJ867413.1	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
10 4_8.Josh_Franco.Citrofortunella_mitis.Site_1	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
11 Penicillium crustosum * Eurotiomycetes * HM469424.1	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
12 1_14.Iris_Zamudio.Citrus_aurantium.Site_2	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
13 Dothideomycetes sp. * GQ153254.1	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
14 3_15.Nereyda_Garcia.Citrus_x_paradisi.Site_4	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
15 Cladosporium cladosporioides * Dothideomycetes * AY251074.2	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
16 Cladosporium sp. * Dothideomycetes * JF449608.1	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
17 Cladosporium sp. * Dothideomycetes * KP143685.1	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
18 Fungal sp. * KJ543763.1	C	T	C	A	G	T	T	T	A	A	C	T	G	T	T	A	C	G	T
19 Aureobasidium sp. * Dothideomycetes * FN665417.1	C	T	C	A	G	T	T	T	A	A	C	T	G	T	T	A	C	G	T

resampled data

bootstrap support

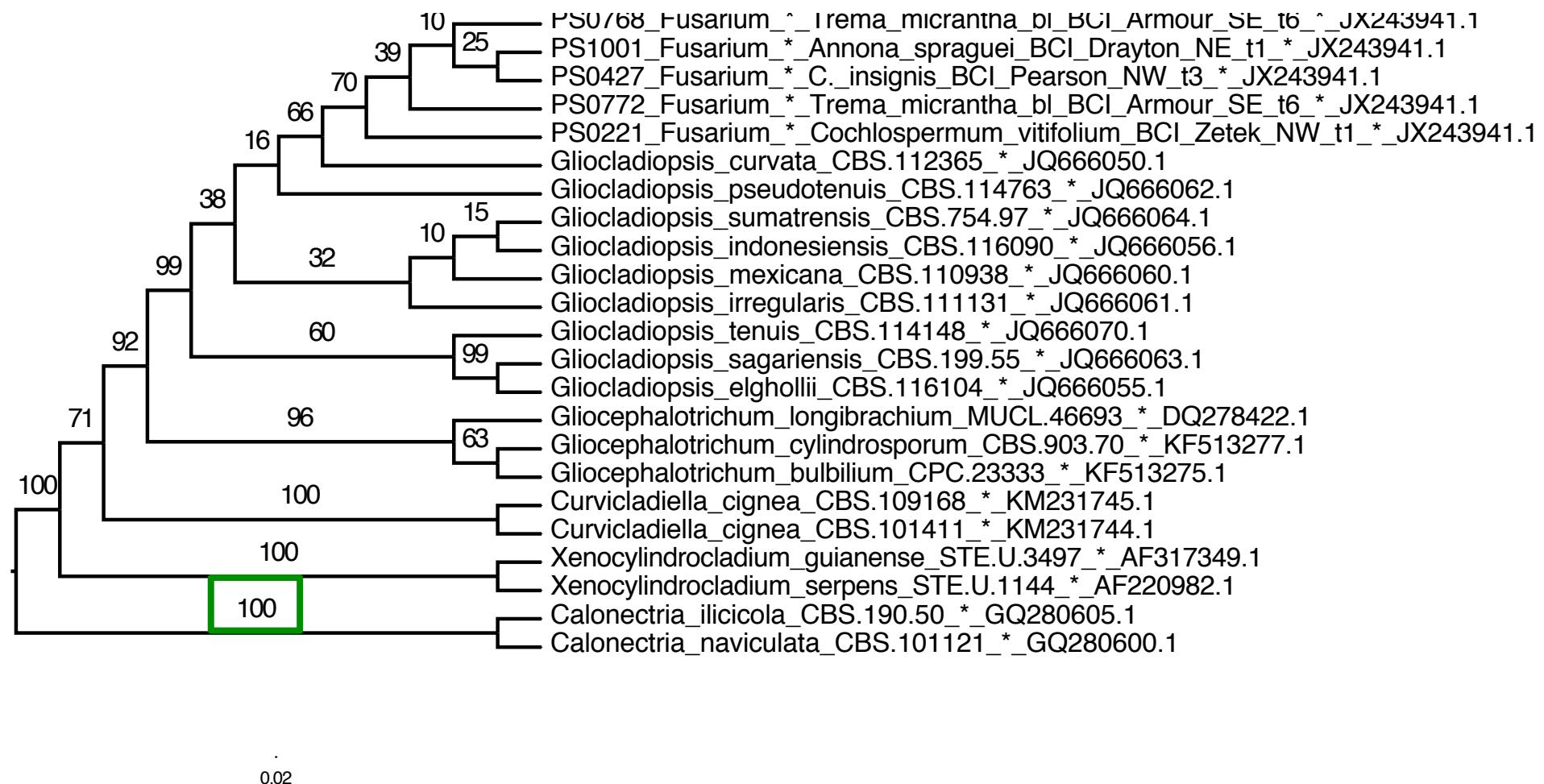
Taxon \ Character	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	599	600	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	651	652	653	654	655	656	657	658	659	660	661	662	663	664	665	666	667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684	685	686	687	688	689	690	691	692	693	694	695	696	697	698	699	700	701	702	703	704	705	706	707	708	709	710	711	712	713	714	715	716	717	718	719	720	721	722	723	724	725	726	727	728	729	730	731	732	733	734	735	736	737	738	739	740	741	742	743	744	745	746	747	748	749	750	751	752	753	754	755	756	757	758	759	760	761	762	763	764	765	766	767	768	769	770	771	772	773	774	775	776	777	778	779	780	781	782	783	784	785	786	787	788	789	790	791	792	793	794	795	796	797	798	799	800	801	802	803	804	805	806	807	808	809	8010	8011	8012	8013	8014	8015	8016	8017	8018	8019	8020	8021	8022	8023	8024	8025	8026	8027	8028	8029	8030	8031	8032	8033	8034	8035	8036	8037	8038	8039	8040	8041	8042	8043	8044	8045	8046	8047	8048	8049	8050	8051	8052	8053	8054	8055	8056	8057	8058	8059	8060	8061	8062	8063	8064	8065	8066	8067	8068	8069	8070	8071	8072	8073	8074	8075	8076	8077	8078	8079	8080	8081	8082	8083	8084	8085	8086	8087	8088	8089	8090	8091	8092	8093	8094	8095	8096	8097	8098	8099	80100	80101	80102	80103	80104	80105	80106	80107	80108	80109	80110	80111	80112	80113	80114	80115	80116	80117	80118	80119	80120	80121	80122	80123	80124	80125	80126	80127	80128	80129	80130	80131	80132	80133	80134	80135	80136	80137	80138	80139	80140	80141	80142	80143	80144	80145	80146	80147	80148	80149	80150	80151	80152	80153	80154	80155	80156	80157	80158	80159	80160	80161	80162	80163	80164	80165	80166	80167	80168	80169	80170	80171	80172	80173	80174	80175	80176	80177	80178	80179	80180	80181	80182	80183	80184	80185	80186	80187	80188	80189	80190	80191	80192	80193	80194	80195	80196	80197	80198	80199	80200	80201	80202	80203	80204	80205	80206	80207	80208	80209	80210	80211	80212	80213	80214	80215	80216	80217	80218	80219	80220	80221	80222	80223	80224	80225	80226	80227	80228	80229	80230	80231	80232	80233	80234	80235	80236	80237	80238	80239	80240	80241	80242	80243	80244	80245	80246	80247	80248	80249	80250	80251	80252	80253	80254	80255	80256	80257	80258	80259	80260	80261	80262	80263	80264	80265	80266	80267	80268	80269	80270	80271	80272	80273	80274	80275	80276	80277	80278	80279	80280	80281	80282	80283	80284	80285	80286	80287	80288	80289	80290	80291	80292	80293	80294	80295	80296	80297	80298	80299	80300	80301	80302	80303	80304	80305	80306	80307	80308	80309	80310	80311	80312	80313	80314	80315	80316	80317	80318	80319	80320	80321	80322	80323	80324	80325	80326	80327	80328	80329	80330	80331	80332	80333	80334	80335	80336	80337	80338	80339	80340	80341	80342	80343	80344	80345	80346	80347	80348	80349	80350	80351	80352	80353	80354	80355	80356	80357	80358	80359	80360	80361	80362	80363	80364	80365	80366	80367	80368	80369	80370	80371	80372	80373	80374	80375	80376	80377	80378	80379	80380	80381	80382	80383	80384	80385	80386	80387	80388	80389	80390	80391	80392	80393	80394	80395	80396	80397	80398	80399	80400	80401	80402	80403	80404	80405	80406	80407	80408	80409	80410	80411	80412	80413	80414	80415	80416	80417	80418	80419	80420	80421	80422	80423	80424	80425	80426	80427	80428	80429	80430	80431	80432	80433	80434	80435	80436	80437	80438	80439	80440	80441	80442	80443	80444	80445	80446	80447	80448	80449	80450	80451	80452	80453	80454	80455	80456	80457	80458	80459	80460	80461	80462	80463	80464	80465	80466	80467	80468	80469	80470	80471	80472	80473	80474	80475	80476	80477	80478	80479	80480	80481	80482	80483	80484	80485	80486	80487	80488	80489	80490	80491	80492	80493	80494	80495	80496	80497	80498	80499	80500	80501	80502	80503	80504	80505	80506	80507	80508	80509	80510	80511	80512	80513	80514	80515	80516	80517	80518	80519	80520	80521	80522	80523	80524	80525	80526	80527	80528	80529	80530	80531	80532	80533	80534	80535	80536	80537	80538	80539	80540	80541	80542	80543	80544	80545	80546	80547	80548	80549	80550	80551	80552	80553	80554	80555	80556	80557	80558	80559	80560	80561	80562	80563	80564	80565	80566	80567	80568	80569	80570	80571	80572	80573	80574	80575	80576	80577	80578	80579	80580	80581	80582	80583	80584	80585	80586	80587	80588	80589	80590	80591	80592	80593	80594	80595	8059

bootstrap support

bootstrapping is a method that resamples a single dataset with replacement in order to provide a measure of confidence in the original data.

once the data are resampled the tree is inferred, and this is done typically 1,000 times.

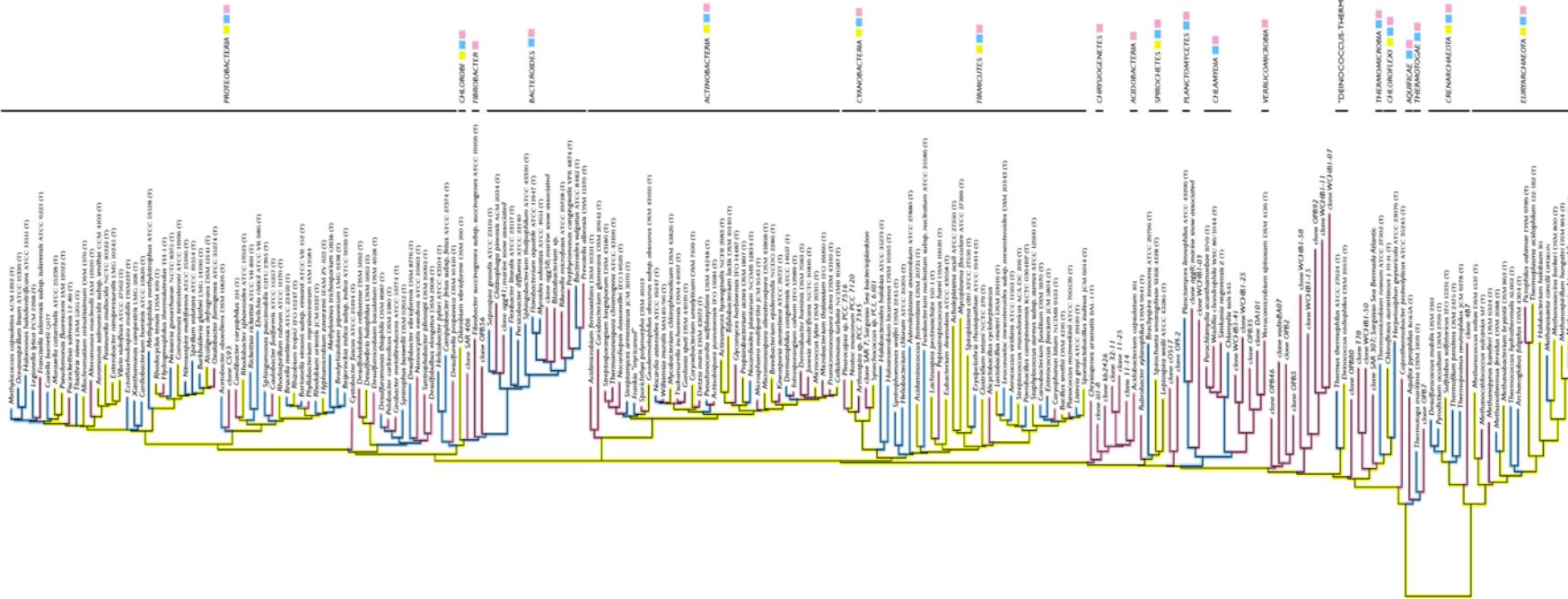
glio_its_MLinvar viewed in FigTree



other considerations

other considerations

- constraint trees



The RDP-II Backbone Tree for Reference 8. The tree was inferred from a distance matrix generated in PAUP* with the Wagner (weighted by ploidy) method. The 177 sequences represent about 18% of the 203 economic bacteria families (20 of the 25 phyla) in the tree. The 1997 edition of the *Bergey's Manual of Systematic Bacteriology* was used as the source of taxonomic names. The tree was rooted with *Thiotricha*. The DNA sequences were first subjected to phylogenetic analysis based on 16S rRNA sequence in 1990. Yellow and blue taxa were included in the first RDP tree in 1992.

other considerations

- available alignments and trees

The screenshot shows the homepage of the TreeBASE website. At the top, there is a dark blue header bar with the TreeBASE logo and the tagline "A Database of Phylogenetic Knowledge". Below the header, the main content area has a white background. On the left side, there is a sidebar with a vertical navigation menu. The menu items include: Search TreeBASE, Submission Tutorial, Submit, About (which is currently selected, indicated by a blue border), Overview, Technology, People, Partnerships, References, NSF Data Management, Data Access, Journals, and Contact. To the right of the sidebar, the main content area features a large title "Welcome to TreeBASE". Below the title, there is a paragraph of text describing the purpose and scope of TreeBASE. Further down, there is another paragraph providing specific statistics about the current dataset, followed by a section listing new features in the current release. At the bottom of the page, there are icons for social media (Twitter and LinkedIn) and a circular badge that says "Covered by DATA CITATION INDEX".

Welcome to TreeBASE

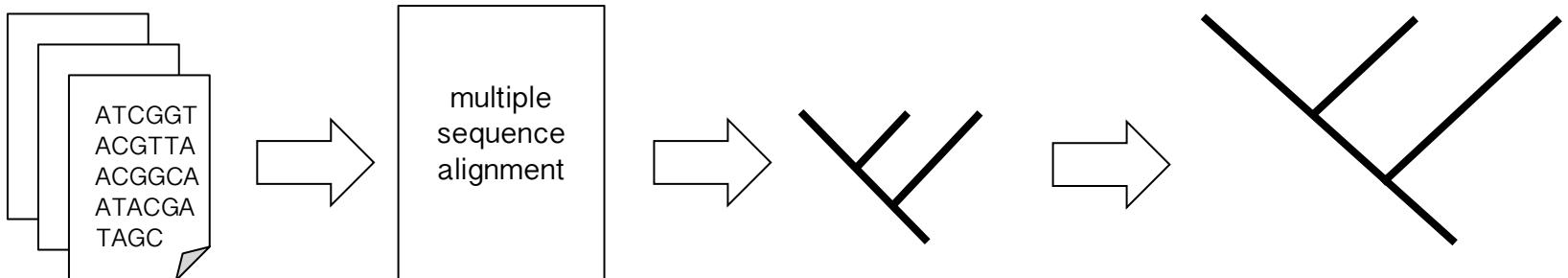
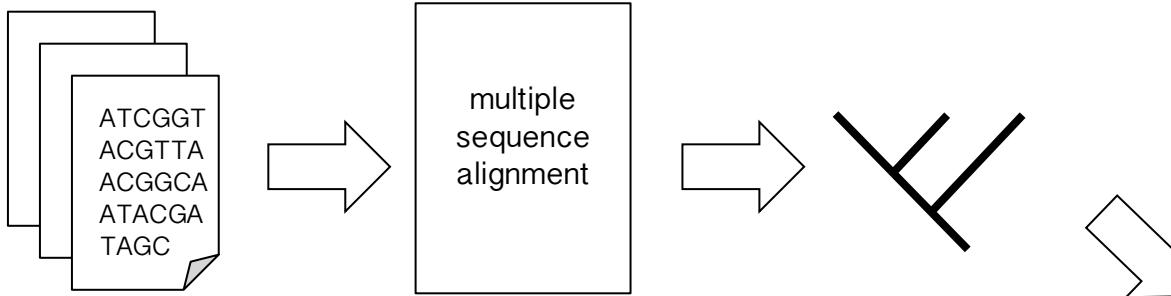
TreeBASE is a repository of phylogenetic information, specifically user-submitted phylogenetic trees and the data used to generate them. TreeBASE accepts all kinds of phylogenetic data (e.g., trees of species, trees of populations, trees of genes) representing all biotic taxa. Data in TreeBASE are exposed to the public if they are used in a publication that is in press or published in a peer-reviewed scientific journal, book, conference proceedings, or thesis. Data used in publications that are in preparation or in review can be submitted to TreeBASE but will not be available to the public until they have passed peer review. Aside from the submitter, such data are only available to the publication editors or reviewers using a special access URL. TreeBASE is produced and governed by the [The Phyloinformatics Research Foundation, Inc.](#)

As of April 2014, TreeBASE contains data for 4,076 publications written by 8,777 different authors. These studies analyzed 8,233 matrices and resulted in 12,817 trees with 761,460 taxon labels that mapped to 104,593 distinct taxa.

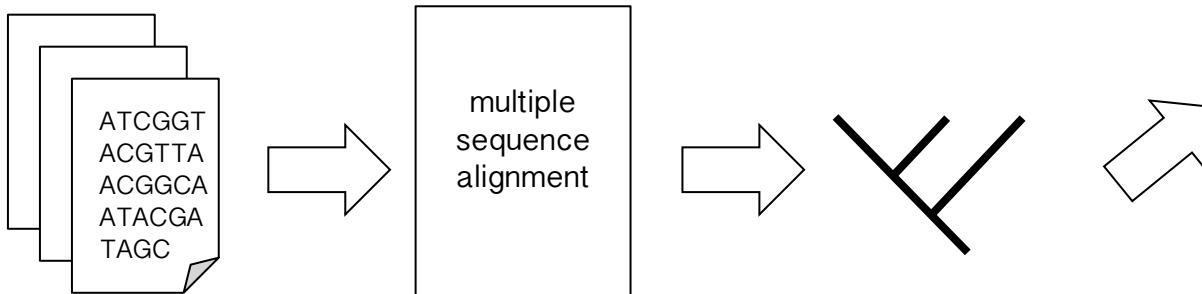
The current release includes a host of new features and improvements over the previous TreeBASE prototype. New features include:

- Richer annotation of metadata (journal DOIs, specimen georeferences, Genbank accession numbers, etc)
- A mapping between taxon labels and taxonomic names in uBio and NCBI for improved normalization of names
- The ability to visualize and edit trees using PhyloWidget
- The ability to search on tree topology
- Persistent and resolvable URLs for data objects in TreeBASE (i.e. studies, trees, matrices) serve as both globally unique identification numbers and resource locators. These can be included in articles and on researcher's websites, making access to TreeBASE data only a click away
- Data are delivered in several serializations, including **NEXUS** and **NeXML**
- A special URL gives journal editors and reviewers anonymous advanced access to data
- Programmatic access to the data using the **PhyloWS API**. Queries are expressed in URLs using PhyloWS syntax and can return results in RDF as RSS 1.0 feeds, which means that users can set their favorite RSS Reader to fetch all new TreeBASE studies that satisfy a particular query (e.g. return all studies published in "Systematic Biology," or return all trees that include "Homo sapiens," etc)

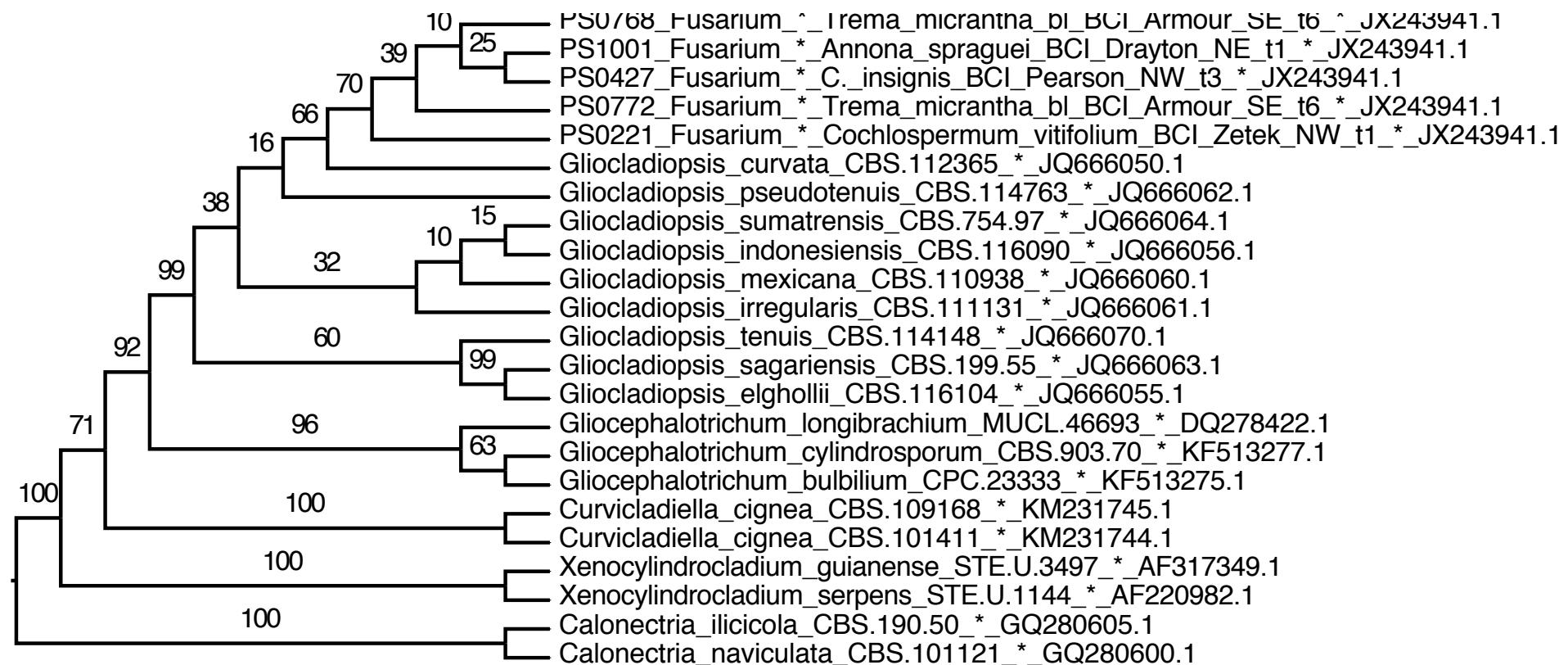
updated workflow



supertree
algorithms

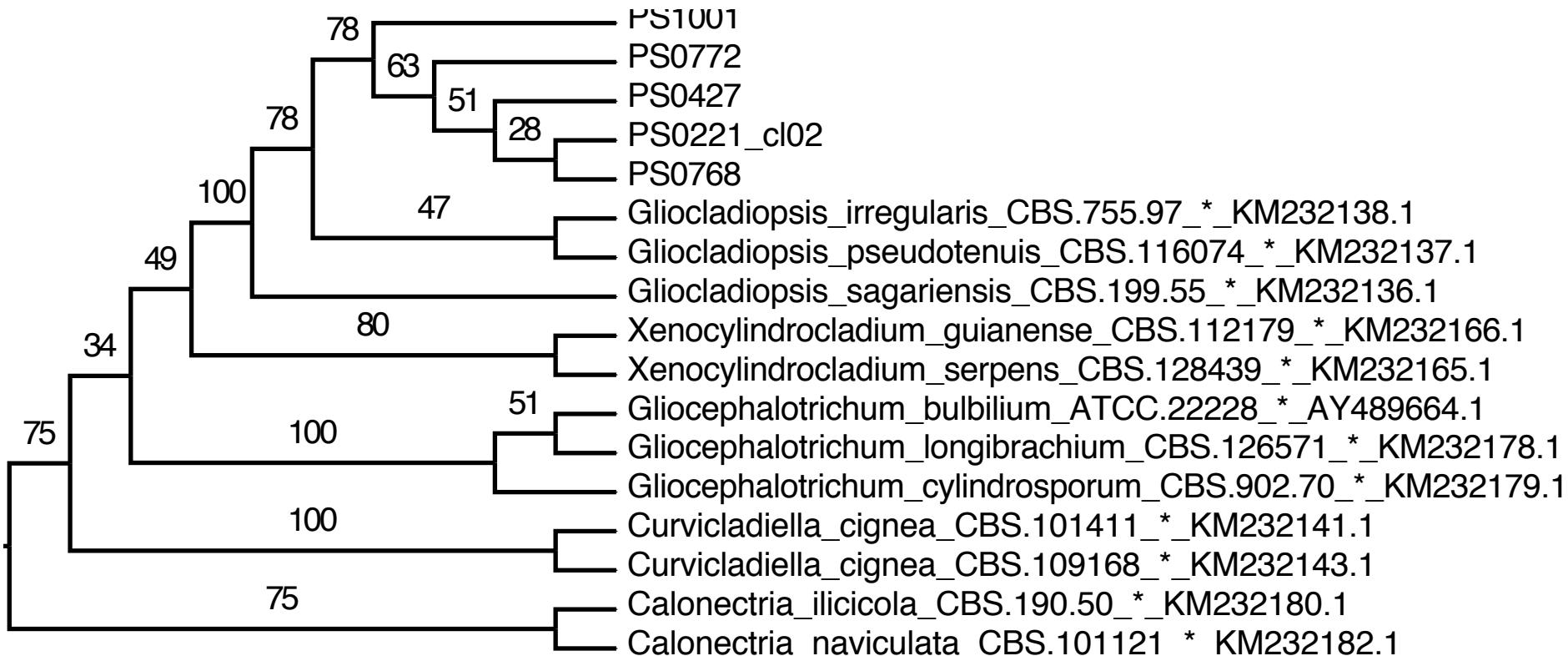


ITS 23 taxa



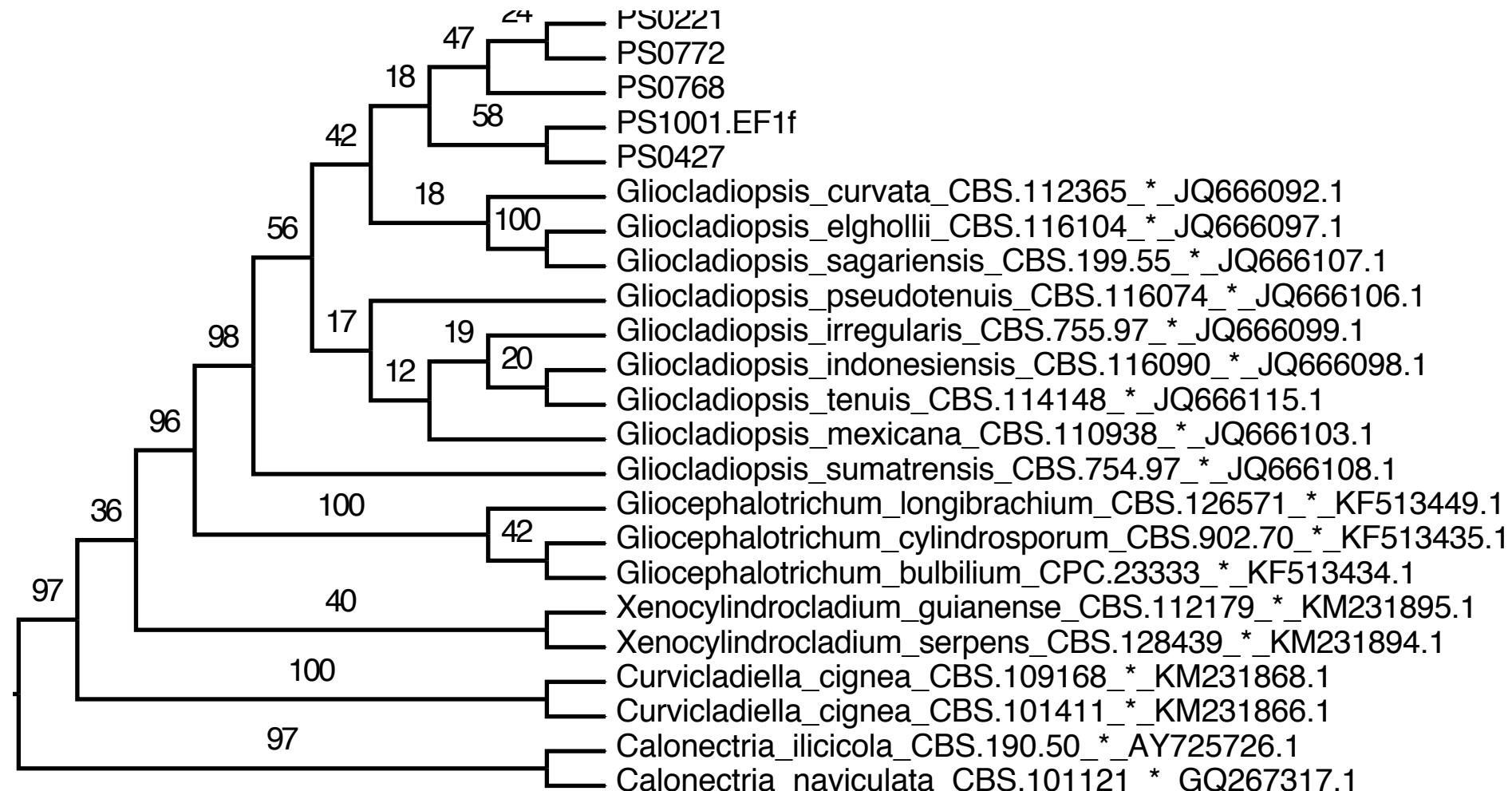
RPB1

17 taxa



TEF

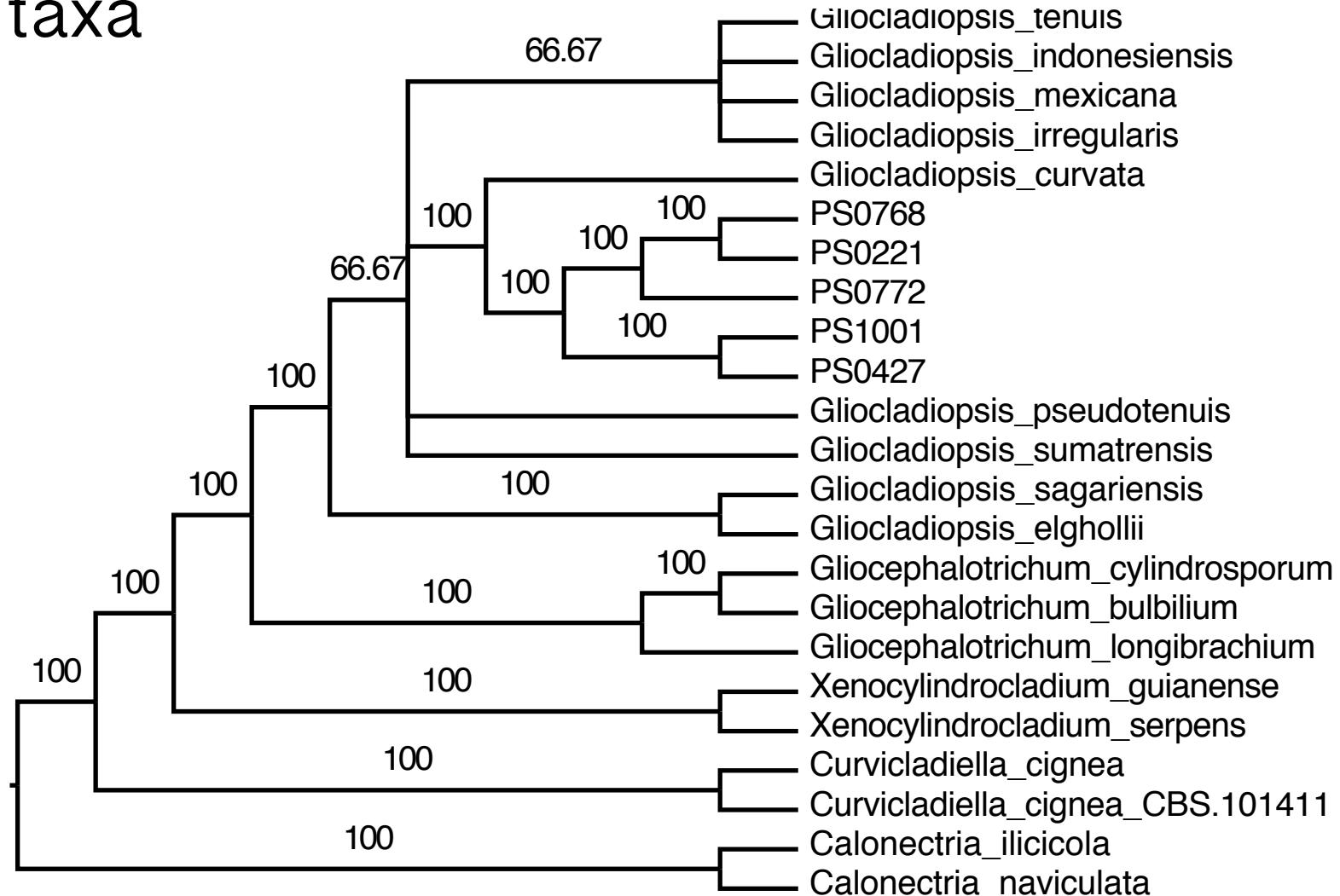
23 taxa



supertree demo

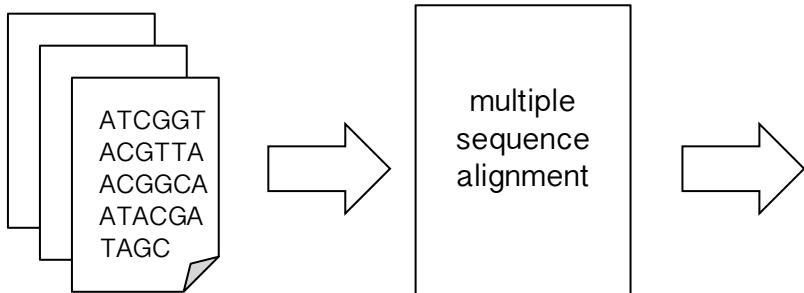
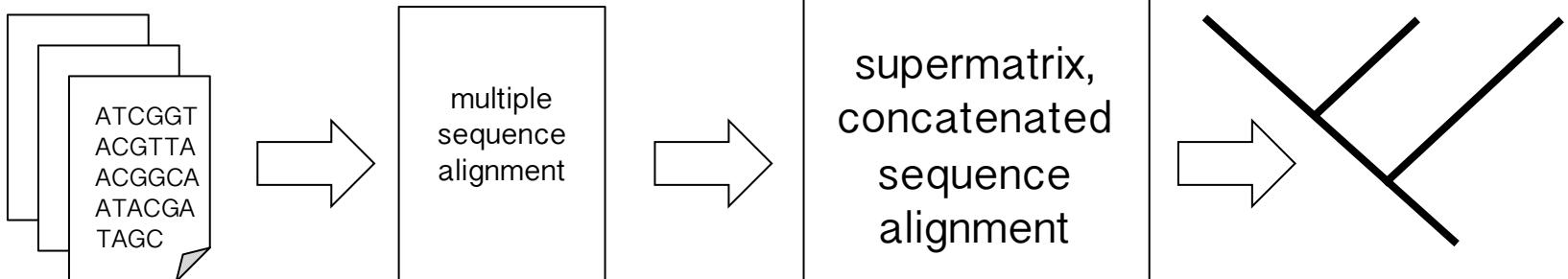
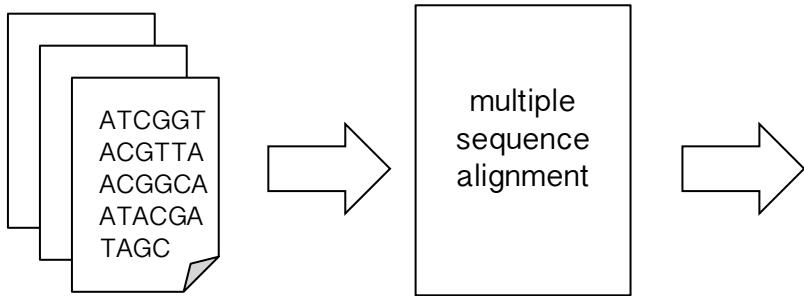
supertree

23 taxa



50% majority rule consensus of 21 equally parsimonious trees

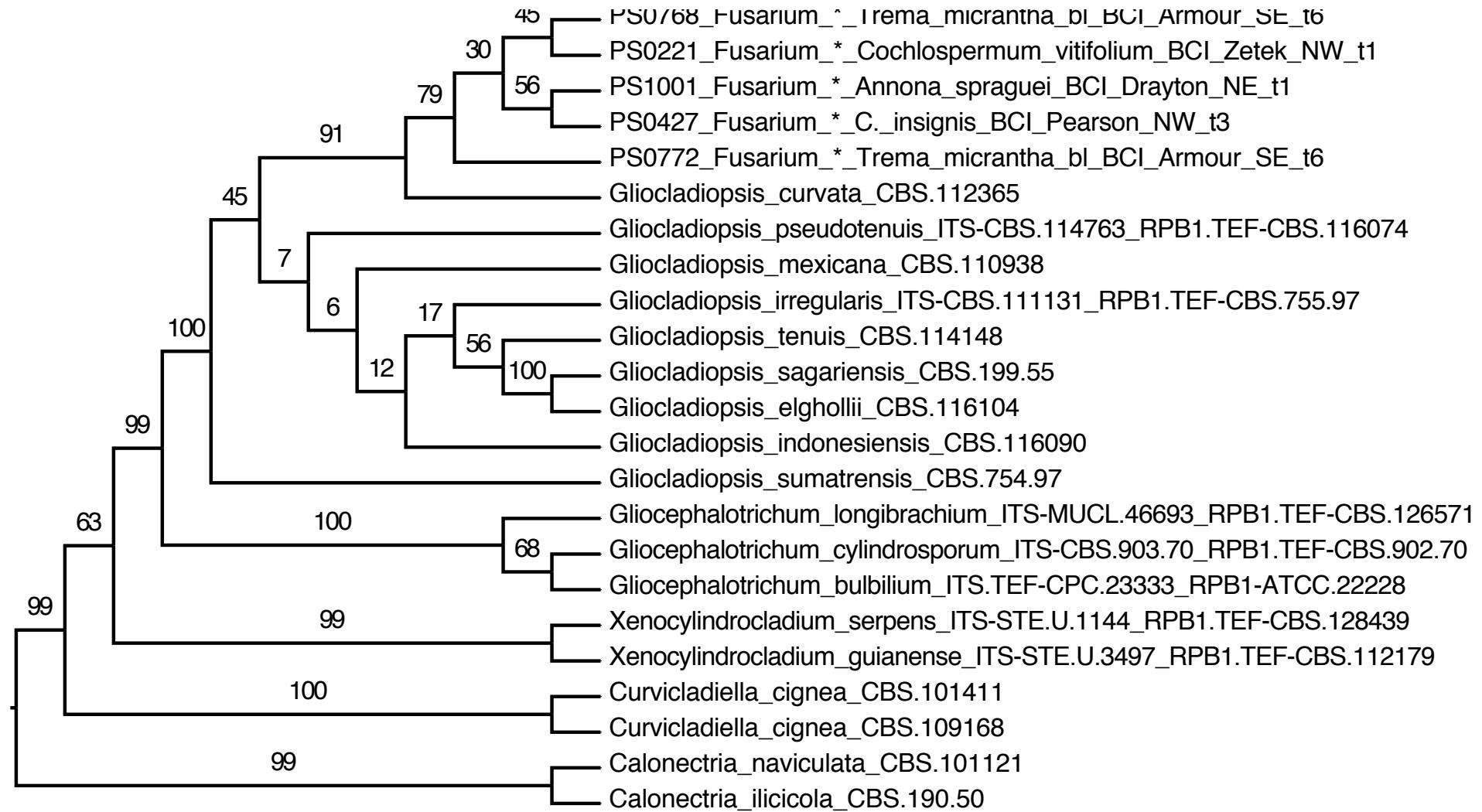
updated workflow



uses
conventional
algorithms

supermatrix

23 taxa, missing data

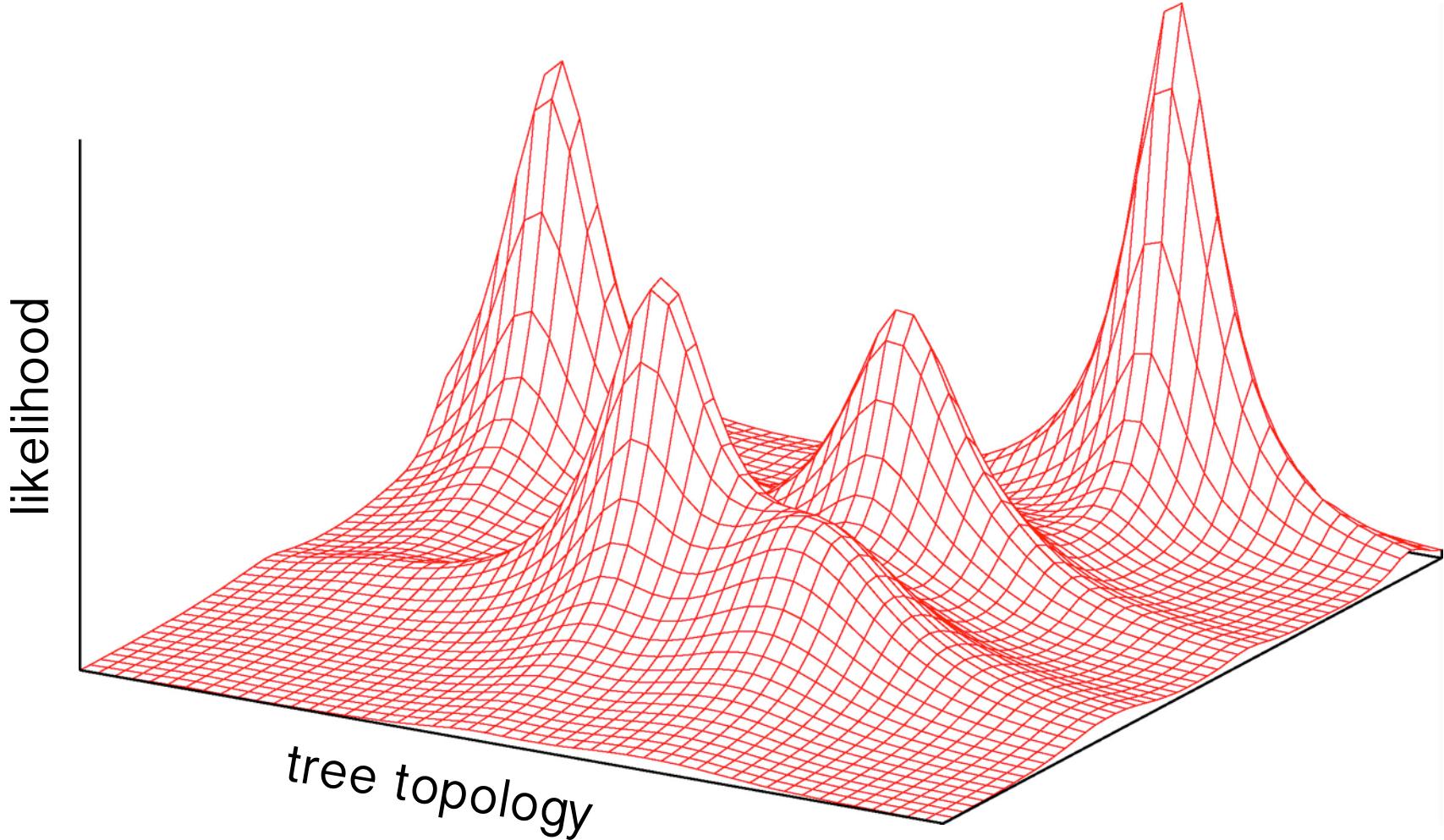


ML tree, reduced support

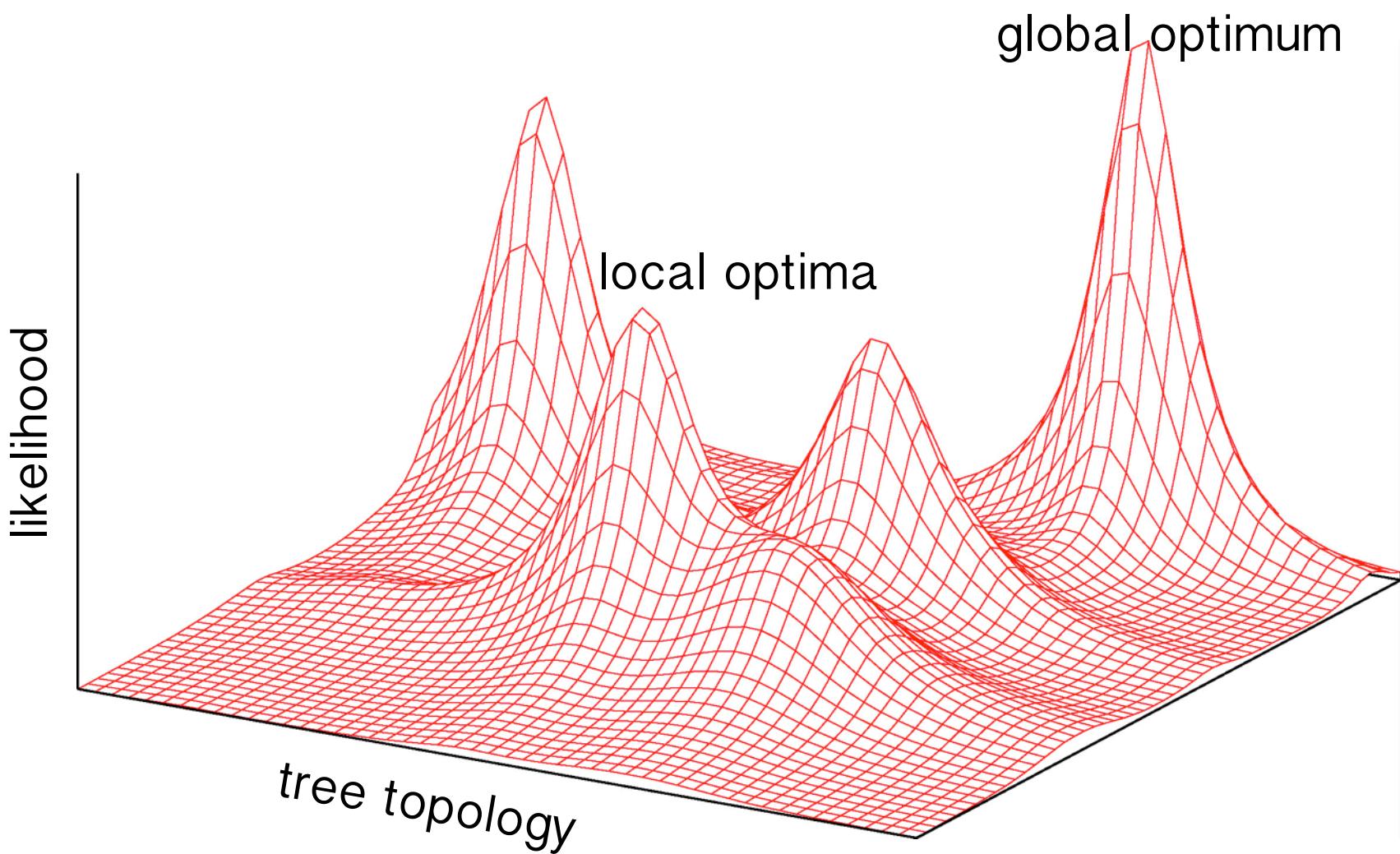
sidebar: terraces in treespace

the number of rooted
and unrooted trees for
 n taxa

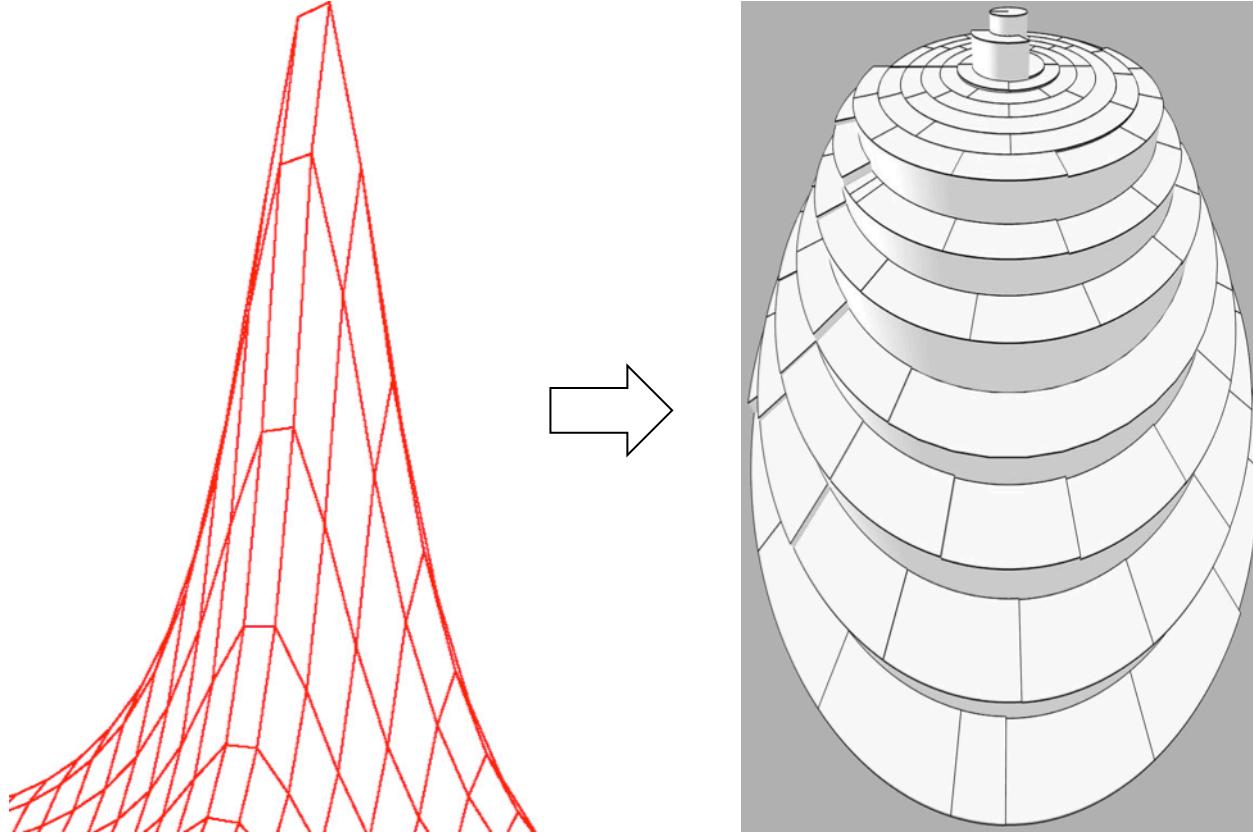
- 3, 1, 3
- 4, 3, 15
- 5, 15, 105
- 6, 105, 945
- 7, 945, 10395
- 8, 10395, 135135
- 9, 135135, 2027025
- 10, 2027025, 34459425
- 11, 34459425, 654729075
- 12, 654729075, 13749310575
- 13, 13749310575, 316234143225
- 14, 316234143225, 7905853580625
- 15, 7905853580625, 213458046676875
- 16, 213458046676875, 6190283353629375
- 17, 6190283353629375, 191898783962510625
- 18, 191898783962510625, 6332659870762850625
- 19, 6332659870762850625, 221643095476699771875
- 20, 221643095476699771875, 8200794532637891559375
- 21, 8200794532637891559375, 319830986772877770815625
- 22, 319830986772877770815625, 13113070457687988603440625
- 23, 13113070457687988603440625, 563862029680583509947946875
- 24, 563862029680583509947946875, 25373791335626257947657609375
- 25, 25373791335626257947657609375, 1192568192774434123539907640625



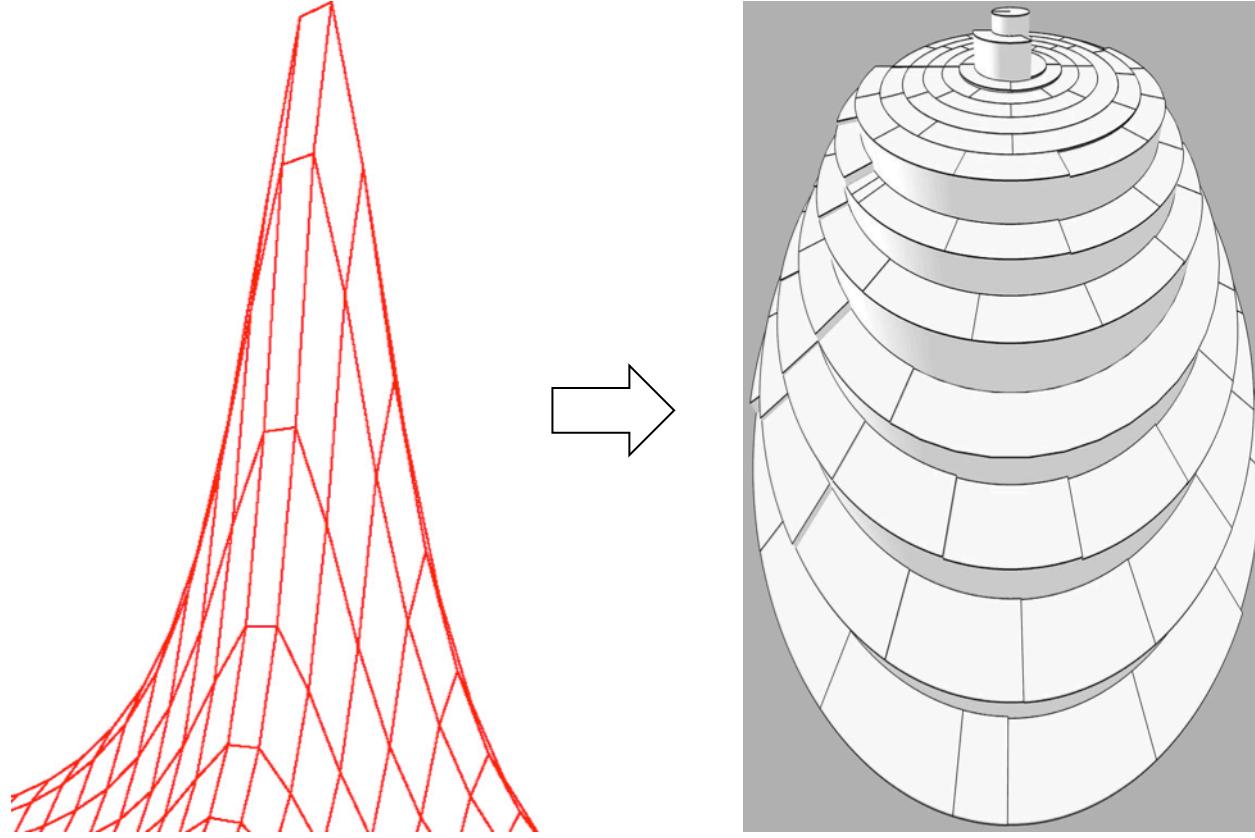
- each point represents a tree
- nearby trees are similar in topology
- each tree has a unique score (unique optimal parameters)



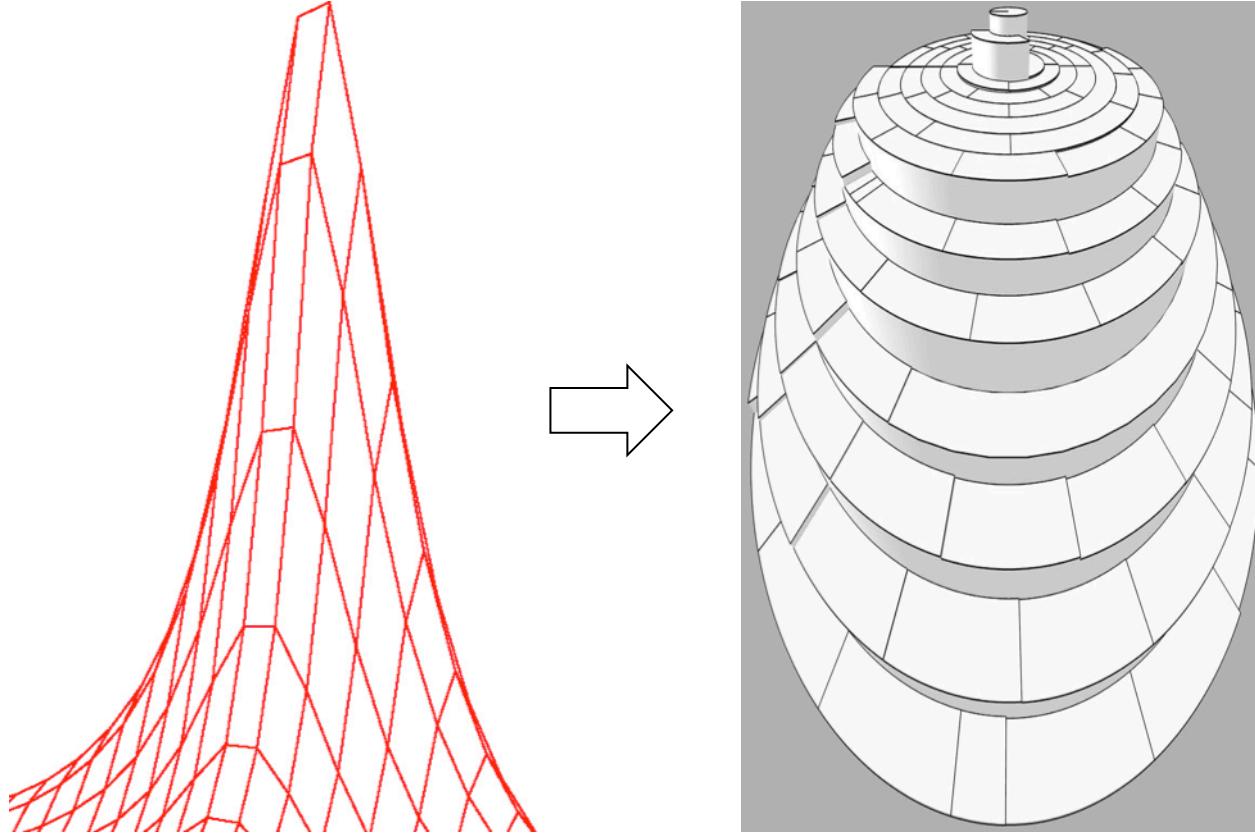
- each point represents a tree
- nearby trees are similar in topology
- each tree has a unique score (unique optimal parameters)



- missing data in multilocus studies creates terraces in the tree space



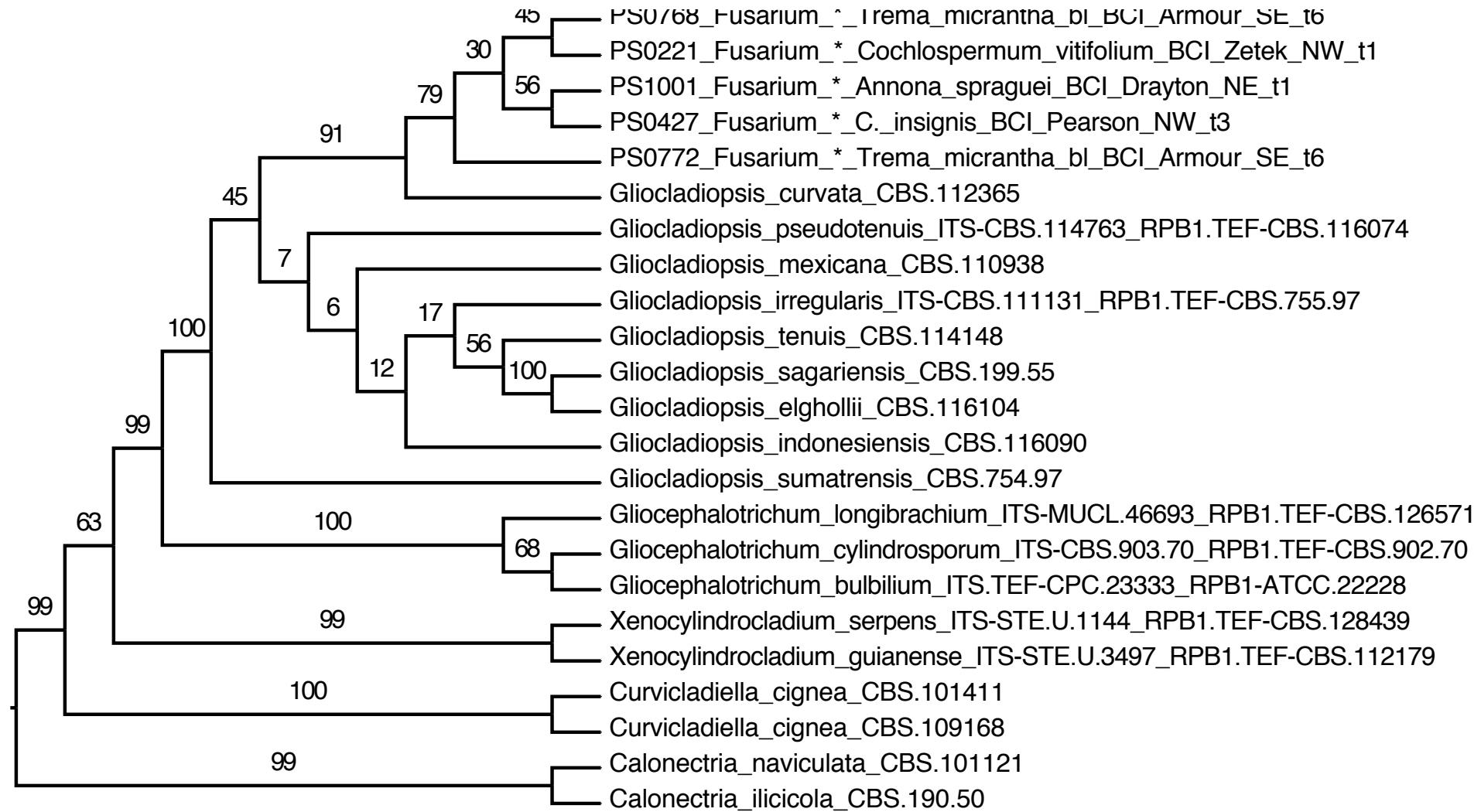
- missing data in multilocus studies creates terraces in the tree space
- each terrace contains many trees with the same likelihood score (e.g., 10^{36} trees for 1,500 taxa, 3 loci, ~40% missing data)



- more work needed to fully understand implications of terraces (Dr. Mike Sanderson, EEB)
- for now, try to use complete datasets

supermatrix

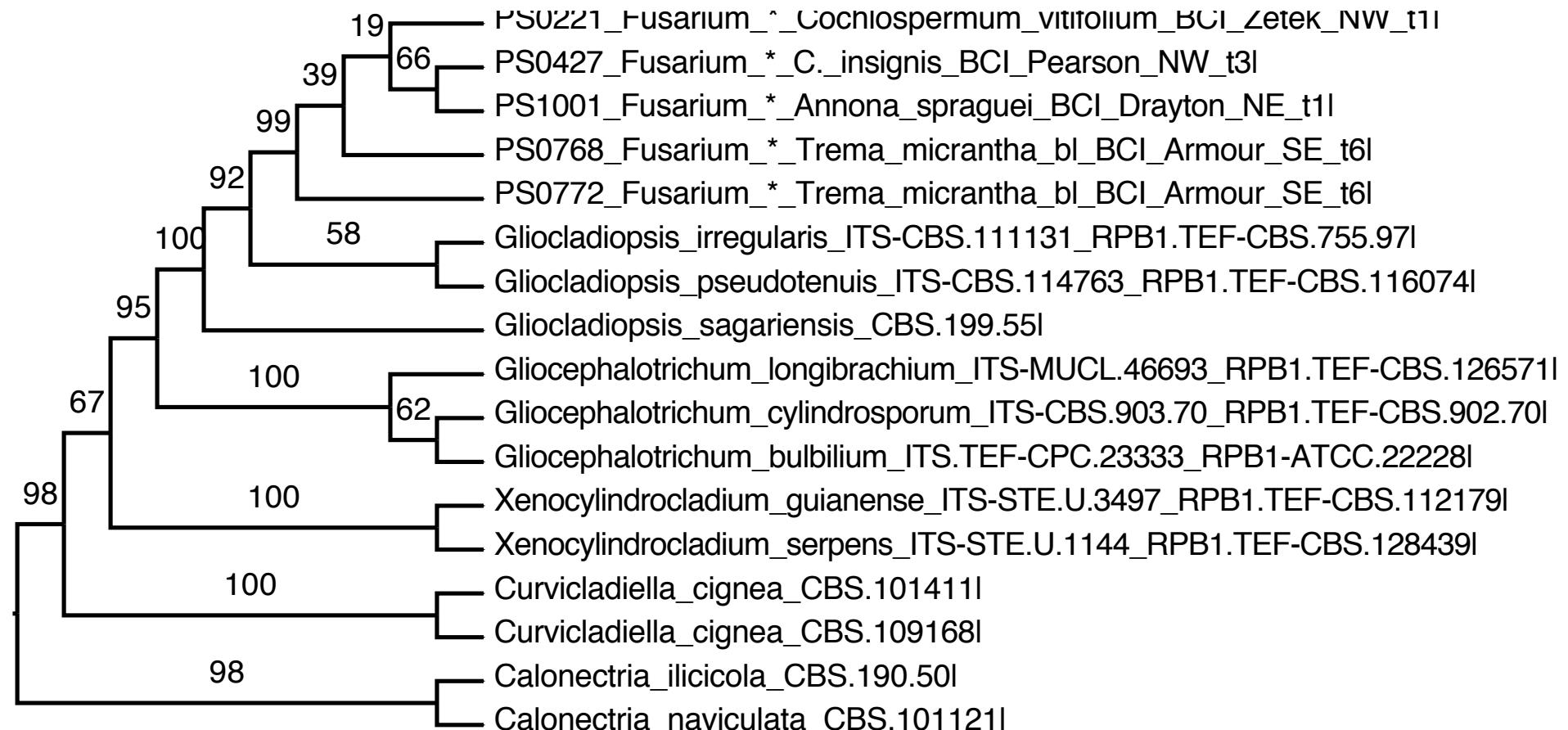
23 taxa, missing data



supermatrix demo

supermatrix

17 taxa, no missing data



analysis of phylogenies

- phylogenetic diversity

analysis of phylogenies

- phylogenetic diversity
- phylogenetic signal

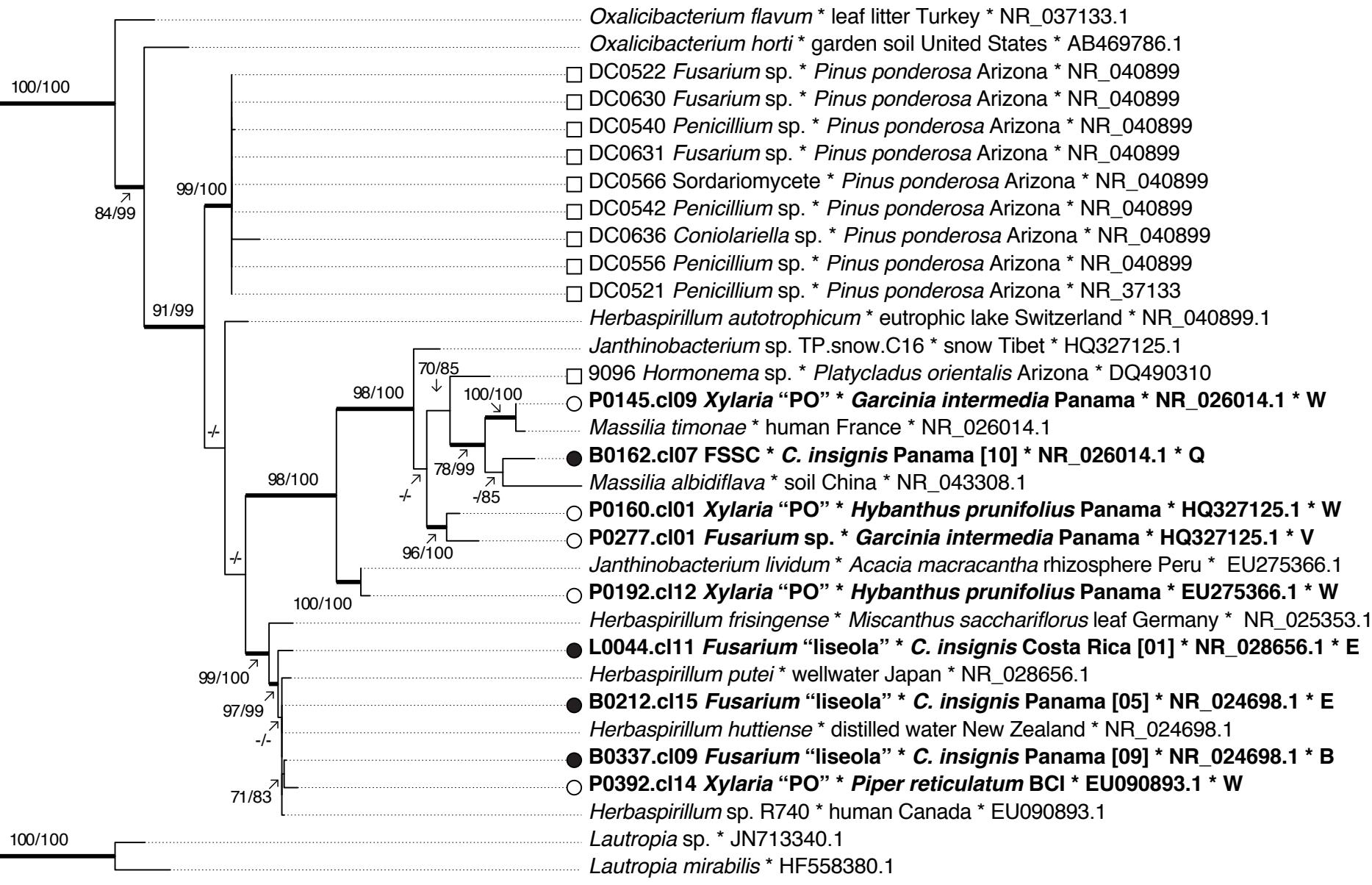
analysis of phylogenies

- phylogenetic diversity
- phylogenetic signal
- phylogenetic beta-diversity

analysis of phylogenies

- phylogenetic diversity
 - phylogenetic signal
 - phylogenetic beta-diversity
-
- ancestral state reconstruction (Mesquite)
 - molecular clock calibration (BEAST)

Oxalobacteriaceae (Betaproteobacteria)



analysis of phylogenies demo