Results

Novel markers for identification of C. botulinum strains

RE -TaqI

Variability in the fragments observed in the *rrs* sequences with RE TaqI allowed us to segregate the 13 strains into four groups. The first group was composed of strains Alaska E43 and Eklund 17B with an RE pattern of 88-585-198-55; 88-585-199 nts being common to them (Table S4). The two strains can be distinguished on the basis of RE pattern of 18-70-585-198-55 nts recorded only with Eklund 17B. The second and third groups composed of BKT015925 and ATCC 3502, respectively, had unique RE digestion patterns of 879-582 and 90-782 nts, respectively (Table S4). The fourth group composed of the rest of 9 strains were indistinguishable and had a single clear cut fragment of 872 nts.

RE-Hpy188III

RE-HPY188III could provide evidence for clear cut segregation of 13 *C. botulinum* strains into 4 groups. The digestion patterns of Alaska E43 and Eklund 17B were similar to each other, 502-96-88-261-76-274-86 nts, which formed a group independent of the rest of the strains (Table S5). The second group consisted of only BKT015925, which had a unique RE digestion pattern of 57-65-183-196-97-88-265-82-278 nts. The *C. botulinum* strain ATCC 3502 showed three different RE – Hpy188III digestion patterns. Of the three digestion patterns, two were unique to ATCC 3502: (i) 196-96-88-337-274 and (ii) 249-196-96-88-337-275, whereas the third RE pattern, 57-249-196-96-88-337-274 nts was shared by a large number of strains (Table S5). The fourth group with unique RE digestion pattern was consisting of ATCC 19397, which had one unique pattern 308-196-96-88-337-274, but shared its second RE digestion pattern with rest of the strains (Table S5).

RE-Tsp509I

As in the previous cases detected above, all the *rrs* copies of Alaska E43 (22/22) and Eklund 17B (22/22) showed a RE digestion pattern of similar kind, such that these two strains formed a group independent of the rest of the strains. The *C. botulinum* strain BKT015925 could be distinguished from the rest of strains on the basis of its unique RE digestion pattern with Tsp509I: 581-67-287-46 nts (Table S6). It thus formed second group. In the third group, we could place Kyoto with unique RE digestion patterns. It showed high variability in the RE digestion, where the unique RE patterns: (i) 461-244-40 and (ii) 183-461-244-40 accounted for 2 of the 18 copies of *rrs*. *C. botulinum* strains Kyoto share the rest four RE digestion patterns with other *C. botulinum* strains with varying frequencies (Table S6).

RE-HpyCH4III

Digestion of *rrs* of *C. botulinum* with RE HpyCH4III, allowed us to segregate them into 4 groups. As observed with RE-TaqI, Alaska E43 and Eklund 17B formed a group independent of the rest of the strains. However, they could be distinguished from each other on the basis of a 370 nts fragment, which was found only in 8/22 *rrs* copies of Eklund 17B (Table S4). The second group consisted of only HO4402065, with a unique RE pattern of 128–261–518-184 nts. The third clear cut groups was formed by BKT015525, where all of the 10 copies of *rrs* had a unique RE-HpyCH4III digestion pattern – 126-261-151-565 nts (Table S4). RE-HpyCH4III was not helpful in distinguishing rest of the *C. botulinum* strains, as they shared their digestion pattern.

RE - Hpy188I

With RE-Hpy188I, unique digestion pattern was recorded with BKT015925: 153-716 (nts). *C. botulinum* strains Eklund 17B, ATCC 3502, ATCC 19397, Hall had two RE- Hpy188I digestion patterns each (Table S5). However, in each of these cases, they share only one of their RE patterns with each other, which permits their easy segregation. ATCC 3502 appeared to be a link between ATCC 19397 and Eklund 17B, since it shares one of its RE patterns with each one of them. Similarly Eklund 17B appears to be a link between ATCC 3502 and Alaska E43, on similar grounds.

RE -NlaIV

A unique RE – NlaIV digestion Pattern was recorded with BKT015925. Kyoto had three types of RE patterns, however, they donot match exactly with those recorded with any other strains. The RE digestion patterns of Kyoto are shared by 230613 on one hand and other strains except Alaska E43 and Eklund 17B, on the other (Table S6).

Other REs

HPYCH4V was useful for distinguishing Kyoto in a clear cut manner (Table S5). Tsp45I proved helpful in segregating Alaska E43 and Eklund 17B strains (Table S4). Since it could digest *rrs* sequences of only these two strains. BstUI did not prove very helpful in distinguishing *C. botilinum* strains (Table S5).

Nucleotide changes

Loss of RE sites

A few cases of loss of RE sites due to NCs were observed due to transitions in *rrs* at different positions: 40, 67, 68, 74, 95, 122, 129, 211, 218, 223, 246, 250, 254, 257, 261, 408, 435, 440, 442, 974, 981, 983, 989, 993, 997, 1000, 1004, 1006, 1420 and 1427 (Table S8 and Fig. 2). It reflects on the susceptibility of different regions of *rrs* to spontaneous mutations. Here, transitions led to the loss of RE sites: (i) G>A led to the loss of RE sites for (1) AciI-(C'C**G**C), (2) AluI-(A**G**'CT), (3) BspMI-(ACCT**G**Cnnnn'nnnn), (4) CviJI-(r**G**'Cy), (5) HindIII-(A'A**G**CTT), and (6) MaeIII-('**G**TnAC); (ii) A>G resulted in the loss of RE sites for (1) AleI–(C**A**Cnn'nnGTG), (2) BfaI–(C'T**A**G), (3) MaeIII–('GTn**A**C), (4) MslI–(C**A**ynn'nnrTG) and (5) Tsp45I–('GTs**A**C); (iii) T>C transition resulted in the loss of (1) AleI–(CACnn'nnG**T**G), (2) MaeIII–('G**T**nAC), (3) MslI–(CAynn'nnr**T**G), and (4) Tsp45I–('G**T**sAC); and (iv) C>T transition lead to the loss of (1) MaeIII–('G**T**nAC) and (2) XmnI–(GAAnn'nnTT**C**). A perusal of all the transitions reveals that (i) all of them led to the loss of MaeIII–('**G**TnAC) site, (ii) A>G and/or T>C resulted in loss of sites for 3 REs (1) AleI, (2) MslI and (3) Tsp45I and (iii) G>A transitions resulted in loss of sites for multiple REs. Among these cases there was two transversion of C>A at position 983/989 in one *rrs* copy each of Alaska E43 and Eklund 17B, which led to the loss of TaqI–(T'**C**GA) site.

The affect of these losses in RE sites were most evident among 5-10 *rrs* copies of *C. botulinum* strains ATCC 3502, ATCC 19397 and Eklund 17B, followed by those in 2-4 *rrs* copies each of the rest of the strains (Table S8). Loch Maree was the only strain where no loss of RE sites was recorded.

Cladistic analysis and RE digestion pattern

A perusal of Table S10 reveals that it is possible to establish a close relationship among the 13 *C. botulinum* strains, on the basis of systematic absence or presence of certain HKGs in them. On the basis of those 11 genes which are not present in all the *C. botulinum* strains, it can be concluded that (i) ATCC3502 is unique, since it is the only strain which contains all these 11 genes, (ii) Kyoto, Loch Maree, Okra, and 657 form a second group with similarity to each other, with respect to the presence of *abrB, cphB, dltD, sfsA,* and *spmB* and absence of *arcA1, ccpA, dltB, mdeA, psd1*, and *psd2*, (iii) Langeland could be separated from group (ii) by the absence of gene *spmB*, (iv) ATCC19397 and Hall could be separated from group (ii) by the absence of gene *dltD* and from group (iii) due to the absence of *dltD* and presence of *spmB*, (v) Alska E43and Eklund 17B formed a very distinct group as they (a) don't contain genes *abrB* which were present in the rest of the 8 strains except H04402 065, BKT015925 and 230613 (b) but do contain gene *dltB* which was absent in rest of the strains except ATCC3502, Alaska E43 and Eklund 17B. These 2 strains can be distinguished from each other on the basis of presence of gene *sfsA* only in Alaska E43.

A critical analysis of cladistic tree drawn on the basis of concatenated 7 HKGs of 13 strains, revealed the following relationships: (i) Alaska E43 with Eklund 17B, (ii) ATCC19397, Hall, ATCC 3502 and H0442065 (iii) Okra with Kyoto (iv) 230613 with Langeland (Fig. 1). In addition, strain Loch Maree and 657 is intermediate to groups (i) (iii) and (iv), where as BKT015925 is independent. A comparison of cladistic analysis with RE digestion patterns of the *rrs* gene reveals that (i) Alaska E43 and Eklund 17B and (ii) BKT015925 formed two distinct groups (Tables S4-S6).

Table **S1**

Characteristics of sequenced genomes of *Clostridium botulinum* strains [a]

Clostridium botulinum genome					*rrs*[b]			
Strain (Accession no.)	Origin	Size (Mbp)[c]	G+C %mol	RDP Accession no.	G+C %mol	Intragenomic heterogeneity[d] (%)		Copies[e]
						rrs (NS)	Pairwise	copies[e]
Alaska E43, CP001078	Alaska, US[a]	3.66	27.4	S001094751, S001094753, S001094755, S001094758, S001094760, S001094762, S001094764, S001094766, S001094768, S001094770, S001094772, S002287743, S002288016, S002288105, S002288383, S002289948, S002290040, S002290286, S002290445, S002290927, S002291044, S002291294	51.52 - 52.09	0.39	0.06 - 0.33	22
Eklund 17B, CP001056	US[a]	3.85	27.5	S001094703, S001094705, S001094707, S001094710, S001094712, S001094714, S001094716, S001094718, S001094720, S001094722, S001094724, S002287177, S002287243, S002288156, S002288548, S002289273, S002289462, S002290154, S002290241, S002290247, S002290368, S002290659	51.52- 52.09	0.53	0.06 - 0.46	22
ATCC 3502, AM412317	Unknown, US[a]	3.9	28.2	S000858476, S000858478, S000858480, S000858482, S000858484, S000858486, S000858488, S000858491, S000858493, S002287460, S002287580, S002288802, S002288844, S002289360, S002289503, S002290042, S002290179, S002290425	52.05 - 52.60	0.97	0.06 - 0.55	18
Okra, CP000939	US[i]	4.11	28.2	S001014392, S001014394, S001014396, S001014398, S001014400, S001014402, S001014404, S001014407, S001014409[j], S002287372, S002287425, S002288021, S002288411, S002288499, S002289495, S002289731, S002290009, S002290479	51.92 - 52.53	0.46	0.06 - 0.33	18
Langeland, CP000728	Denmark[a]	4.01	28.3	S000891606, S000891608, S000891610, S000891612, S000891614, S000891616, S000891618, S000891621, S000891623, S002287184, S002287265, S002287727, S002288829, S002288840, S002289298, S002289587, S002290679, S002290804	52.05 - 52.60	0.53	0.06 - 0.39	18
657, CP001083	Texas, US[f]	4.26	28.0	S001416080, S001416082, S001416084, S001416086, S001416088, S001416090, S001416092, S001416095, S001416097, S002287409, S002287812, S002289183, S002289899, S002289967, S002290081, S002290186, S002291184, S002291313	51.98 - 52.53	0.59	0.13 - 0.39	18
Kyoto, CP001581	Japan[g]	4.16	28.2	S001350513, S001350515, S001350516, S001350517, S001350518, S001350519, S001350521, S001350522, S001350523, S002288023, S002288706, S002288916, S002288993, S002289181, S002289518, S002289629, S002290216, S002290406	52.05 - 52.60	0.59	0.06 - 0.46	18
Loch Maree, CP000962	Scotland[h]	4.26	28.1	S001014499, S002290134, S001014494, S002287648, S001014496, S002291211, S001014484, S001014492, S002290255, S002289199, S001014486, S001014488, S001014490, S002287052, S002287399, S002288157, S001014501, S002290692,	52.18 - 52.73	0.59	0.06 - 0.46	18
Hall, CP000727	Unknown[k]	3.76	28.2	S000891594, S002290424, S000891592, S002287900, S000891590, S002288659, S000891605, S000891598, S002290087, S002289150, S000891603, S002291106, S000891600, S002290636, S000891596, S002290871,	52.05 - 52.60	0.79	0.06 - 0.53	16
ATCC 19397, CP000726	Unknown[k]	3.86	28.2	S000891578, S002287857, S000891576, S002287510, S000891574, S002291146, S000891589, S000891582, S002290387,	52.05 - 52.60	0.79	0.06 - 0.53	16

				S002290229, S000891587, S002289359, S000891584, S002290093, S000891580, S002288115,				
BKT015925, CP002410	Sweden[l]	3.21	28.2	S002441238, S002441241, S002441243, S002441245, S002441247, S002441249, S002441251, S002441253, S002441255, S002441257	52.12 - 52.45	0.59	0.06- 0.39	10
H04402 065, FR773526	United Kingdom[m]	3.92	28.2	S002408083, S002408085, S002408087, S002408089, S002408091, S002408093, S002408095, S002408098, S002408100	52.26 - 52.53	0.93	0.06- 0.39	9
230613, CP002011	China[a]	4.01	28.3	S002165509, S002165511, S002165507, S002165513, S002165517, S002165519, S002165515, S002165522, S002165524	52.33 - 52.53	0.26	0.06- 0.26	9
							Total	212

[a] http://www.ncbi.nlm.nih.gov/genomes/MICROBES/microbial_taxtree.html

[b] http://rdp.cme.msu.edu/genome/

[c] http://tools.neb.com/~vincze/genomes/

[d] Percent variation range calculated on the basis of pair-wise comparison of *rrs* of a genome

[e] http://ribosome.mmg.msu.edu/rrndb/search.php

[f] Isolated from an infant botulism case in Texas **(9).**

[g] Associated with an infant botulism case in Japan **(9).**

[h] Isolated from a food poisoning incident in Scotland **(9).**

[i] Most likely originated from food borne botulism incident in the US **(9).**

[j] Reference *rrs* used for detecting nucleotide changes in the rest of the *rrs* sequences.

[k] Most likely originated from food borne botulism cases in the western US **(9).**

[l] Isolated from an poultry broilers case in Sweden **(34)**

[m] Isolated from an infant botulism case in United Kingdom. **(10)**

Table **S2**

Intra- and intergenomic heterogeneity among *rrs* of *Clostridium botulinum* strains

RDP Accession no.	**Similarity of *rrs* to other genomes**	
	RDP Accession no. and strain	**Heterogeneity**
***Clostridium botulinum* ATCC 19397 (CP000726)**[a]		
S000891574	S000891590 **Hall (CP000727)**	Intergenomic
S000891576	S000891592 **Hall (CP000727)**	Intergenomic
S000891578	S000891594 **Hall (CP000727)**	Intergenomic
S000891580	S000891596 **Hall (CP000727)**	Intergenomic
S000891582	S000891598 **Hall (CP000727)**	Intergenomic
S000891584	S000891600 **Hall (CP000727)**	Intergenomic
S000891587	S000891603 **Hall (CP000727)**	Intergenomic
S000891589	S000891605 **Hall (CP000727)**	Intergenomic
S002287857	S002290424 **Hall (CP000727)**	Intergenomic
S002290093	S002290636 **Hall (CP000727)**	Intergenomic
S002287510	S002287900 **Hall (CP000727)** S002289503 **ATCC 3502 (AM412317)**	Intergenomic
S002288115	S002290871 **Hall (CP000727)** S002287580 **ATCC 3502 (AM412317)**	Intergenomic
S002289359	S002291106 **Hall (CP000727)** S002287460 **ATCC 3502 (AM412317)**	Intergenomic
S002290387	S002290087 **Hall (CP000727)** S002288844 **ATCC 3502 (AM412317)**	Intergenomic
S002291146	S002288659 **Hall (CP000727)** S002290425 **ATCC 3502 (AM412317)**	Intergenomic
S002288802	S002289150 **Hall (CP000727)** S002289360 **ATCC 3502 (AM412317)**	Intergenomic
	S002290229 ATCC 19397	Intragenomic
***C. botulinum* 657 (CP001083)**		
S001416080	S001416084, S001416086	Intragenomic
S001416082	S001416090	Intragenomic
S002290081	S002290186	Intragenomic
S002289899	S002291184, S002291313	Intragenomic
***C. botulinum* Loch Maree (CP000962)**		
S001014484	S001014492	Intragenomic
S001014486	S001014488, S001014490	Intragenomic
S002287052	S002287399, S002288157	Intragenomic
S002289199	S002290255	Intragenomic
***C. botulinum* Alaska E43 (CP001078)**		
S001094751	S001094764, S001094768	Intragenomic
S001094753	S001094760	Intragenomic
S001094766	S001094772	Intragenomic
S002291044	S002291294	Intragenomic
S002288016	S002289948	Intragenomic
	S002290241 **Eklund 17B (CP001056)**	Intergenomic
S002288383	S002290286, S002290445	Intragenomic
S002287743	S002289273 S002287243 **Eklund 17B (CP001056)**	Intergenomic
***C. botulinum* Eklund 17B (CP001056)**		
S001094710	S001094724	Intragenomic
S001094714	S001094718	Intragenomic
S001094716	S001094722	Intragenomic
S002288706	S002289181, S002289518	Intragenomic
S002287177	S002288548	Intragenomic
S002290247	S002290368	Intragenomic
S001350523	S002408083, S002408091 **H04402 065 (FR773526)**	Intergenomic
S002290241	S002288016, S002289948 **Alaska E43 (CP001078)**	Interegenomic
S002287243	S002289273	Intragenomic
	S002287743 **Alaska E43 (CP001078)**	Interegenomic
***C. botulinum* Kyoto (CP001581)**		
S001350516	S001350518, S001350522	Intragenomic
S001350517	S001014409, S001014402, S001014398 **Okra (CP000939)**	Intergenomic
	S000891612 S000891618 **Langeland (CP000728)**	Intergenomic
	S002408100 **H04402 065 (FR773526)**	Intergenomic
S002288916	S002289495, S002289731, S002290479 **Okra (CP000939)**	Intergenomic
	S002287265, S002289587 **Langeland (CP000728)**	Intergenomic
	S002290179 **ATCC 3502 (AM412317)**	Intergenomic
***C. botulinum* Okra (CP000939)**		
S001014396	S000891606 **Langeland (CP000728)**	Intergenomic
S001014409	S001014402, S001014398	Intragenomic
	S001350517 **Kyoto (CP001581)**	Intergenomic

	S000891612, S000891618 **Langeland (CP000728)**	Intergenomic
	S002408100 **H04402 065 (FR773526)**	Intergenomic
S002289495	S002289731, S002290479	Intragenomic
	S002288916 **Kyoto (CP001581)**	Intergenomic
	S002287265, S002289587 **Langeland (CP000728)**	Intergenomic
	S002290179 **ATCC 3502 (AM412317)**	Intergenomic
S002287372	S002290804 **Langeland (CP000728)**	Intergenomic
***C. botulinum* Langeland (CP000728)**		
S000891606	S001014396 **Okra (CP000939)**	Intergenomic
S000891621	S000891623	Intragenomic
S002290804	S002287372 **Okra (CP000939)**	Intergenomic
S002287265	S002289587	Intragenomic
	S002288916 **Kyoto (CP001581)**	Intergenomic
	S002289495, S002289731, S002290479 **Okra (CP000939)**	Intergenomic
	S002290179 **ATCC 3502 (AM412317)**	Intergenomic
S002288829	S002288840	Intragenomic
S000891612	S000891618	Intragenomic
	S001014398 S001014402 S001014409 **Okra (CP000939)**	Intergenomic
	S001350517 **Kyoto (CP001581)**	Intergenomic
	S002408100	Intragenomic
***C. botulinum* ATCC 3502 (AM412317)**		
S000858484	S000858493	Intragenomic
S002289503	S002287510 **ATCC 19397 (CP000726)**	Intergenomic
	S002287900 **Hall (CP000727)**	Intergenomic
S002290425	S002291146 **ATCC 19397 (CP000726)**	Intergenomic
	S002288659 **Hall (CP000727)**	Intergenomic
S002288844	S002290387 **ATCC 19397 (CP000726)**	Intergenomic
	S002290087 **Hall (CP000727)**	Intergenomic
S002287580	S002288115 **ATCC 19397 (CP000726)**	Intergenomic
	S002290871 **Hall (CP000727)**	Intergenomic
S002287460	S002289359 **ATCC 19397 (CP000726)**	Intergenomic
	S002291106 **Hall (CP000727)**	Intergenomic
S002289360	S002288802	Intragenomic
	S002290229 **ATCC 19397 (CP000726)**	Intergenomic
	S002289150 **Hall (CP000727)**	Intergenomic
S002290179	S002289495, S002289731, S002290479 **Okra (CP000939)**	Intergenomic
	S002287265 S002289587 **Langeland (CP000728)**	Intergenomic
	S002288916 **Kyoto (CP001581)**	Intergenomic
S002288802	S002289360	Intragenomic
***C. botulinum* 230613 (CP002011)**		
S002165507	S002165509, S002165511, S002165513, S002165517, S002165519	Intragenomic
S002165522	S002165524	Intragenomic
***C. botulinum* BKT015925 (CP002410)**		
S002441238	S002441249, S002441257	Intragenomic
***C. botulinum* H04402 065 (FR773526)**		
S002408083	S002408091	Intragenomic
	S001350523 **Kyoto (CP001581)**	Intergenomic
S002408100	S001014398, S001014402, S001014409 **Okra (CP000939)**	Intergenomic
	S001350517 **Kyoto (CP001581)**	Intergenomic
	S000891612, S000891618 **Langeland (CP000728)**	Intergenomic

[a] All the 16 *rrs* of *C. botulinum* ATCC 19397 CP000726) and Hall (CP000727) were similar with respect to their nucleotide sequences.

Table **S3**
List of Restriction Endonucleases used in this study

Enzyme with generate recognition		
Cutters	RE with recognition site	Total
4 base cutters	AciI(C'CGC), AluI(AG'CT), BfaI(C'TAG), BstKTI(GAT'C), BstUI(CG'CG), Csp6I(G'TAC), DpnI(GA'TC), FatI('CATG), HaeIII(GG'CC), HhaI(GCG'C), HinPI('CGC), HpaII(C'CGG), HpyCH4IV(A'CGT), HpyCH4V(TG'CA), MboI('GATC), MseI(T'TAA), NlaIII(CATG), RsaI(GT'AC), TaiI(ACGT'), TaqI(T'CGA), Tsp509I('AATT)	21
5 base cutters	NA	
6 base cutters	AatII(GACGT'C), ApaI(GGGCC'C), AseI(AT'TAAT), BspEI(T'CCGGA), BsrGI(T'GTACA), BssSI(C'ACGAG), BstBI(TT'CGAA), EagI(C'GGCCG), EcoRI(G'AATTC), EcoRV(GAT'ATC), FspI(TGC'GCA), HindIII(A'AGCTT), HpaI(GTT'AAC), NruI(TCG'CGA), PmlI(CAC'GTG), PsiI(TTA'TAA), PspOMI(G'GGCCC), SacII(CCGC'GG), SmaI(CCC'GGG), SnaBI(TAC'GTA), SspI(AAT'ATT), XbaI(T'CTAGA), XmaI(C'CCGGG), ZraI(GAC'GTC)	24
Large recognition site (>6 base cutters)	BbvCI(CC'TCAGC)	1
Enzyme with degenerate recognition		
4 base cutters	CviJI(rG'Cy)	1
5 base cutters	AflIII(A'CryGT), AvaII(G'GwCC), BstNI(CC'wGG), DdeI(C'TnAG), Fnu4HI(GC'nGC), HinfI(G'AnTC), Hpy188I(TCn'GA), HpyCH4III(ACn'GT), MaeIII('GTnAC), NciI(CC'sGG), PspGI('CCwGG), Sau96I(G'GnCC), ScrFI(CC'nGG), StyDI('CCnGG), TauI(GCsG'C), TfiI(G'AwTC), TseI(G'CwGC), Tsp45I('GTsAC)	18
6 base cutters	ApoI(r'AATTy), AvaI(C'yCGrG), BanII(GrGCy'C), Bme1580I(GkGCm'C), BsaAI(yAC'GTr), BsaHI(Gr'CGyC), BsaJI(C'CnnGG), BsaWI(w'CCGGw), BsiEI(CGry'CG), BsiHKAI(GwGCw'C), Bsp12876I(GdGCh'C), BsrI(ACTGGn'), BtgI(C'CryGG), CacI(GCn'NGC), EaeI(y'GGCCr), HincII(GTy'rAC), Hpy8I(GTn'nAC), Hpy188III(TC'nnGA), MspA1I(CmG'CkG), NlaIV(GGn'nCC), NspI(rCATG'y), TatI(w'GTACw)	22
Large recognition site (>6 base cutters)	AlwI(GGATCnnnn'n), BaeI(ACnnnnGTAyCnnnnnnnnnnnn'), BaeI(GrTACnnnnGTnnnnnnnnnnnnnnn'), BbsI(GAAGACnn'nnnn), BbvI(GCAGCnnnnnnnn'nnnn), BceAI(ACGGCnnnnnnnnnnnn'nn), BcgI(CGAnnnnnnTGCnnnnnnnnnnnn'), BcgI(GCAnnnnnnTCGnnnnnnnnnnnn'), BciVI(GTATCCnnnnnn'), BmrI(ACTGGGnnnnn'), BpmI(CTGGAGnnnnnnnnnnnnnnnn'), Bpu10I(CC'TnAGC), BseMII(CTCAGnnnnnnnnnn'), BseRI(GAGGAGnnnnnnnnnn'), BsgI(GTGCAGnnnnnnnnnnnnnnnn'), BslI(CCnnnnn'nnGG), BsmI(GAATGCn'), BsmAI(GTCTCn'nnnn), BspCNI(CTCAGnnnnnnnnn'), BspMI(ACCTGCnnnn'nnnn), BsrDI(GCAATGnn'), BstEII(G'GTnACC), BstF5I(GGATGnn'), BtsI(GCAGTGnn'), EciI(GGCGGAnnnnnnnnnnn'), Eco57MI(CTGrAGnnnnnnnnnnnnnnnn'), EcoNI(CCTnn'nnnAGG), EcoO109I(rG'GnCCy), FauI(CCCGCnnnn'nn), FokI(GGATGnnnnnnnnn'nnnn), HgaI(GACGCnnnnn'nnnnn), Hin4I(GAynnnnnvTCnnnnnnnnnnnnn'), Hin4I(GAbnnnnnrTCnnnnnnnnnnnnn'), HphI(GGTGAnnnnnnnn'), HpyF10VI(GCnnnnnn'nGC), MboII(GAAGAnnnnnnnn'), MlyI(GAGTCnnnnn'), MnlI(CCTCnnnnnnn'), MslI(CAynn'nnrTG), MwoI(GCnnnnn'nnGC), PleI(GAGTCnnnn'n), SfaNI(GCATCnnnnn'nnnn), StyI(C'CwwGG), TaqII(CACCCAnnnnnnnnnnn'), TspDTI(ATGAAnnnnnnnnnnn'), TspGWI(ACGGAnnnnnnnnnnn'), TspRI(nnCAsTGnn'), Tth111I(GACn'nnGTC), XcmI(CCAnnnnn'nnnnTGG), XmnI(GAAnn'nnTTC)	50
	Total	137

[a] RE action sites were observable with only 137/ 241 REs initially employed.

Table **S4**

In silico Restriction Enzyme activity in *rrs* gene of *Clostridium botulinum* strains: HpyCH4III, TaqI and Tsp45I

***Clostridium botulinum* strains**^{a}	**Restriction Enzyme digestion fragments (nucleotides, nts)**												
	HpyCH4III												
Alaska E43 (22/22), Eklund 17B (14/22)			■	261	■	149	■						
Eklund 17B (8/22)			■	261	■	149	■	370	■				
ATCC3502 (18/18), ATCC19397/Hall (16/16), Kyoto (18/18), Langeland (18/18), Loch Maree (18/18), Okra (18/18), 657 (18/18), 230613 (9/9), H04402 065 (8/9)	■	128	■	261	■	150	■	368	■	184		■	
H04402 065 (1/9)	■	128	■	261	■		518		■	184		■	
BKT015925 (10/10)	■	128	■	261	■	151	■	565				■	
	TaqI												
Alaska E43 (20/22), Eklund 17B (16/22)	■		88		■	585	■	198	■	55	■		
Alaska E43 (2/22), Eklund 17B (2/22)	■		88		■	585	■	199	■				
Eklund (4/22)	■	18	■	70	■	585	■	198	■	55	■		
ATCC3502 (16/18), ATCC19397/Hall (16/16), Kyoto (18/18), Langeland (18/18), Loch Maree (18/18), Okra (18/18), 657 (18/18), 230613 (9/9), H04402 065 (9/9)	■				872				■				
BKT015925 (10/10)	■				879				■	522			■
ATCC3502 (2/18)	■		90		■		782		■				
	Tsp45I												
Alaska E43 (10/22)	■	9	■	371		■							
Eklund 17B (20/22)		9		370	1								
Alaska E43 (12/22), Eklund 17B (2/22)			■	371		■							
ATCC3502 (18/18), ATCC19397/Hall (16/16), H04402 065 (9/9) Kyoto (18/18), Langeland (18/18), Loch Maree (18/18), Okra (18/18), 657 (18/18), 230613 (9/9), BKT015925 (10/10)	NA												

Table **S5**

In silico Restriction Enzyme activity in *rrs* gene of *Clostridium botulinum* strains: BstUI, Hpy188III and HpyCH4V

***Clostridium botulinum* strains**[a]	**Restriction Enzyme digestion fragments (nucleotides, nts)**
	BstUI
Alaska E43 (22/22), Eklund 17B (14/22), BKT015925 (8/10)	■ 166 ■ 100 ■ 444 ■ 200 ■ 94 ■ 86 ■ 15 ■
Eklund 17B (8/22), BKT015925 (2/10)	■ 166 ■ 100 ■ 444 ■ 200 ■ 180 ■ 15 ■
ATCC3502 (18/18), ATCC19397/Hall (16/16), Kyoto (18/18), Langeland (18/18), Loch Maree (18/18), Okra (18/18), 657 (18/18), H04402 065 (9/9), 230613 (9/9)	■ 166 ■ 100 ■ 445 ■ 379 ■ 15 ■
	Hpy188III
Alaska E43 (22/22), Eklund 17B (22/22)	■ 502 ■ 96 ■ 88 ■ 261 ■ 76 ■ 274 ■ 86 ■
ATCC19397/Hall (10/16)	■ 308 ■ 196 ■ 96 ■ 88 ■ 337 ■ 274 ■
ATCC19397/Hall (6/16), ATCC3502 (4/18), H04402 065 (9/9), Kyoto (18/18), Langeland (18/18), Loch Maree (18/18), Okra (18/18), 657 (18/18), 230613 (9/9)	■ 57 ■ 249 ■ 196 ■ 96 ■ 88 ■ 337 ■ 274 ■
BKT015925 (10/10)	■ 57 ■ 69 183 ■ 196 ■ 97 ■ 88 ■ 265 ■ 82 278 ■
ATCC3502 (9/18)	■ 249 ■ 196 ■ 96 ■ 88 ■ 337 ■ 275 ■
ATCC3502 (5/18)	■ 196 ■ 96 ■ 88 ■ 337 ■ 274 ■
	HpyCH4V
Alaska E43 (22/22), Eklund 17B (22/22)	■ 309 ■ 22 ■ 217 ■ 25 ■ 395 ■ 204 ■
ATCC3502 (2/18), ATCC19397/Hall (2/16), Langeland (2/18), Loch Maree (4/18), Okra (4/18), H04402 065 (3/9)	■ 548 ■ 25 ■ 346 ■ 49 ■ 46 ■ 157 ■ 58 ■
ATCC3502 (16/18), ATCC19397/Hall (14/16), Langeland (16/18), Loch Maree (14/18), Okra (14/18), 230613 (9/9), H04402 065 (6/9)	■ 548 ■ 25 ■ 346 ■ 49 ■ 46 ■ 215 ■
Kyoto (8/9), 657 (18/18)	■ 548 ■ 25 ■ 346 ■ 49 ■ 46 ■ 215 ■
Kyoto (2/18)	■ 133 ■ 415 ■ 25 ■ 346 ■ 49 ■ 46 ■ 215 ■
BKT015925 (10/10)	■ 552 ■ 25 ■ 405 ■

a: Values within parentheses are the frequency of *rrs* copies with the RE pattern out of the total number of copies in the strain.

Symbol (Filled square) indicates RE site in the *rrs* gene.

Table **S6**

In silico Restriction Enzyme activity in *rrs* gene of *Clostridium botulinum* strains: Tsp509I, Hpy188I and NlaIV

***Clostridium botulinum* strains**[a]	**Restriction Enzyme digestion fragments (nucleotides, nts)**																	
Tsp509I																		
Alaska E43 (22/22), Eklund 17B (22/22)											■		244	■	40	■	46	■
ATCC 3502 (9/18), ATCC 19397/Hall (8/16), Langeland (9/18), Loch Maree (9/18), Okra (9/18), 657 (7/18), Kyoto (7/18), 230613 (9/9), H04402 065 (9/9)											■		244	■	40			
ATCC 3502 (9/18), ATCC 19397/Hall (8/16), Langeland (9/18), Loch Maree (9/18), Okra (9/18), Kyoto (7/18), 657 (9/18)			■					644			■		244	■	40	■		
Kyoto (1/18), 657 (2/18), BKT015925 (1/10)			■	88	■			556			■		244	■	40	■		
Kyoto (1/8), 657 (2/18)					■			556			■		244	■	40			
Kyoto (1/8)							■	461			■		244	■	40	■		
Kyoto (1/8)			■	183			■	461			■		244	■	40	■		
BKT015925 (9/10)			■			581			■	67	■		287			■	46	■
Hpy188I																		
Alaska E43 (22/22), Eklund 17B (18/22)	■			338			■					854				■	56	■
ATCC19397/Hall (2/16)	■	70	■	267			■					856				■	56	■
ATCC3502 (16/18), ATCC19397/Hall (14/16), Kyoto (18/18), Langeland (18/18), Loch Maree (18/18), Okra (18/18), 657 (18/18), 230613 (9/9), H04402 065 (9/9)			■	267			■					854				■	56	■
ATCC3502 (2/18), Eklund 17B (4/22)							■					854				■	56	■
BKT015925(10/10)							■	153					716			■		
NlaIV																		
ATCC19397/Hall (8/16), Langeland (14/18), ATCC 3502 (12/18),), H04402 065 (8/9)	■					672					■	1	■		524	■		
ATCC19397/Hall (8/16), Langeland (4/18), ATCC 3502 (6/18), H04402 065 (1/9)	■	229			■	443					■	1	■		524	■		
657 (18/18), 230613 (7/9), Kyoto (16/18), Okra (18/18)	■					672					■	1	■		524	■		
Alaska E43 (22/22), Eklund 17B (22/22), Loch Maree (18/18)	■	229			■	443					■	1	■					
Kyoto(2/18), 230613 (2/9)	■	229			■	443					■	1	■					
BKT015925(10/10)		230			■	356	■	14	■	76	■	1	■		536			■

a: Values within parentheses are the frequency of *rrs* copies with the RE pattern out of the total number of copies in the strain.

Symbol (Filled square) indicates RE site in the *rrs* gene.

Table S7.

Affect on restriction enzyme sites in *rrs* of sequenced genomes of different *Clostridium botulinum* strains due to single nucleotide polymorphism.

rrs C. botulnum	**Nucleotide substitution**	**Restriction Enzyme**			
		Name and sequence[a]	**No. of sites**		**Change**
			Pre - S	**Post - S**	
***Clostridium botulinum* Langeland (CP000728)**					
S000891614	67**G**>A	AluI - (A**G**'CT)	5	4	-1
	67**G**>A	HindIII – (A'A**G**CTT)	1	0	-1
	67**G**>A, 1420**G**>A	CviJI - (r**G**'Cy)	27	25	-2
	1420**G**>A	NlaIV - (GG**n**'nCC)	4	4	0
S000891616	1420G>A	CviJI - (r**G**'Cy)	27	26	-1
	1420**G**>A	NlaIV - (GG**n**'nCC)	4	4	0
S000891606	1420G>A	CviJI - (rG'Cy)	27	26	-1
	1420**G**>A	NlaIV - (GGn'nCC)	4	4	0
S000891610	1220 C>**T**	HpyCH4V – (**T**G'CA)	7	8	+1
S002287184	**1227C>T**	HpyCH4V – (**T**G'CA)	7	8	+1
S000891621	67**G**>A	AluI - (AG'CT)	5	4	-1
	67**G**>A	CviJI - (rG'Cy)	27	26	-1
	67**G**>A	HindIII – (A'AGCTT)	1	0	-1
	435**A**>G, 454**T**>C	BfaI – (C'**TA**G)	8	6	-2
	454T>**C**	Acc65I – (G'GTAC**C**)	0	1	+1, Novel
	454T>**C**	BanI – (G'GyrC**C**)	0	1	+1, Novel
	454T>**C**	NlaIV – (GGn'nC**C**)	4	5	+1
	454T>**C**	KpnI – (GGTAC'**C**)	0	1	+1, Novel
S000891623	67**G**>A	AluI - (AG'CT)	5	4	-1
	67**G**>A	CviJI - (rG'Cy)	27	26	-1
	67**G**>A	HindIII – (A'AGCTT)	1	0	-1
	435**A**>G, 454**T**>C	BfaI – (C'**TA**G)	8	6	-2
	454T>**C**	Acc65I – (G'GTAC**C**)	0	1	+1, Novel
	454T>**C**	BanI – (G'GyrC**C**)	0	1	+1, Novel
	454T>**C**	NlaIV – (GGn'nC**C**)	4	5	+1
	454T>**C**	KpnI – (GGTAC'**C**)	0	1	+1, Novel
S002288829	461T>C	BfaI – (C'**TA**G)	8	6	-2
	461T>C	Acc65I – (G'GTAC**C**)	0	1	+1, Novel
	461T>C	BanI – (G'GyrC**C**)	0	1	+1, Novel
	461T>C	NlaIV – (GGn'nC**C**)	4	5	+1
	461T>C	KpnI – (GGTAC'**C**)	0	1	+1, Novel
S002288840	461T>C	BfaI – (C'**TA**G)	8	6	-2
	461T>C	Acc65I – (G'GTAC**C**)	0	1	+1, Novel
	461T>C	BanI – (G'GyrC**C**)	0	1	+1, Novel
	461T>C	NlaIV – (GGn'nC**C**)	4	5	+1
	461T>C	KpnI – (GGTAC'**C**)	0	1	+1, Novel
	461T>C				
S002288829	74G>A	AluI - (AG'CT)	5	4	-1
	74G>A	CviJI - (rG'Cy)	27	26	-1
	74G>A	HindIII – (A'AGCTT)	1	0	-1
S002288840	74G>A	AluI - (AG'CT)	5	4	-1
	74G>A	CviJI - (rG'Cy)	27	26	-1
	74G>A	HindIII – (A'AGCTT)	1	0	-1
S002290679	74G>A	AluI - (AG'CT)	5	4	-1
	74G>A	CviJI - (rG'Cy)	27	26	-1
	74G>A	HindIII – (A'AGCTT)	1	0	-1
S002288829	442**A**>G	BfaI – (C'**TA**G)	8	6	-2
S002288840	442**A**>G	BfaI – (C'**TA**G)	8	6	-2
S002290804	1427G>A	CviJI - (r**G**'Cy)	27	26	-1
	1427G>A	NlaIV - (GG**n**'nCC)	4	4	0
S002290679	1427G>A	CviJI - (r**G**'Cy)	27	26	-1
	1427G>A	NlaIV - (GG**n**'nCC)	4	4	0

S002289298	1427G>A	CviJI - (r**G**'Cy)	27	26	-1
	1427G>A	NlaIV - (GG**n**'nCC)	4	4	0
***Clostridium botulinum* ATCC 3502 (AM412317)**					
S000858486	223**G**>A	MaeIII – ('**G**TnAC)	9	8	-1
S000858478	40**G**>A	AluI - (A**G**'CT)	5	4	-1
	40**G**>A	CviJI - (r**G**'Cy)	24	23	-1
	40**G**>A	HindIII – (A'A**G**CTT)	1	0	-1
S000858480	52G>**A**	Hpy188III - (TC'nn_G**A**)	6	7	+1
S000858491	40**G**>A	AluI - (A**G**'CT)	5	4	-1
	40**G**>A	CviJI - (r**G**'Cy)	24	23	-1
	40**G**>A	HindIII – (A'A**G**CTT)	1	0	-1
	52G>**A**	Hpy188III – (TC'nnG**A**)	6	7	+1
	408**A**>G, 427**T**>C	BfaI – (C'**TA**G)	8	6	-2
	427T>**C**	NlaIV – (GGn'nC**C**)	4	5	+1
	427T>**C**	KpnI – (GGTAC'**C**)	0	1	+1, Novel
	1193C>**T**	HpyCH4V – (TG'CA)	7	8	+1
S000858488	52G>**A**	Hpy188III – (TC'nnGA)	6	7	+1
	118**C**>T	BstBI – (TT'CGAA)	1	2	+1
	118**C**>T	TaqI - (T'CGA)	2	3	+1
	118**C**>T	Hpy188I – (T**C**n'GA)	4	3	-1
	363**C**>T	Fnu4HI – (G**C**'nGC)	14	13	-1
	363**C**>T	AciI – (**C**'CGC)	20	19	-1
	363**C**>T	TauI – (G**C**sG'C)	8	7	-1
	363C>**T**	HpyCH4IV – (A'CG**T**)	6	7	+1
	363C>**T**	TaiI – (ACG**T**')	6	7	+1
	363C>**T**	Hpy99I – (CG**w**CG')	0	1	+1, Novel
	408**A**>G, 427**T**>C	BfaI – (C'**TA**G)	8	6	-2
	427T>**C**	NlaIV – (GGn'nC**C**)	4	5	+1
	427T> **C**	KpnI – (GGTAC'**C**)	0	1	+1, Novel
S002290042	152C>T	BstBI – (TT'CGAA)	1	2	-1
	152C>T	TaqI - (T'CGA)	2	3	+1
	152C>T	Hpy188I – (T**C**n'GA)	4	3	-1
S002289503	74G>A	AluI - (A**G**'CT)	5	4	-1
	74G>A	CviJI - (r**G**'Cy)	24	23	-1
	74G>A	HindIII – (A'A**G**CTT)	1	0	-1
S002287460	74G>A	AluI - (A**G**'CT)	5	4	-1
	74G>A	CviJI - (r**G**'Cy)	24	23	-1
	74G>A	HindIII – (A'A**G**CTT)	1	0	-1
S002287580	129 G>A	BspMI – (**A**CCTGCnnnn'nnnn)	2	1	-1
S000858482	95G>A	BspMI – (**A**CCTGCnnnn'nnnn)	2	1	-1
S002288844	257 G>A	MaeIII – ('**G**TnAC)	9	8	-1
S002287460	440 A>G	BfaI – (C'**TA**G)	8	6	-2
S000858488	440 A>G	BfaI – (C'**TA**G)	8	6	-2
S002287580	440 A>G	BfaI – (C'**TA**G)	8	6	-2
S002287580	461T>C	BfaI – (C'**TA**G)	8	6	-2
S002290042	461T>C	BfaI – (C'**TA**G)	8	6	-2
S002287460	461T>C	BfaI – (C'**TA**G)	8	6	-2
S002290042	397C>T	Fnu4HI – (G**C**'nGC)	14	13	-1
	397C>T	AciI – (**C**'CGC)	20	19	-1
	397C>T	TauI – (G**C**sG'C)	8	7	-1
	397C>T	HpyCH4IV – (A'CG**T**)	6	7	+1
	397C>T	TaiI – (ACG**T**')	6	7	+1
	397C>T	Hpy99I – (CG**w**CG')	0	1	+1, Novel
***Clostridium botulinum* 657 (CP001083)**					
S001416082	84G>**A**	Tsp509I – ('**A**ATT)	3	4	+1
S001416090	84G>**A**	Tsp509I – ('**A**ATT)	3	4	+1
S001416088	80A>**G**	EcoNI - (CCTnn'n_nnA**G**G)	2	3	+1
	80A>**G**	BslI – (CCnn_nnn'nn**G**G)	7	8	+1
S001416095	222**A**>G	AvaII – (G'G**w**CC)	2	1	-1
	222A>**G**	Sau96I – (G'G**n**CC)	5	6	+1
	222A>**G**	NlaIV – (GG**n**'nCC)	4	4	0

	222A>**G**	PspOMI – (G'G**G**CCC)	1	2	+1
	222A>**G**	CviJI – (r**G**'Cy)	26	27	+1
	222A>**G**	HaeIII – (G**G**'CC)	4	5	+1
	222A>**G**	ApaI – (GG**G**CC'C)	1	2	+1
	222A>**G**	BanII – (Gr**G**Cy'C)	1	2	+1
	222A>**G**	Bme1580I – (Gk**G**Cm'C)	1	2	+1
	222A>**G**	Bsp1286I – (Gd**G**Ch'C)	2	3	+1
	222A>**G**	Cac8I – (**G**Cn'nGC)	5	6	+1
S001416092	254**C**>T	MaeIII -('GTnA**C**)	9	8	-1
	997**C**>T	XmnI – (GAAnn'nnTT**C**)	3	2	-1
	1420A>**G**	NlaIV – (GGn'**n**CC)	4	4	0
	1420A>**G**	CviJI – (r**G**'Cy)	26	27	+1
	1420A>**G**	Cac8I – (GCn'**n**GC)	5	6	+1
S001416097	254**C**>T	MaeIII – ('GTnA**C**)	9	8	-1
S002289967	229A>G	AvaII – (G'G**w**CC)	2	1	-1
	229A>G	Sau96I – (G'G**n**CC)	5	6	+1
	229A>G	NlaIV – (GG**n**'nCC)	4	4	0
	229A>G	PspOMI – (G'G**G**CCC)	1	2	+1
	229A>G	CviJI – (r**G**'Cy)	26	27	+1
	229A>G	HaeIII – (G**G**'CC)	4	5	+1
	229A>G	ApaI – (GG**G**CC'C)	1	2	+1
	229A>G	BanII – (Gr**G**Cy'C)	1	2	+1
	229A>G	Bme1580I – (Gk**G**Cm'C)	1	2	+1
	229A>G	Bsp1286I – (Gd**G**Ch'C)	2	3	+1
	229A>G	Cac8I – (**G**Cn'nGC)	5	6	+1
S002289183	1427A>G	NlaIV – (GGn'**n**CC)	4	4	0
	1427A>G	CviJI – (r**G**'Cy)	26	27	+1
	1427A>G	Cac8I – (GCn'**n**GC)	5	6	+1
S002287409	87G>**A**	EcoNI - (CCTnn'n_nnA**G**G)	2	3	+1
	87G>**A**	BslI – (CCnn_nnn'nn**G**G)	7	8	+1
S002290186	91G>**A**	Tsp509I – ('**A**ATT)	3	4	+1
S002290081	91G>**A**	Tsp509I – ('**A**ATT)	3	4	+1
S002287812	261C>T	MaeIII -('GTnA**C**)	9	8	-1
S002289183	261C>T	MaeIII -('GTnA**C**)	9	8	-1
S002289183	1004C>T	XmnI – (GAAnn'nnTT**C**)	3	2	-1
***Clostridium botulinum* Alaska E43 (CP001078)**[b]					
S001094760	993G>**A**	MaeIII – ('GTn**A**C)	4	6	+2
	993G>**A**	Tsp45I – ('GTsAC)	1	3	+2
	993G>**A**	AleI – (CACnn'nnGTG)	0	1	+1, Novel
	993G>**A**	MslI – (CAynn'nnrTG)	0	1	+1, Novel
	1000C>**T**	HphI – (GG**T**GAnnnnnnn_n')	6	7	+1
S001094753	993G>**A**	MaeIII – ('GTn**A**C)	4	6	+2
	993G>**A**	Tsp45I – ('GTsAC)	1	3	+2
	993G>**A**	AleI – (CACnn'nnGTG)	0	1	+1, Novel
	993G>**A**	MslI – (CAynn'nnrTG)	0	1	+1, Novel
	1000C>**T**	HphI – (GG**T**GAnnnnnnn_n')	6	7	+1
S001094755	983**C**>A	TaqI – (T'**C**GA)	5	4	-1
	993G>**A**	MaeIII – ('GTn**A**C)	4	6	+2
	993G>**A**	Tsp45I – ('GTs**A**C)	1	3	+2
	993G>**A**	AleI – (C**A**Cnn'nnGTG)	0	1	+1, Novel
	993G>**A**	MslI – (C**A**ynn'nnrTG)	0	1	+1, Novel
	1000C>**T**	HphI – (GG**T**GAnnnnnnn_n')	6	7	+1
S001094758	993G>**A**	MaeIII – ('GTn**A**C)	4	6	+2
	993G>**A**	Tsp45I – ('GTs**A**C)	1	3	+2
	993G>**A**	AleI – (C**A**Cnn'nnGTG)	0	1	+1, Novel
	993G>**A**	MslI – (C**A**ynn'nnrTG)	0	1	+1, Novel
	1000C>**T**	HphI – (GG**T**GAnnnnnnn_n')	6	7	+1
S001094762	993G>**A**	MaeIII – ('GTn**A**C)	4	6	+2
	993G>**A**	Tsp45I – ('GTs**A**C)	1	3	+2
	993G>**A**	AleI – (CACnn'nnGTG)	0	1	+1, Novel
	993G>**A**	MslI – (C**A**ynn'nnrTG)	0	1	+1, Novel
	1000C>T	HphI – (GGTGAnnnnnnn_n')	6	7	+1
S002289948[C]	999G>**A**	MaeIII – ('GTn**A**C)	4	6	+2
	999G>**A**	Tsp45I – ('GTsAC)	1	3	+2

	999G>**A**	AleI – (CACnn'nnGTG)	0	1	+1, Novel
	999G>**A**	MslI – (CAynn'nnrTG)	0	1	+1, Novel
S002290927	999G>**A**	MaeIII – ('GTn**A**C)	4	6	+2
	999G>**A**	Tsp45I – ('GTsAC)	1	3	+2
	999G>**A**	AleI – (CACnn'nnGTG)	0	1	+1, Novel
	999G>**A**	MslI – (CAynn'nnrTG)	0	1	+1, Novel
S002287743	999G>**A**	MaeIII – ('GTn**A**C)	4	6	+2
	999G>**A**	Tsp45I – ('GTsAC)	1	3	+2
	999G>**A**	AleI – (CACnn'nnGTG)	0	1	+1, Novel
	999G>**A**	MslI – (CAynn'nnrTG)	0	1	+1, Novel
S002288016	999G>**A**	MaeIII – ('GTn**A**C)	4	6	+2
	999G>**A**	Tsp45I – ('GTsAC)	1	3	+2
	999G>**A**	AleI – (CACnn'nnGTG)	0	1	+1, Novel
	999G>**A**	MslI – (CAynn'nnrTG)	0	1	+1, Novel
S002288105	999G>**A**	MaeIII – ('GTn**A**C)	4	6	+2
	999G>**A**	Tsp45I – ('GTsAC)	1	3	+2
	999G>**A**	AleI – (CACnn'nnGTG)	0	1	+1, Novel
	999G>**A**	MslI – (CAynn'nnrTG)	0	1	+1, Novel
S002288105	989C>A	TaqI – (T'**C**GA)	5	4	-1
***Clostridium botulinum* ATCC 19397 (CP000726) / Hall (CP000727)**					
S000891582 / S000891598	250**G**>A	MaeIII – ('**G**TnAC)	9	8	-1
S000891576 / S000891592	67**G**>A	AluI - (A**G**'CT)	5	4	-1
	67**G**>A	HindIII – (A'A**G**CTT)	1	0	-1
	67**G**>A	CviJI - (r**G**'Cy)	27	25	-2
S000891578 / S000891594	79G>**A**	Hpy188III – (TC'nnG**A**)	7	8	+1
	435**A**>G 454**T**>C	BfaI – (C'**TA**G)	8	6	-2
	454T>**C**	Acc65I – (G'GTAC**C**)	0	1	+1, Novel
	454T>**C**	BanI – (G'GyrC**C**)	0	1	+1, Novel
	454T>**C**	NlaIV – (GGn'nC**C**)	4	5	+1
	454T>**C**	KpnI – (GGTAC'**C**)	0	1	+1, Novel
S000891587 / S000891603	67**G**>A	HindIII – (A'A**G**CTT)	1	0	-1
	79G>**A**	Hpy188III – (TC'nnG**A**)	7	8	+1
	435**A**>G, 454**T**>C	BfaI – (C'T**A**G)	8	6	-2
	454T>**C**	Acc65I – (G'GTAC**C**)	0	1	+1, Novel
	454T>**C**	BanI – (G'GyrC**C**)	0	1	+1, Novel
	454T>**C**	NlaIV – (GGn'nC**C**)	4	5	+1
	454T>**C**	KpnI – (GGTAC'**C**)	0	1	+1, Novel
	1220**C**>T	HgaI – (GA**C**GCnnnnn'nnnnn)	3	2	-1
	1220C>**T**	HpyCH4V - (**T**G'CA)	7	8	+1
S000891584 / S000891600	67**G**>A	HindIII – (A'A**G**CTT)	1	0	-1
	78G>**A**, 79G>**A**	Hpy188I – (TC_**n**'**G**A)	4	5	+1
	454T>**C**	Acc65I – (G'GTAC**C**)	0	1	+1, Novel
	454T>**C**	BanI – (G'GyrC**C**)	0	1	+1, Novel
	454T>**C**	NlaIV – (GGn'nC**C**)	4	5	+1
	454T>**C**	KpnI – (GGTAC'**C**)	0	1	+1, Novel
S000891580 / S000891596	79G>**A**	Hpy188III – (TC'nnGA)	7	8	+1
	122**G**>A	BspMI – (**A**CCTGCnnnn'nnnn)	2	1	-1
	454T>**C**	Acc65I – (G'GTAC**C**)	0	1	+1, Novel
	454T>**C**	BfaI	8	6	-2
	454T>**C**	BanI – (G'GyrC**C**)	0	1	+1, Novel
	454T>**C**	NlaIV – (GGn'nC**C**)	4	5	+1
	454T>**C**	KpnI – (GGTAC'**C**)	0	1	+1, Novel
S002289359	1227C>T	HgaI – (GA**C**GCnnnnn'nnnnn)	3	2	-1
	1227C>T	HpyCH4V - (**T**G'CA)	7	8	+1
S002287857[c]	461T>C	BfaI – (C'**TA**G)	8	6	-2
	461T>C	Acc65I – (G'GTAC**C**)	0	1	+1, Novel
	461T>C	BanI – (G'GyrC**C**)	0	1	+1, Novel
	461T>C	NlaIV – (GGn'nC**C**)	4	5	+1

	461T>C	KpnI – (GGTAC'**C**)	0	1	+1, Novel
S002288115[c]	461T>C	BfaI – (C'**TA**G)	8	6	-2
	461T>C	Acc65I – (G'GTAC**C**)	0	1	+1, Novel
	461T>C	BanI – (G'GyrC**C**)	0	1	+1, Novel
	461T>C	NlaIV – (GGn'nC**C**)	4	5	+1
	461T>C	KpnI – (GGTAC'**C**)	0	1	+1, Novel
S002289359[c]	461T>C	BfaI – (C'**TA**G)	8	6	-2
	461T>C	Acc65I – (G'GTAC**C**)	0	1	+1, Novel
	461T>C	BanI – (G'GyrC**C**)	0	1	+1, Novel
	461T>C	NlaIV – (GGn'nC**C**)	4	5	+1
	461T>C	KpnI – (GGTAC'**C**)	0	1	+1, Novel
S002287857	86G>**A**	Hpy188III – (TC'nnG**A**)	7	8	+1
S002287510[d]	74G>A	AluI - (A**G**'CT)	5	4	-1
	74G>A	HindIII – (A'A**G**CTT)	1	0	-1
	74G>A	CviJI - (r**G**'Cy)	27	25	-2
S002289359[d]	74G>A	AluI - (A**G**'CT)	5	4	-1
	74G>A	HindIII – (A'A**G**CTT)	1	0	-1
	74G>A	CviJI - (r**G**'Cy)	27	25	-2
S002290093[d]	74G>A	AluI - (A**G**'CT)	5	4	-1
	74G>A	HindIII – (A'A**G**CTT)	1	0	-1
	74G>A	CviJI - (r**G**'Cy)	27	25	-2
S002288115[d]	129 G>A	BspMI – (**A**CCTGCnnnn'nnnn)	2	1	-1
S002290387	257 G>A	MaeIII – ('**G**TnAC)	9	8	-1
S002287857[d],	442A>G	BfaI – (C'**TA**G)	8	6	-2
S002288115[d],	442A>G	BfaI – (C'**TA**G)	8	6	-2
S002290093[d],	442A>G	BfaI – (C'**TA**G)	8	6	-2
S002289359[d]	442A>G	BfaI – (C'**TA**G)	8	6	-2
Clostridium botulinum **Eklund 17B (CP001056)**					
S001094712	211**G**>A	AciI – (C'C**G**C)	20	19	-1
	983**C**>A	TaqI – (T'**C**GA)	5	4	-1
S001094703	68**G**>A	AluI - (**AG'CT**)	7	6	-1
	68**G**>A	HindIII – (A'A**G**CTT)	1	0	-1
	68**G**>A	CviJI - (r**G**'Cy)	27	26	-1
S001094716	68**G**>A	A luI - (A**G'**CT)	7	6	-1
	68**G**>A	HindIII – (A'A**G**CTT)	1	0	-1
	68**G**>A	CviJI - (r**G**'Cy)	27	26	-1
S001094722	68**G**>A	A luI - (A**G'**CT)	7	6	-1
	68**G**>A	HindIII – (A'A**G**CTT)	1	0	-1
	68**G**>A	CviJI - (r**G**'Cy)	27	26	-1
S001094714	68**G**>A	AluI - (**AG'CT**)	7	6	-1
	68**G**>A	HindIII – (A'A**G**CTT)	1	0	-1
	68**G**>A	CviJI - (r**G**'Cy)	27	26	-1
	1232**C**>T	AciI – (C'CG**C**)	20	19	-1
	1232**C**>T	BstUI – (CG'**C**G)	8	7	-1
	1232C>**T**	HpyCH4III – (ACn'G**T**)	3	4	+1
S001094718	68**G**>A	AluI - (**AG'CT**)	7	6	-1
	68**G**>A	HindIII – (A'A**G**CTT)	1	0	-1
	68**G**>A	CviJI - (r**G**'Cy)	27	26	-1
	1232**C**>T	AciI – (C'CG**C**)	20	19	-1
	1232**C**>T	BstUI – (CG'**C**G)	8	7	-1
	1232C>**T**	HpyCH4III – (ACn'G**T**)	3	4	+1
S001094707	77**G**>A	Hpy188I – (TCn'**G**A)	4	3	-1
	77G>**A**	BstBI – (TT'CG**A**A)	2	3	+1
	77G>**A**	TaqI – (T'CG**A**)	5	6	+1
	993**A**>G	MaeIII – ('GTn**A**C)	6	4	-2
	993**A**>G, 1000**T**>C	Tsp45I – ('GTs**A**C)	3	1	-2
	993**A**>G, 1000**T**>C	AleI – (C**A**Cnn'nnGTG)	1	0	-1
	993**A**>G	MslI – (C**A**ynn'nnrTG)	1	0	-1
	1000**T**>C	MaeIII – ('G**T**nAC)	6	4	-2
	1000**T**>C	MslI – (CAynn'nnr**T**G)	1	0	-1

	1232**C**>T	AciI – (C'CG**C**)	20	19	-1
	1232**C**>T	BstUI – (CG'**C**G)	8	7	-1
	1232C>**T**	HpyCH4III – (AC_n'G**T**)	3	4	+1
S001094720	77**G**>A	Hpy188I – (TC_n'**G**A)	4	3	-1
	77G>**A**	BstBI – (TT'C**G**AA	2	3	+1
	77G>**A**	TaqI – (T'C**G**A)	5	6	+1
	1232**C**>T	AciI – (C'CGC)	20	19	-1
	1232**C**>T	BstUI – (CG'**C**G)	8	7	-1
	1232C>**T**	HpyCH4III – (ACn'G**T**)	3	4	+1
S002288156	83G>A	Hpy188I – (TCn'**G**A)	4	3	-1
	83G>A	BstBI – (TT'CG**A**A)	2	3	+1
	83G>A	TaqI – (T'CG**A**)	5	6	+1
S002289462	1238C>T	AciI – (C'CGC)	20	19	-1
	1238C>T	BstUI – (CG'**C**G)	8	7	-1
	1238C>T	HpyCH4III – (ACn'G**T**)	3	4	+1
S002290368	1238C>T	AciI – (C'CGC)	20	19	-1
	1238C>T	BstUI – (CG'**C**G)	8	7	-1
	1238C>T	HpyCH4III – (ACn'G**T**)	3	4	+1
S002290247	1238C>T	AciI – (C'CGC)	20	19	-1
	1238C>T	BstUI – (CG'**C**G)	8	7	-1
	1238C>T	HpyCH4III – (ACn'G**T**)	3	4	+1
S002288156	1238C>T	AciI – (C'CGC)	20	19	-1
	1238C>T	BstUI – (CG'**C**G)	8	7	-1
	1238C>T	HpyCH4III – (ACn'G**T**)	3	4	+1
S002290154	74G>A	AluI - (**AG'CT**)	7	6	-1
	74G>A	HindIII – (A'A**G**CTT)	1	0	-1
	74G>A	CviJI - (r**G**'Cy)	27	26	-1
S002290368	74G>A	AluI - (**AG'CT**)	7	6	-1
	74G>A	HindIII – (A'A**G**CTT)	1	0	-1
	74G>A	CviJI - (r**G**'Cy)	27	26	-1
S002290247	74G>A	AluI - (**AG'CT**)	7	6	-1
	74G>A	HindIII – (A'A**G**CTT)	1	0	-1
	74G>A	CviJI - (r**G**'Cy)	27	26	-1
S002287177	74G>A	AluI - (**AG'CT**)	7	6	-1
	74G>A	HindIII – (A'A**G**CTT)	1	0	-1
	74G>A	CviJI - (r**G**'Cy)	27	26	-1
S002290659	218G>A	AciI – (C'C**G**C)	20	19	-1
S002290659	989 C>A	TaqI – (T'**C**GA)	5	4	-1
S002289462	1006T>C	Tsp45I – ('GTs**A**C)	3	1	-2
	1006T>C	AleI – (C**A**Cnn'nnGTG)	1	0	-1
	1006T>C	MaeIII – ('G**T**nAC)	6	4	-2
	1006T>C	MslI – (CAynn'nnr**T**G)	1	0	-1
***Clostridium botulinum* Loch Maree (CP000962)**					
S001014494	67A>**G**	AluI - (A**G**'CT)	4	5	+1
	67A>**G**	CviJI - (r**G**'Cy)	26	27	+1
	67A>**G**	HindIII – (A'A**G**CTT)	0	1	+1, Novel
S001014499	67A>**G**	AluI - (A**G**'CT)	4	5	+1
	67A>**G**	CviJI - (r**G**'Cy)	26	27	+1
	67A>**G**	HindIII – (A'A**G**CTT)	0	1	+1, Novel
S001014484	1220 C>**T**	HpyCH4V – (**T**G'CA)	7	8	+1
S001014492	1220 C>**T**	HpyCH4V – (**T**G'CA)	7	8	+1
S001014501	82T>**C**	AciI – C’CG**C**/G'**C**G**G**	21	23	+2
	1317A**>G**	AciI – C’CG**C**/G'**C**G**G**	21	23	+2
	1319**A**>G	Hpy188III – (TC'nnG**A**)	8	7	-1
	1319**A**>G	NruI – (TCG'C**G**A)	1	0	-1
	1319**A**>G	HinfI – (G'**A**nTC)	3	2	-1
	1319**A**>G	TfiI – (G'**A**wTC)	2	1	-1
	1319A>**G**	MboI – ('**G**ATC)	3	4	+1
	1319A>**G**	DpnI – (**G**A'TC)	3	4	+1
	1319A>**G**	BstKTI – (**G**AT'C)	3	4	+1
	1319A>**G**	AlwI – (G**G**ATCnn)	2	3	+1
S002290692	1326A>G	AciI – C’CG**C**/G'**C**G**G**	21	23	+2

	1326A>G	Hpy188III – (TC'nnG**A**)	8	7	-1
	1326A>G	NruI – (TCG'C**G**A)	1	0	-1
	1326A>G	HinfI – (G'**A**nTC)	3	2	-1
	1326A>G	TfiI – (G'**A**wTC)	2	1	-1
	1326A>G	MboI – ('**G**ATC)	3	4	+1
	1326A>G	DpnI – (**G**A'TC)	3	4	+1
	1326A>G	BstKTI – (**G**AT'C)	3	4	+1
	1326A>G	AlwI – (G**G**ATCnn)	2	3	+1
S002287648	74A>**G**	AluI - (A**G**'CT)	4	5	+1
	74A>**G**	CviJI - (r**G**'Cy)	26	27	+1
	74A>**G**	HindIII – (A'A**G**CTT)	0	1	+1, Novel
S002290134	74A>**G**	AluI - (A**G**'CT)	4	5	+1
	74A>**G**	CviJI - (r**G**'Cy)	26	27	+1
	74A>**G**	HindIII – (A'A**G**CTT)	0	1	+1, Novel
S002290692	89T>**C**	AciI – C’CG**C**/G'**C**G**G**	21	23	+2
S002290255	1227C>**T**	HpyCH4V – (**T**G'CA)	7	8	+1
S002289199	1227C>**T**	HpyCH4V – (**T**G'CA)	7	8	+1
***Clostridium botulinum* Kyoto (CP001581)**					
S001350513	84 G>**A**	Tsp509I – ('**A**ATT)	3	4	+1
S001350515	435 **A**>G, 454 **T**>C	BfaI – (C'T**A**G)	8	6	-2
	454 T>**C**	Acc65I – (G'GTAC**C**)	0	1	+1, Novel
	454 T>**C**	BanI – (G'GyrC**C**)	0	1	+1, Novel
	454 T>**C**	NlaIV – (GGn'nC**C**)	4	5	+1
	454 T>**C**	KpnI – (GGTAC'**C**)	0	1	+1, Novel
S002290406	461T>C	BfaI – (C'T**A**G)	8	6	-2
	461T>C	Acc65I – (G'GTAC**C**)	0	1	+1, Novel
	461T>C	BanI – (G'GyrC**C**)	0	1	+1, Novel
	461T>C	NlaIV – (GGn'nC**C**)	4	5	+1
	461T>C	KpnI – (GGTAC'**C**)	0	1	+1, Novel
S002288993	91G>**A**	Tsp509I – ('**A**ATT)	3	4	+1
S002290406	442A>G	BfaI – (C'T**A**G)	8	6	-2

***Clostridium botulinum* Okra (CP000939)**					
S001014394	1420 **G**>A	CviJI - (r**G**'Cy)	27	26	-1
	1420 **G**>A	NlaIV - (GGn'**n**CC)	4	4	0
S001014396	1420 **G**>A	CviJI - (r**G**'Cy)	27	26	-1
	1420 **G**>A	NlaIV - (GGn'**n**CC)	4	4	0
S001014407	974 **G**>A, 1420 **G**>A	CviJI - (r**G**'Cy)	27	25	-2
	1220C>**T**	HpyCH4V – (**T**G'CA)	7	8	+1
	1420 **G**>A	NlaIV - (GGn'**n**CC)	4	4	0
S001014404	1220C>**T**	HpyCH4V – (**T**G'CA)	7	8	+1
	171T>**C**, 174A>**G**	FatI – ('**C**AT**G**)	7	8	+1
	171T>**C**, 174A>**G**	NlaIII – (**C**AT**G**')	7	8	+1
S002288021	178T>C, 181A>G	FatI – ('**C**AT**G**)	7	8	+1
S002288021	1227C>**T**	HpyCH4V – (**T**G'CA)	7	8	+1
S002288411	1227C>**T**	HpyCH4V – (**T**G'CA)	7	8	+1
S002288411	981 G>A	CviJI - (r**G**'Cy)	27	25	-2
S002288411	1000 A>G	Tsp45I – ('GTs**A**C)	3	1	-2
	1000 A>G	AleI – (C**A**Cnn'nnGTG)	1	0	-1
	1000 A>G	MaeIII – ('G**T**nAC)	6	4	-2
	1000 A>G	MslI – (CAynn'nnr**T**G)	1	0	-1
S001014407	993A>G	Tsp45I – ('GTs**A**C)	3	1	-2
	993A>G	AleI – (C**A**Cnn'nnGTG)	1	0	-1
	993A>G	MaeIII – ('G**T**nAC)	6	4	-2
	993A>G	MslI – (CAynn'nnr**T**G)	1	0	-1
S002287425	1427 G>A	CviJI - (r**G**'Cy)	27	26	-1
	1427 G>A	NlaIV - (GGn'**n**CC)	4	4	0

S002288411	1427 G>A	CviJI - (r**G**'Cy)	27	26	-1
	1427 G>A	NlaIV - (GGn'**n**CC)	4	4	0
H04402 065					
S002408098	454T>C	Acc65I – (G'GTAC**C**)	0	1	+1, Novel
	454T>C	BanI – (G'GyrC**C**)	0	1	+1, Novel
	454T>C	NlaIV – (GGn'nC**C**)	4	5	+1
	454T>C	KpnI – (GGTAC'**C**)	0	1	+1, Novel
S002408093	861G>A	HpyCH4III -AC_n'GT	6	5	-1
	861G>A	MboI-'GATC_	3	3	0
	861G>A	DpnI-GA'TC	3	4	+1
	861G>A	BsiEI-CG_ry'CG	2	2	0
	861G>A	PvuI-, CG_AT'CG	0	1	+1 Novel
	861G>A	BstKTI-G_AT'C	3	4	+1
	861G>A				
S002408093	67A>G	AluI - (A**G**'CT)	5	4	-1
	67**G**>A	HindIII – (A'A**G**CTT)	1	0	-1
	67**G**>A	CviJI - (r**G**'Cy)	27	25	-2
S002408095	82T>C	AciI – C’CG**C**/G'**C**G**G**	21	23	+2
S002408089	1220C>**T**	HpyCH4V – (**T**G'CA)	7	8	+1
S002408093	1220C>**T**	HpyCH4V – (**T**G'CA)	7	8	+1
S002408095	1220C>**T**	HpyCH4V – (**T**G'CA)	7	8	+1
230613					
S002165522[d]	454T>C	Acc65I – (G'GTAC**C**)	0	1	+1, Novel
	454T>C	BanI – (G'GyrC**C**)	0	1	+1, Novel
	454T>C	NlaIV – (GGn'nC**C**)	4	5	+1
	454T>C	KpnI – (GGTAC'**C**)	0	1	+1, Novel
S002165524	454T>C	Acc65I – (G'GTAC**C**)	0	1	+1, Novel
	454T>C	BanI – (G'GyrC**C**)	0	1	+1, Novel
	454T>C	NlaIV – (GGn'nC**C**)	4	5	+1
	454T>C	KpnI – (GGTAC'**C**)	0	1	+1, Novel
S002165515	67G >A	AluI - (A**G**'CT)	5	4	-1
	67G >A	HindIII – (A'A**G**CTT)	1	0	-1
	67G >A	CviJI - (r**G**'Cy)	27	25	-2
S002165522	67G >A	AluI - (A**G**'CT)	5	4	-1
	67G >A	HindIII – (A'A**G**CTT)	1	0	-1
	67G >A	CviJI - (r**G**'Cy)	27	25	-2
S002165524	67G >A	AluI - (A**G**'CT)	5	4	-1
	67G >A	HindIII – (A'A**G**CTT)	1	0	-1
	67G >A	CviJI - (r**G**'Cy)	27	25	-2
S002165522	435A>G	BfaI – (C'**TA**G)	8	6	-2
S002165524	435A>G	BfaI – (C'**TA**G)	8	6	-2
BKT015925					
S002441243	581C>T	ApoI-r'AATT_y	2	2	0
	581C>T	Tsp509I-'AATT	6	6	0
	581C>T	EcoRI-G'AATT_C	1	0	-1
S002441251	73A>**G**	AluI - (A**G**'CT)	5	4	-1
	73A>**G**	HindIII – (A'A**G**CTT)	1	0	-1
	73A>**G**	CviJI - (r**G**'Cy)	27	25	-2
S002441255	73A>**G**	AluI - (A**G**'CT)	5	4	-1
	73A>**G**	HindIII – (A'A**G**CTT)	1	0	-1
	73A>**G**	CviJI - (r**G**'Cy)	27	25	-2
S002441247	73A>**G**	**AluI - (AG'CT)**	5	4	-1
	73A>**G**	HindIII – (A'A**G**CTT)	1	0	-1
	73A>**G**	CviJI - (r**G**'Cy)	27	25	-2
S002441251	246C>T	AluI - AG'CT	5	4	-1
S002441255	246C>T	CviJI - rG'Cy	27	25	-2
	246C>T	BfaI - C'TA_G	7	7	0

[a] Y = C or T, R = G or A, W = T or A, K = T or G, S = G or C, M = C or A, D = A or G or T, H = A or C or T

[b] In *C. botulinum* Alaska E43 (CP001078), there were no RE sites at positions 993 and 1000. However, due to nucleotide substitutions at (i) 993G>A, 4 RE sites were generated: 2 at 990 for MaeIII and Tsp45I and 2 novel sites at 993 for AleI and MslI, and (ii) 1000C>T, 3 RE sites were generated: 1 at 998 for MaeIII, 1 at 999 for Tsp45I and 1 at 1000 for HpHI. The impact of these changes were evident as follows: (i) number of MaeIII and Tsp45I sites increased by 2 each, (ii) number of HphI site increased by 1, and (iii) novel sites were created for AleI and MslI.

Table S8

Loss of restriction endonuclease sites due to nucleotide changes in *rrs* of *Clostridium botulinum* strains

***Clostridium botulinum* strain**	**RDP Accession No.**	**Nucleotide Changes[a]**	**Affected Restriction Endonuclease[b]**
ATCC 3502	S000858478, S000858491	40G>A	Alu I - (AG'CT), CviJI - (rG'Cy), HindIII – (A'AGCTT)
ATCC 19397	S000891576[d], S000891584[d], S000891587[d]	67G>A	
Langeland	S000891614, S000891621[c]	67G >A	
230613	S002165515, S002165522, S002165524	67G >A	
H04402 065	S002408093, S002408095, S002408098	67G >A	
Eklund 17B	S001094703, S001094714[c], S001094716[c]	68G>A	
ATCC 3502	S002289503, S002287460	74G>A	
Langeland	S002288829[C], S002290679	74G>A	
ATCC 19397	S002287510[d], S002289359[d], S002290093[d]	74G>A	
Eklund 17B	S002290154, S002290368[c,] S002287177	74 G>A	
ATCC 3502	S000858482	95G>A	BspMI – (ACCTGCnnnn'nnnn)
ATCC 19397	S000891580[d]	122G>A	
	S002288115[d]	129 G>A	
ATCC 3502	S002287580	129 G>A	
Eklund 17B	S001094712	211G>A	AciI – (C'CGC)
	S002290659	218G>A	
ATCC 3502	S000858486	223G>A	MaeIII -('GTnAC)
BKT015925	S002441251, S002441255	246C>T	AluI - AG'CT, CviJI - rG'Cy, BfaI - C'TA_G
H04402 065	S002408085	254C>T	MaeIII -('GTnAC)
ATCC 3502	S002288844	257 G>A	
ATCC 19397	S000891582[d]	250G>A	
657	S001416092, S001416097	254C>T	
ATCC 19397	S002290387	257 G>A	
657	S002287812, S002289183	261C>T	
ATCC 3502	S000858482, S000858488, S000858491	408A>G	BfaI – (C'TAG)
Langeland	S000891621[c]	435A>G	
Kyoto	S001350515	435A>G	
230613	S002165522, S002165524	435A>G	
ATCC 19397	S000891578[d], S000891580[d], S000891584[d], S000891587[d]	435A>G	
ATCC 3502	S002287460, S000858488, S002287580	440 A>G	
Langeland	S002288829[C]	442A>G	
Kyoto	S002290406	442A>G	
ATCC 19397	S002287857[d], S002288115[d], S002290093[d], S002289359[d]	442A>G	
Okra	S001014407	974G>A	CviJI - (rG'Cy)
	S002288411	981 G>A	
Eklund 17B	S001094712	983C>A	TaqI – (T'CGA)
	S002290659	989 C>A	
Alaska E43	S001094755	983C>A	TaqI – (T'CGA)
	S002288105	989C>A	
Okra	S001014407	993A>G	AleI – (CACnn'nnGTG), MaeIII – ('GTnAC), MslI – (CAynn'nnrTG), Tsp45I – ('GTsAC)
657	S001416092	997C>T	XmnI – (GAAnn'nnTTC)
Okra	S002288411	1000 A>G	AleI – (CACnn'nnGTG), MaeIII – ('GTnAC) MslI – (CAynn'nnrTG), Tsp45I – ('GTsAC)
Eklund 17B	S001094707	1000T>C	
657	S002289183	1004C>T	XmnI – (GAAnn'nnTTC)
Eklund 17B	S002289462	1006T>C	AleI – (CACnn'nnGTG), MaeIII – ('GTnAC) MslI – (CAynn'nnrTG), Tsp45I – ('GTsAC)

Langeland	S000891606, S000891614, S000891616	1420G>A	CviJI- rG'Cy
Okra	S001014394, S001014407	1420G>A	
H04402 065	S002408095	1420G>A	
Okra	S001014394, S001014407	1420G>A	
	S002287425, S002288411	1427 G>A	
Langeland	S002290804, S002290679, S002289298	1427G>A	

[a] Designated by a “>”-character, 40G>A denotes that at nucleotide 40 a G is changed to a A.

[b] y = C or T, r = G or A, s = G or C.

[c] Intragenomic sequences (see Table S2) with changes at similar position.

[d] Intergenomic sequences (see Table S2) with changes at similar position.

Table **S9**

Affected Restriction enzyme (RE) sites in *rrs* of sequenced genomes of different *Clostridium botulinum*

***C. botulinum* strains**	***rrs***		**Changes in RE sites**[a]		
	Total	**Affected**	**Novel**	**Addition**	**Loss**
ATCC 19397/HALL	16	12	21	44	44
Alaska E43	22	10	20	45	2
Langeland	18	13	12	6	36
Kyoto	18	4	6	4	6
230613	9	3	6	2	16
Loch Maree	18	10	4	28	8
ATCC 3502	18	12	4	13	41
H04402065	9	6	3	8	5
657	18	12	0	30	8
BKT015925	10	4	0	0	12
Okra	18	7	0	7	20
Eklund 17B	22	15	0	14	65

a: Details are presented in Additional Table S7.

Table S10

Location of genes in the 10 completely sequenced genomes of *Clostridium botulinum* used for Cladistic analysis

Gene	***Clostridium botulinum* strains**											
	ATCC 3502/ AM412317	**Loch Maree/ CP000962**	**Kyoto/ CP001581**	**Okra/ CP000939**	**657/ CP001083**	**Langeland/ CP000728**	**ATCC 19397/ CP000726**	**Hall/ CP000727**	**Eklund 17B/ CP001056**	**Alaska E43/ CP001078**	**FR773526/ H04402 065**	**CP002410/ BKT015925**	**CP002011/230613**
AbrB	95044..95289	88829..89074	88931..89177	94701..94946	88952..89197	89739..89984	87847..88092	87849..88094	NA	NA	NA	NA	NA
CphB	154869..155669	148290..149090	148752..149552	154658..155458	151949..152749	149675..150475	147671..148471	147673..148473	NA	NA	NA	NA	149938..150738
DltB	68317..68508	NA	NA	NA	NA	NA	NA	NA	687020..688174	653993..655147	NA	NA	NA
DltD	68937..69329	62683..63105	63230..63388	68848..69105	62164..63210	63212..64258	NA	NA	684195..685379	651168..652352	NA	NA	NA
SfsA	41636..42328	35845..36537	36050..36742	41727..42419	35626..36318	36993..37685	41651..42343	41651..42343	NA	2190482..2191174	NA	2730964..2731656	37061..37753
SpmB	101717..102235	95501..96019	95606..96124	101374..101892	95626..96144	NA	94520..95038	94522..95040	110541..111068	100563..101090	NA	2640944..2641468	NA
arcA1	97129..98340	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
CcpA	131661..132659	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
MdeA	35687..36889	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
psd1	24804..25562, 25565..25690	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
psd2	30610..31497	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
AtpA	185799..186479	181244..182758	182386..183900	188341..189855	185326..186840	182856..184370	180668..182182	180670..182184	495241..496755	490312..491826	NA	2465450..2466958	183117..184631
AtpC	186519..186758	185058..185459	186200..186601	192155..192556	189140..189541	186670..187071	184482..184883	184484..184885	499068..499472	494139..494543	NA	NA	186931..187332
BofA	41225..41485	35434..35694	35639..35899	41316..41576	35215..35475	36582..36842	41240..41500	41240..41500	3641493..3641741	3485449..3485697	NA	NA	36650..36910
CphA	131661..132659	149134..151755	149596..152217	155502..158123	152793..155414	150519..153140	148515..151136	148517..151138	470168..472786	465274..467892	NA	NA	150782..153403
DnaA	1..1347	14-1351	157..1494	38..1375	164..1510	1..1338	1..1347	1-1347	175..1545	127..1497	1..1338	2771582..2772934	NA
DnaN	1610..2713	1614-2717	1757..2860	1638..2741	1773..2876	1601..2704	1610..2713	1610..2713	1806..2906	1758..2858	1601..2704	2770223..2771326	1601..2704
GyrA	6342..8852	6347-8842	6499..9039	6370..8955	6506..8986	6343..8838	6342..8867	6342-8867	6549..9098	6501..9050	NA	2764060..2766609	6343..8802
GyrB	4403..6316	4408-6321	4560..6473	4431..6344	4567..6480	4404..6317	4403..6316	4403-6316	4619..6526	4571..6478	4394..6307	2766631..2768538	4404..6317
IspE	158427..159269	151848..152690	152310..153152	158216..159058	155507..156349	153233..154075	151229..152071	151231..152073	472967..473809	468072..468914	NA	NA	153496..154338
IspF	98426..98893	92210..92677	92315..92782	98083..98550	92333..92800	93120..93587	91229..91696	91231..91698	357965..358438	351409..351882	NA	2664831..2665304	93385..93852

ksgA	111827..112663	105618..106445	105725..106552	111623..112450	109255..110082	106660..107487	104639..105466	104641..105468	129582..130433	119905..120756	NA	2619392..2620237	NA
MetG	105378..107315	99161..101098	99267..101204	105035..106972	102797..104734	100072..102009	98181..100118	98183..100120	124646..126583	114812..116749	99037..100974	2629554..2631494	100337..102274
MnaA	183767..184924	177145..178302	178286..179443	184242..185399	181227..182384	178757..179914	176569..177726	176571..177728	489360..490508	484596..485744	NA	NA	179019..180176
NrdD	122969..125110	116731..118872	116913..119054	122752..124893	120449..122590	117774..119915	115772..117913	115774..117915	138621..140726	128841..130946	NA	NA	NA
NrdG	125127..125642	118889..119404	119071..119586	124910..125425	122607..123122	119932..120456	117930..118445	117932..118447	147033..147542	137327..137836	NA	NA	120198..120722
PrfA	176432..177508	169804..170880	170319..171395	176406..177482	173334..174410	171424..172500	169234..170310	169236..170312	484556..485638	479796..480878	170061..171137	2478221..2479294	171686..172762
PyrG	169487..171088	163040..164641	163374..164975	169409..171010	166570..168171	164427..166028	162289..163890	162291..163892	476625..478226	471842..473443	163114..164715	2484517..2486082	164690..166291
RecF	2998..4101	3012-4106	3164..4258	3035..4129	3171..4265	3008..4102	3007..4101	3007-4101	3223..4308	3175..4260	NA	2768847..2769929	3008..4102
RecR	40159..40755	34368-34964	34573..35169	40250..40846	34149..34745	35516..36112	40174..40770	40174-40770	3642272..3642868	3486398..3486994	NA	2739192..2739797	35584..36180
Rho	172052..173497	165424..166869	165939..167384	172026..173471	168954..170399	167044..168489	164854..166299	164856..166301	478609..480234	473822..475462	165681..167126	2482325..2483731	167306..168751
RpiB	180520..180957	173891-174328	174407..174844	180493..180930	177423..177860	175511..175948	173322..173759	173324-173761	487499..487948	482739..483188	NA	2474752..2475201	175773..176210
RpmE	173537..173755	166909..167127	167424..167642	173511..173729	170439..170657	168529..168747	166339..166557	166341..166559	481522..481731	476759..476968	NA	NA	168791..169009
SerS	22071..23622	626050-627327	22539..23819	629175..630452	609017..610294	22267..23547	576859..578136	597295-598572	17472..17685	17662..18939	NA	2749181..2750461	NA
SpmA	101129..101704	94913..95488	95018..95593	**100786..101361**	95038..95613	95823..96398	93932..94507	93934..94509	109950..110525	99972..100547	NA	2641491..2642066	NA
TagO	182719..183753	176097..177131	177238..178272	183194..184228	180179..181213	177709..178743	175521..176555	175523..176557	2932031..2932981	(2736344..2737294	NA	NA	177971..179005
TdK	173982..174557	167354..167929	167869..168444	173956..174531	170884..171459	168974..169549	166784..167359	166786..167361	481909..482496	477146..477733	NA	NA	169236..169811
Upp	180987..181616	174358..174987	174874..175503	180960..181589	177890..178519	175978..176607	173789..174418	173791..174420	487994..488623	483234..483863	NA	2474080..2474709	176240..176869

http://www.ncbi.nlm.nih.gov/genomes/MICROBES/microbial_taxtree.html

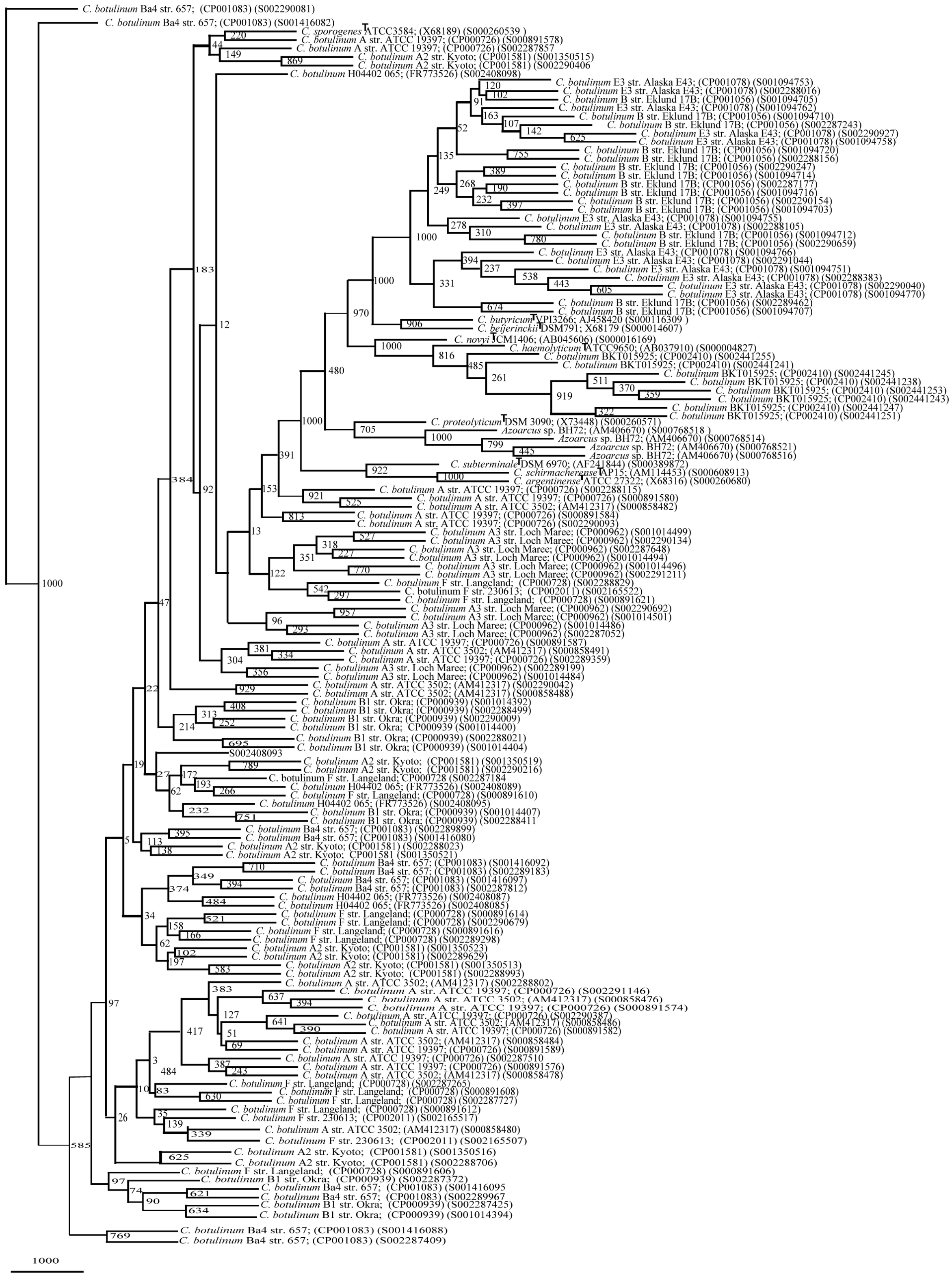

Fig S1. Phylogenetic tree of 130 rrs of thirteen completely sequenced genomes of *Clostridium botulinum* strains and their associates. File contains *rrs* sequences of *Clostridium* species: *C. argentinense* and *C. butyricum* (as neurotoxin producers) and *C. beijerinckii*, *C. haemolyticum*, *C. novyi*, *C. proteolyticus*, *C. chirmacherense*, *C. sporogenes* and *C. subterminale* (as non-producers of neurotoxin). *C. botulinum* strains ATCC 19397 and Hall had exactly similar nucleotide sequences in all the copies of rrs (Table S2), hence the figure carries the sequences of ATCC 19397 only. The tree was rooted using *Azoarcus* sp. BH72.