Novel ciliate genetic code variants including the reassignment of all three stop codons to sense codons in *C. magnum*

Supplementary Information

Supplementary Table S1. Absolute substitution count of each amino acid identified at TAA per species with stop codon count.

TAA	A	C	D	E	F	G	H	I	K	L	M	N	P	Q	R	S	T	V	W	Y
Anophryoides haemophila - 11,848	17285	1623	14847	29652	3756	9422	8341	6750	27678	11975	5235	13548	7381	78550	16157	16791	13325	8539	1165	4229
Aristerostoma sp.ATCC - 5,046	13123	1543	12463	23334	3959	8237	5250	5298	20950	9559	4423	9491	7068	45021	16060	14601	10702	7279	844	4098
Blepharisma japonicum - 233	23	8	11	13	40	25	8	19	23	29	35	10	26	64	23	31	24	18	4	9
Campanella umbellaria - 45,644	125319	8556	253397	953229	16068	75543	35122	31327	145035	52789	20064	80293	61883	159688	87219	102955	80266	51116	6547	22015
Carchesium polypinum - 7,877	22184	1490	35350	120401	3498	11458	6815	6005	27493	12475	5314	14060	9067	59827	17898	19622	14183	9820	1249	4395
Climacostomum virens - 1,379	126	14	78	211	67	99	43	48	200	164	45	58	78	121	133	96	82	96	100	235
Colpoda aspera - 371	48	11	34	40	14	63	16	40	61	48	8	30	30	79	47	65	43	28	8	35
Condylostoma magnum - 33,173	129735	17439	114984	195753	39815	94124	48865	66341	168526	112746	37029	84291	65585	321755	120915	112301	92649	80416	10700	38980
Euplotes focardii - 468	378	64	376	803	60	189	113	292	629	434	151	349	142	1540	459	297	325	285	59	105
Euplotes crassus - 326	61	41	46	90	31	88	120	73	133	227	28	79	47	56	46	88	81	59	2	39
Fabrea salina - 228	27	11	20	23	11	23	4	18	39	22	22	9	13	16	19	27	13	91	3	43
Favella ehrenbergii - 2,902	6162	989	6488	11126	1474	4776	2452	2411	9653	4595	1956	4963	3080	20289	7085	7528	4840	3789	352	1541
Litonotus pictus - 85	12	3	6	9	1	9	1	5	10	6	0	3	1	10	3	15	6	11	0	1
Mesodinium pulex - 3,697	8702	1594	5235	6442	23596	4065	6376	7983	6165	14576	4165	4624	3528	4186	5921	5668	6191	8972	5317	65512
Myrionecta rubra - 8,430	18607	5051	12104	16825	63504	12470	20626	17073	19537	35174	9677	14267	9640	11678	21005	18202	20469	19300	18360	176999
Paralembus digitiformis - 19,177	23613	2537	19315	39125	5351	11900	9585	8444	37384	18148	7532	18456	10517	130410	24578	23132	18050	11978	1513	6679
Paramecium bursaria - 26,330	59923	6056	55988	104888	13215	34555	24457	21491	87435	42138	17199	45600	28873	240271	58795	55901	42763	30595	3910	16049
Platyophrya macrostoma - 207	36	0	12	22	10	18	2	15	19	18	3	10	22	13	8	24	15	17	1	71
Protocruzia adherens - 305	29	2	18	43	13	21	4	7	16	31	8	13	19	29	12	30	19	24	1	8
Pseudokeronopsis sp.OXSA - 5,693	14102	1862	12243	22637	3831	9765	5782	5775	21382	9837	4143	10715	6416	51836	15031	16047	11362	7860	847	3739
Strombidinopsis acuminata - 10,655	34974	3366	32823	60687	8827	21871	14908	12762	47755	22815	9320	25315	17093	103384	34872	34253	24819	18002	2612	10688
Strombidium inclinatum - 3,196	7386	1073	6965	13826	1754	4242	3047	3472	12432	5730	2081	5436	5000	24633	8812	7980	5513	4256	544	2290
Tiarina fusus - 3,717	7524	576	5546	11241	2017	4382	3305	2611	9715	5376	2323	5633	3332	42556	6094	7307	5256	3332	332	2264
Uronema sp.Bbcil - 24,394	32282	3713	29512	56524	7051	18880	15388	11419	50014	23146	10138	25652	15363	151845	31681	33949	25591	16555	2007	8659

Supplementary Table S2. Absolute substitution count of each amino acid identified at TAG per species with stop codon count.

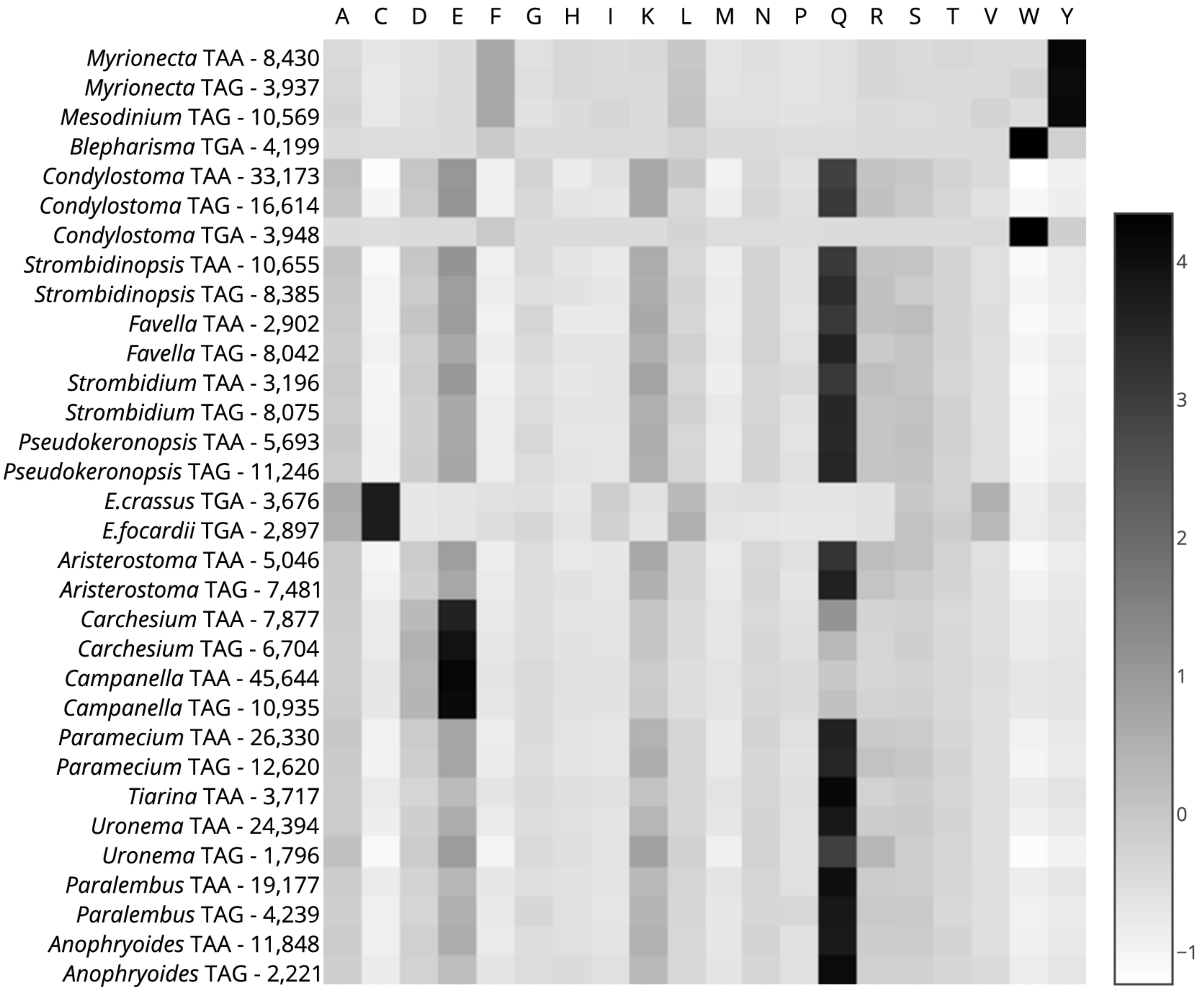
TAG	**A**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**K**	**L**	**M**	**N**	**P**	**Q**	**R**	**S**	**T**	**V**	**W**	**Y**
Anophryoides haemophila - 2,221	2539	353	2275	3883	888	1529	1624	1170	4310	1888	835	2005	1107	14511	2460	2472	2036	1571	167	822
Aristerostoma sp.ATCC - 7,481	20899	1861	18499	36297	5364	11993	9397	7956	33140	15459	6618	15656	8724	87407	23823	20333	16235	11340	1388	5877
Blepharisma japonicum - 114	24	2	12	32	61	28	3	24	20	44	5	14	37	17	15	67	32	26	68	14
Campanella umbellaria - 10,935	30950	1919	57310	204178	3897	16439	8392	7555	34863	13313	4977	20656	15137	44439	24386	26579	19396	11594	1625	5820
Carchesium polypinum - 6,704	16139	1269	30665	95046	3366	8828	5582	5099	20746	8965	3653	12160	7458	27385	12644	15668	10742	7907	991	4096
Climacostomum virens - 1,108	205	44	136	109	74	124	37	78	273	249	38	111	105	161	279	114	117	170	16	43
Colpoda aspera - 140	22	2	7	10	9	33	0	6	349	10	1	5	5	33	28	15	11	10	1	10
Condylostoma magnum - 16,614	63664	6129	59307	115776	13578	42131	26769	23566	95483	40945	16398	44399	33860	202183	67984	59092	45188	30744	4315	14610
Euplotes focardii - 341	560	47	341	746	92	191	296	167	854	355	105	325	499	2326	586	387	366	279	52	96
Euplotes crassus - 182	56	48	163	91	164	38	84	28	98	126	48	65	41	31	57	92	37	42	3	52
Fabrea salina - 95	19	1	14	16	36	13	2	7	12	12	2	11	17	1	9	4	8	4	13	11
Favella ehrenbergii - 8,042	16103	2448	15723	28672	4155	11363	7896	7413	26071	15081	4903	14266	9243	67165	17176	19522	13577	9858	958	5428
Litonotus pictus - 25	1	0	3	1	1	1	0	2	1	4	0	2	0	2	0	1	0	0	0	0
Mesodinium pulex - 10,569	26736	5951	13729	20740	88148	12426	24730	24651	20385	49318	14124	14106	10300	13837	20227	19279	17522	28257	20854	261373
Myrionecta rubra - 3,937	10874	2518	6201	8702	31984	7693	10622	9312	10014	18738	5309	6934	4602	6823	10972	9961	8856	9294	12629	85771
Paralembus digitiformis - 4,239	5071	540	4177	9130	1475	4061	2487	1949	8469	4082	1723	4135	3882	25255	5795	5803	3735	2936	308	1222
Paramecium bursaria - 12,620	32160	3202	29047	56531	8685	19135	13801	12618	50903	23515	10474	23591	14829	127059	37341	33071	25248	17587	2094	9350
Platyophrya macrostoma - 337	49	6	15	43	14	37	6	26	39	30	5	15	28	16	16	31	59	25	1	76
Protocruzia adherens - 142	12	5	8	14	5	5	3	3	6	7	0	8	12	20	16	8	7	1	2	2
Pseudokeronopsis sp.OXSA - 11,246	27802	3758	26357	49143	7434	17183	13062	11644	44391	21878	8681	22624	14759	111534	31085	32888	23024	15970	1627	8209
Strombidinopsis acuminata - 8,385	26663	3064	24233	48008	6471	14677	12562	9932	39981	20444	7648	20249	12635	95555	30313	24745	20469	13731	1993	7038
Strombidium inclinatum - 8,075	18461	2819	16979	30689	5667	12044	9305	9017	28877	16428	7687	14931	9808	69855	21316	21965	16267	11101	1873	6312
Tiarina fusus - 1,079	4133	273	2859	6209	616	1937	1582	1222	4633	2132	1317	2329	1264	15853	3732	2666	2416	1796	346	1057
Uronema sp.Bbcil - 1,796	4191	486	3428	6556	806	2565	2206	1800	6206	3140	1094	3012	2050	11551	4909	3336	2915	2223	178	997

Supplementary Table S3. Absolute substitution count of each amino acid identified at TGA per species with stop codon count.

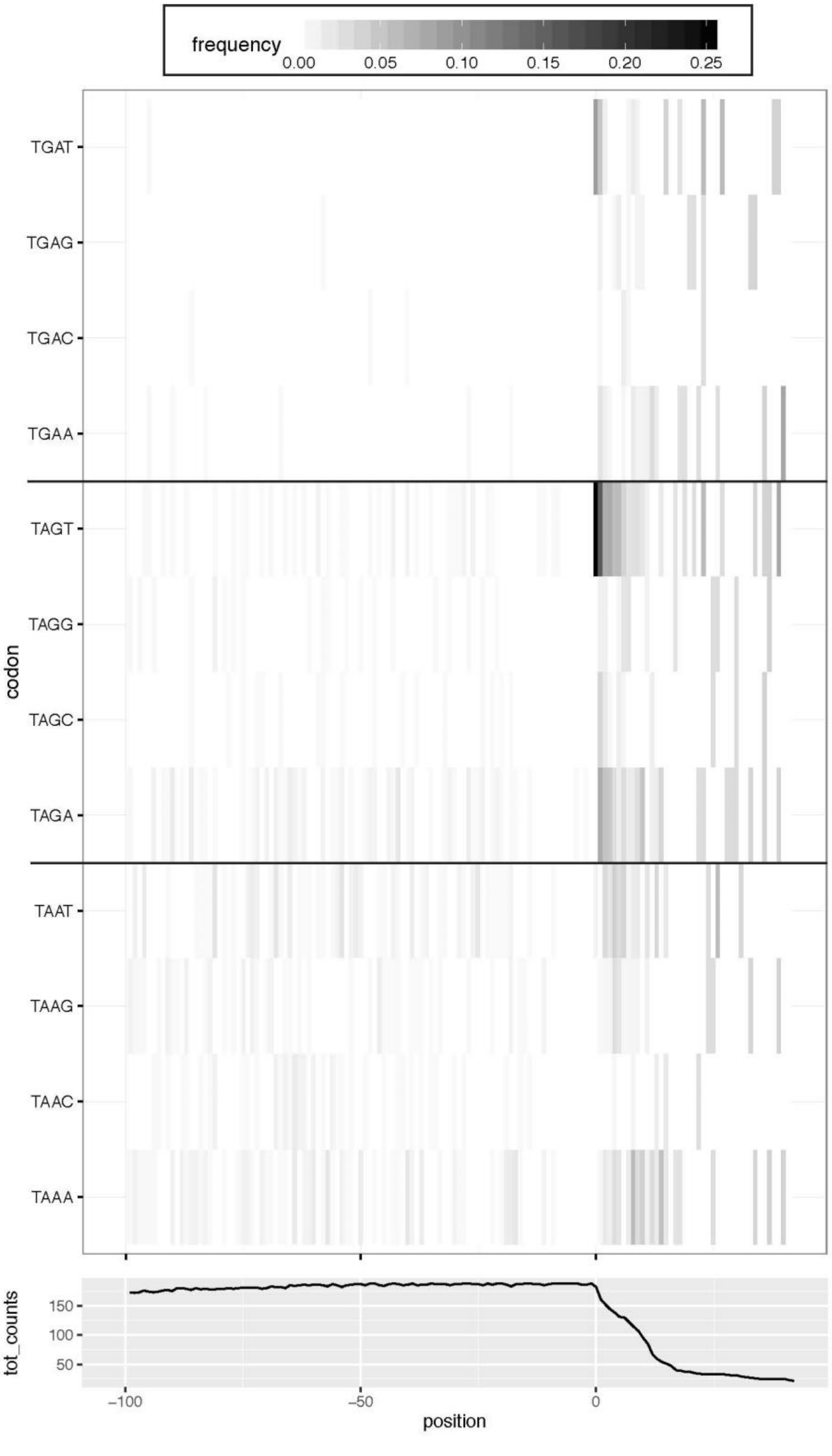
TGA	**A**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**K**	**L**	**M**	**N**	**P**	**Q**	**R**	**S**	**T**	**V**	**W**	**Y**
Anophryoides haemophila - 67	3	2	4	14	4	5	1	3	13	10	2	5	2	12	2	3	4	2	0	1
Aristerostoma sp.ATCC - 167	15	5	14	23	4	21	2	11	12	13	4	11	17	25	7	12	6	8	2	1
Blepharisma japonicum - 4199	3875	1845	2508	3342	17553	2911	2777	5262	4049	10711	2852	2527	2188	2242	4169	3616	3060	4304	153638	12857
Campanella umbellaria - 722	100	36	55	63	53	124	49	83	59	171	116	29	69	44	104	80	30	63	86	35
Carchesium polypinum - 215	21	45	22	6	41	18	12	16	20	23	10	3	17	9	5	19	17	46	83	24
Climacostomum virens - 1,664	82	64	120	53	43	177	10	146	38	143	57	50	48	43	227	55	53	128	7	15
Colpoda aspera - 255	27	4	15	28	40	49	16	32	27	47	6	9	40	26	27	53	10	29	77	35
Condylostoma magnum - 3,948	2602	1204	1944	2840	13556	2114	1943	3261	2544	6788	1571	1709	1234	2103	2938	2527	2790	3042	110685	9520
Euplotes focardii - 2,897	13950	39423	2119	2980	4190	5933	2589	6805	2806	14003	2711	2389	2058	2365	3125	10122	7705	11982	1021	2997
Euplotes crassus - 3676	19172	51261	2229	3306	3849	4747	2329	9383	4395	14913	4512	4587	2976	3016	3804	11203	8395	17637	650	3931
Fabrea salina - 55	15	235	3	1	0	64	1	3	1	6	0	3	1	2	12	2	4	14	0	1
Favella ehrenbergii - 239	16	42	21	31	5	23	7	8	30	22	5	22	9	17	32	27	18	25	2	7
Litonotus pictus - 24	1	5	1	2	0	1	0	1	3	1	1	0	1	1	0	0	1	0	0	0
Mesodinium pulex - 190	16	279	18	38	2	11	3	7	17	26	4	13	23	7	12	20	6	23	13	5
Myrionecta rubra - 100	14	8	7	16	7	25	3	2	11	10	0	5	14	5	21	13	7	9	2	3
Paralembus digitiformis - 182	22	10	14	29	6	34	2	39	25	26	8	21	7	20	39	21	12	11	13	3
Paramecium bursaria - 394	48	9	21	35	20	58	77	33	152	149	15	22	45	63	34	36	42	44	87	95
Platyophrya macrostoma - 142	13	5	10	24	9	34	7	12	23	8	2	5	10	7	7	11	51	8	1	4
Protocruzia adherens - 157	58	2	14	94	11	11	9	9	70	12	3	13	12	33	72	10	24	16	3	7
Pseudokeronopsis sp.OXSA - 212	114	55	174	219	119	121	121	65	716	101	33	202	18	177	1385	292	144	127	74	191
Strombidinopsis acuminata - 274	77	233	81	98	34	59	71	81	122	195	19	88	29	72	222	92	72	119	23	19
Strombidium inclinatum - 77	2	1	16	8	4	5	5	3	22	3	7	8	9	12	17	5	5	3	0	2
Tiarina fusus - 193	15	232	65	18	7	354	4	1	16	17	0	6	20	12	28	15	18	12	2	9
Uronema sp.Bbcil - 38	2	58	1	3	0	1	0	1	5	4	1	0	3	2	1	2	3	1	1	0

Supplementary Table S4. Absolute substitution count of each amino acid identified at 95% conservation threshold with a minimum of 100 homologous sequences

	A	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**K**	**L**	**M**	**N**	**P**	**Q**	**R**	**S**	**T**	**V**	**W**	**Y**
Myrionecta-TAA-115	0	0	141	0	661	0	1160	0	0	0	0	0	0	0	119	0	0	0	184	17807
Myrionecta-TAG-50	0	0	0	0	163	0	0	0	0	0	0	0	0	0	243	0	0	0	0	8162
Mesodinium-TAA-31	0	0	0	0	130	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5093
Mesodinium-TAG-143	0	0	0	123	1633	132	0	0	0	0	0	0	0	0	196	0	0	0	0	22659
Blepharisma-TGA-199	0	0	0	0	217	0	0	0	0	0	0	0	0	0	0	0	0	0	37332	0
Condylostoma-TAA-231	439	0	107	707	0	312	239	0	320	0	0	154	107	38328	789	0	0	0	0	0
Condylostoma-TAG-187	0	0	227	964	0	739	307	0	384	239	104	382	109	32006	305	248	0	0	0	109
Condylostoma-TGA-165	0	0	0	0	478	106	0	0	0	0	0	0	0	0	0	0	0	0	29758	0
Strombidinopsis-TAA-102	0	0	0	210	0	223	0	0	103	0	0	0	110	17865	0	0	0	0	0	0
Strombidinopsis-TAG-53	0	0	0	218	0	221	0	0	0	0	0	0	0	9684	0	0	0	0	0	0
Favella-TAA-19	0	0	0	0	0	0	0	0	0	0	0	0	0	3141	0	0	0	0	0	0
Favella-TAG-55	0	0	0	142	0	231	208	0	172	0	0	0	219	9021	155	0	0	0	0	0
Strombidium-TAA-23	0	0	0	180	0	0	0	0	0	0	0	0	0	4027	246	0	0	0	0	0
Strombidium-TAG-71	0	0	108	250	168	0	383	0	0	0	0	0	104	13088	0	0	242	0	0	0
Pseudokeronopsis-TAA-50	0	0	0	0	97	240	209	139	104	109	0	0	0	8436	0	0	0	0	0	0
Pseudokeronopsis-TAG-91	0	0	160	1209	0	0	453	0	110	0	0	154	414	13678	0	0	0	0	0	0
E.crassus-TGA-49	0	7930	0	175	0	0	0	0	0	204	0	0	0	0	0	0	0	0	0	495
E.focardii-TGA-57	217	8345	0	140	0	467	0	0	0	817	0	0	0	0	136	0	0	0	223	0
Aristerostoma-TAA-31	0	0	0	151	0	0	0	0	165	0	0	0	0	4906	223	0	0	0	0	0
Aristerostoma-TAG-58	0	0	0	119	0	234	0	0	0	0	0	0	0	9725	226	112	0	190	0	0
Carchesium-TAA-118	0	0	134	14629	0	0	0	0	0	0	0	0	0	550	0	101	0	0	0	0
Carchesium-TAG-64	0	0	289	8510	0	104	0	0	497	0	0	0	0	1255	0	0	0	0	0	0
Campanella-TAA-955	0	0	4112	171481	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Campanella-TAG-204	225	0	469	34497	0	200	0	0	0	0	0	0	0	3886	0	244	0	0	0	0
Paramecium-TAA-164	0	0	0	485	0	0	0	298	298	127	0	0	0	28294	134	558	0	0	0	0
Paramecium-TAG-91	0	0	0	114	0	0	0	0	0	138	0	0	336	18431	118	0	0	0	0	0
Tiarina-TAA-52	0	0	0	273	197	196	0	0	0	0	0	0	156	8805	0	0	0	0	0	0
Uronema-TAA-137	105	0	0	276	0	132	0	0	234	0	0	248	208	23296	0	0	0	0	0	0
Uronema-TAG-8	0	0	0	0	0	0	0	0	0	0	0	0	0	1476	0	0	0	0	0	0
Paralembus-TAA-134	0	0	0	263	0	0	0	0	591	1086	0	0	190	24192	251	0	0	0	97	0
Paralembus-TAG-31	0	0	0	198	0	586	0	0	0	0	0	392	0	4344	101	0	0	0	0	0
Anophryoides-TAA-64	0	0	0	326	0	108	0	0	216	0	0	0	0	10891	0	0	0	0	0	0
Anophryoides-TAG-20	0	0	0	0	0	0	0	0	0	0	0	0	0	3395	0	0	0	0	0	0



Supplementary Fig. S1. Identification of amino acid specifications of the reassigned codons. Each row corresponds to a single reassigned codon. The organism source of a codon, its identity and the total number of occurrences in aligned sequences are indicated on the left. The frequencies of amino acid substitutions are shown as heatmaps according to their Z-scores.



Supplementary Fig. S2. Frequency of stop quadruplets (stop codon with a nucleotide adjacent at 3’ end) relative to expected positions (zero on axis x) of translation termination in *Condylostoma*. Top panel – frequency of each out of 12 quadruplets. Bottom panel – total number of quadruplets found at corresponding location.