Supplementary Table 1

Summary statistics of 25 human-infective *Cryptosporidium spp*. genome projects including 23 whole genome sequences (WGS).

SPECIES	gp60 SUBTYPE	STANDARD ID	SOURCE	ACCESSION	WGS SIZE (bp)	N50 (Mb)	Host	Collected	Sample origin	Recent Travel
C. cuniculus	VbA37	UKCU2*	This study	PRJNA315496	9,183,765	1.806	Human	2013	United Kingdom	None
C. cuniculus	VaA31	UKCU5#	This study	PRJNA492839	Not Assembled		Human	2013	United Kingdom	None
C. hominis	IbA10G2	UKH1‡	Widmer, G.¥	CryptoDB.org	9,141,398	0	Human	2012	United Kingdom	None
C. hominis	IbA10G2	UKH3#	Hadfield *et al*[8]	PRJNA253834	9,136,308	0.060	Human	2012	United Kingdom	Turkey
C. hominis	IaA14R3	UKH4*	Hadfield *et al*[8]	PRJNA253838	9,158,297	0.167	Human	2013	United Kingdom	None
C. hominis	IbA10G2	UKH5#	Hadfield *et al*[8]	PRJNA253839	9,179,731	0.168	Human	2013	United Kingdom	None
C. hominis	IdA30	UKH6#	This study	PRJNA492838	Not Assembled		Human	2013	United Kingdom	None
C. meleagridis	IIIbA22G1	UKMEL1*	Widmer, G.¥	CryptoDB.org	8,973,224	0	Human	2013	United Kingdom	India
C. meleagridis	IIIgA23G3	UKMEL3*	This study	PRJNA315502	8,732,077	0.062	Human	2013	United Kingdom	None
C. meleagridis	IIIhA7	UKMEL4*	This study	PRJNA315503	8,811,811	0.025	Human	2013	United Kingdom	Spain
C. parvum parvum	IIaA19G1R2	UKP2*	Hadfield. *et al*[8]	PRJNA253836	9,104,817	0.034	Human	2012	United Kingdom	None
C. parvum parvum	IIaA18G2R1	UKP3*	Hadfield *et al*[8]	PRJNA253840	9,085,662	0.126	Human	2013	United Kingdom	None
C. parvum parvum	IIaA15G2R1	UKP4#	Hadfield *et al*[8]	PRJNA253843	9,001,535	0.107	Human	2012	United Kingdom	None
C. parvum parvum	IIaA15G2R1	UKP5#	Hadfield *et al*[8]	PRJNA253845	9,283,240	0.236	Human	2012	United Kingdom	None
C. parvum parvum	IIaA15G2R1	UKP6‡	Hadfield. *et al*[8]	PRJNA253846	9,112,937	0.023	Human	2013	United Kingdom	None
C. parvum parvum	IIaA17G1R1	UKP7*	Hadfield *et al*[8]	PRJNA253847	9,221,025	0.246	Human	2013	United Kingdom	None
C. parvum parvum	IIdA22G1	UKP8*	Hadfield *et al*[8]	PRJNA253848	9,203,336	0.145	Human	2013	United Kingdom	None
C. parvum anthroponosum	IIcA5G3p	UKP12*	This study	PRJNA315504	9,325,214	1.686	Human	2013	United Kingdom	Pakistan
C. parvum anthroponosum	IIcA5G3a	UKP13*	This study	PRJNA315505	9,031,205	1.876	Human	2013	United Kingdom	Pakistan
C. parvum anthroponosum	IIcA5G3a	UKP14*	This study	PRJNA315506	9,432,159	0.944	Human	2014	United Kingdom	Pakistan
C. parvum anthroponosum	IIcA5G3a	UKP15‡	This study	PRJNA315507	9,408,807	0.307	Human	2013	United Kingdom	None
C. parvum parvum	IIcA5G3j	UKP16‡	This study	PRJNA315508	9,308,724	0.240	Human	2013	United Kingdom	None
C. ubiquitum	XIIb	UKUB1*	This study	PRJNA315509	9,060,260	1.812	Human	2013	United Kingdom	None
C. ubiquitum	XIIb	UKUB2*	This study	PRJNA315510	9,069,162	0.907	Human	2012	United Kingdom	None
C. viatorum	XVaA3f	UKVIA1*	This study	PRJNA492837	11,261,626	0.112	Human	2012	United Kingdom	India

* Included in whole genome comparative genomics
‡ Included in whole genome comparative genomics and recombination analysis
# Used only for read mapping onto Iowa II in Figure 3b
¥ Tufts University School of Veterinary Medicine, Medford, Massachusetts (Unpublished genome, CryptoDB.org)

Supplementary Table 2

General Linear Model (GLM) of the pairwise genetic distance (Kxy) between *C. p. parvum* and *C. p. anthroponosum* isolates, with geographic distance as covariate crossed with species. Genetic distances of the gp60 gene between isolates were expressed as Kxy, and these were calculated with the software DnaSP 5.10.1[52]. The geographic distance between isolates (expressed in km as the crow flies) were calculated as the distance between the centre of one country or region to the centre of another using Google Maps (2017). A General Linear Model (GLM) was used to assess differences in the population genetic structure of *C. p. parvum* and *C. p. anthroponosum*. In this model, the pairwise genetic distance (Kxy) was used as the response variable, and species as fixed factor. Species was crossed with geographic distance between sampling points, which was included as a covariate in the model. This interaction term (species x distance) interrogates whether the two regression lines for both species have a similar slope.

Analysis of Variance for Kxy, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Spp	1	10.073	9.492	9.492	13.14	0.001
Km	1	18.729	25.359	25.359	35.10	0.000
spp*km	1	34.209	34.209	34.209	47.34	0.000
Error	79	57.082	57.082	0.723		
Total	82	120.093				

Supplementary Table 3

Linear Model of the pairwise genetic distance (Kxy) of *C. p. parvum* isolates versus distance.

Linear Model of Kxy versus geographic distance

Source	DF	SS	MS	F	P
Regression	1	52.560	52.560	40.63	0.000
Residual Error	26	33.634	1.294		
Total	27	86.193			

R-Sq. = 61.0%

Supplementary Table 4

Description of positively-selected (>1.0 Ka/Ks) protein-coding genes between *C. parvum parvum* UKP6 and *C. parvum anthroponosum* UKP15.

Chromosome	CryptoDB ID[41]	Nucleotide Diversity (π)[44]	Ka/Ks[51]	Length (bp)	Protein localization[50]
1	cgd1_120	0.0212	2.6169	1293	plas
	cgd1_620	0.0053	1.2919	1314	extr
	cgd1_1230	0.0019	1.5228	3195	extr
	cgd1_1400	0.0017	1.0679	2886	plas
	cgd1_1640	0.0006	1.3054	8628	plas
	cgd1_3760	0.0023	2.3799	3477	nucl
2	cgd2_390	0.0286	1.0009	465	extr
	cgd2_430	0.0099	1.1634	612	extr
	cgd2_940	0.0022	1.5374	3189	plas
	cgd2_2900	0.0118	1.5270	513	cyto
	cgd2_4060	0.0023	1.0462	2163	plas
	cgd2_4370	0.0302	3.2666	1122	extr
3	cgd3_1690	0.0061	1.1526	987	extr
	cgd3_1710	0.0112	2.5777	900	extr
	cgd3_1780	0.0224	2.9534	1653	plas
	cgd3_2080	0.0015	1.0061	3930	plas
	cgd3_3650	0.0035	1.1533	1413	cyto
	cgd3_4180	0.0028	1.1498	3603	extr
4	cgd4_3670	0.0078	1.0479	1404	nucl
	cgd4_3750	0.0052	1.8261	1938	plas
5	cgd5_20	0.0040	2.0821	2280	extr
	cgd5_50	0.0067	1.4633	1962	extr
	cgd5_580	0.0026	1.1164	3405	plas
	cgd5_2560	0.0025	1.5475	2451	nucl
6	cgd6_10	0.1302	1.1110	630	extr
	cgd6_40	0.0334	2.6311	555	extr
	cgd6_640	0.0029	1.6584	2727	cyto
	cgd6_1010	0.0028	1.3601	2145	nucl
	cgd6_3600	0.0054	1.0740	747	nucl
	cgd6_3920	0.0042	2.5510	2367	extr
	cgd6_5110	0.0041	1.0056	8109	plas
	cgd6_5270	0.1194	1.2212	480	extr
	cgd6_5410	0.0052	1.0307	2121	extr
	cgd6_5500	0.0091	1.1566	882	extr
7	cgd7_1280	0.0072	1.2224	561	extr
	cgd7_2600	0.0018	1.8389	4416	cyto
8	cgd8_40	0.0452	1.6397	2652	extr
	cgd8_60	0.0186	1.1000	600	extr
	cgd8_290	0.0043	1.3539	1155	mito
	cgd8_380	0.0045	1.2426	1578	cyto
	cgd8_520	0.0092	1.1355	1101	extr
	cgd8_1570	0.0043	1.3895	1410	mito
	cgd8_2450	0.0019	1.5568	3162	nucl
	cgd8_2550	0.0015	1.5955	3951	plas

Supplementary Table 5

Description of hypervariable (<90.0% amino acid identities) protein-coding genes between *C. parvum parvum* UKP6 (IIaA15G2R1) and *C. hominis* UKH4 (IaA14R3).

Chromosome	CryptoDB ID[41]	% AA IDs[44]	InDel Frameshift[44]	Putative Protein Function[49]	Putative Localization[50]	KaKs[51]
1	cgd1_110	83.4	FS	Secreted protein	extr	0.6075
	cgd1_120	77.2		Uncharacterized	plas	1.0254
	cgd1_130	80.8		IWS1-like protein	plas	0.6841
	cgd1_140	60.7	FS	Predicted secreted protein	extr	0.6060
	cgd1_430	56.0	FS	Uncharacterized	extr	0.4024
	cgd1_470	74.7		Mucin	nucl	0.6481
	cgd1_590	84.7	FS	Proteoglycan/mucin	extr	0.5808
	cgd1_620	88.6		Viral A-type inclusion protein	extr	0.7577
	cgd1_680	73.3		Uncharacterized	nucl	0.3329
	cgd1_900	33.3	FS	Uncharacterized	extr	0.2405
	cgd1_1030	54.9	FS	Uncharacterized	cyto	0.2691
	cgd1_1190	79.4	FS	Uncharacterized	extr	99.000
	cgd1_1320	52.0	FS	Developmental protein	extr	0.0010
	cgd1_1440	88.0	FS	Uncharacterized	cyto	0.2990
	cgd1_1510	89.5	FS	Uncharacterized	mito	0.1463
	cgd1_1650	86.8	FS	Uncharacterized	extr	0.0629
	cgd1_1710	89.2	FS	Phosphoglycerate mutase	mito	0.1439
	cgd1_3290	89.4		Carboxylesterase	plas	0.3342
	cgd1_3430	82.5	FS	Uncharacterized	extr	1.3880
	cgd1_3450	38.5	FS	Uncharacterized	nucl	0.2532
	cgd1_3590	86.2		Membrane associated protein	plas	0.3157
	cgd1_3680	36.1	FS	EGF-like domain protein	plas	0.3002
	cgd1_3850	78.5		Uncharacterized	plas	0.6967
	cgd1_3860	22.1	FS	Deoxyuridine 5'-triphosphate nucleotidohydrolase	mito	0.4853
2	cgd2_390	81.3		Mucin	extr	1.4084
	cgd2_400	82.5		Mucin	extr	0.5439
	cgd2_410	82.9		Mucin	extr	1.0176
	cgd2_420	59.5		Mucin	extr	1.4730
	cgd2_430	71.8		Mucin	extr	0.9262
	cgd2_440	78.5		Mucin	extr	4.1629
	cgd2_450	74.8		Mucin	extr	0.5573
	cgd2_840	84.0	FS	Phosphatidylinositol N-acetylglucosaminyltransferase subunit P	cyto	0.1336
	cgd2_1170	68.0	FS	Zinc finger protein ZPR1	cyto	0.0010
	cgd2_1550	83.8	FS	Origin of replication complex subunit 4	cyto	0.0496
	cgd2_1970	86.3	FS	SAM dependent methyltransferase	nucl	0.2756
	cgd2_2110	73.4	FS	Uncharacterized	plas	0.1043
	cgd2_2180	72.4	FS	Uncharacterized	extr	0.2647
	cgd2_2460	55.2	FS	Insulin growth factor-binding protein	cyto	0.1568
	cgd2_2550	87.9		Lipoprotein	plas	0.4952
	cgd2_2560	66.1	FS	Uncharacterized	extr_plas	0.8179

	cgd2_2570	88.3		Uncharacterized	extr	0.4058
	cgd2_2600	89.6		Uncharacterized	extr	0.4483
	cgd2_2650	84.9	FS	Uncharacterized	plas	0.7622
	cgd2_2900	87.7		Uncharacterized	cyto	0.9958
	cgd2_3140	62.4		Mucin	plas	0.1668
	cgd2_3270	83.1	FS	Phosphoglucomutase/phosphomannomutase family protein	E.R.	0.0661
	cgd2_3280	37.9	FS	Aminopeptidase	plas	0.3054
	cgd2_3370	64.2	FS	Proteasome regulatory subunit Rpn12 family	extr	0.2861
	cgd2_3520	83.2		IWS1 like protein	extr	0.4715
	cgd2_3530	85.7		Eukaryotic translation initiation factor	nucl	0.5232
	cgd2_3610	67.8	FS	WD domain containing protein	extr	0.0719
	cgd2_3780	81.5	FS	Mucin	cyto_nucl	0.1400
	cgd2_3820	52.8	FS	Uncharacterized	extr	0.0010
	cgd2_3970	85.8	FS	RNA recognition family protein	extr	0.1737
	cgd2_4020	89.8		Uncharacterized	extr	0.6982
	cgd2_4310	88.8	FS	Uncharacterized	nucl	0.9573
	cgd2_4370	78.0		Early endosome antigen 1	extr	1.1392
	cgd2_4380	69.9	FS	Mucin	extr	0.7805
3	cgd3_10	83.4		Anchor protein	plas	0.4945
	cgd3_170	72.9	FS	DUF947-domain-containing protein	nucl	0.4624
	cgd3_190	73.6		Mucin	plas	0.1696
	cgd3_370	39.8	FS	Uncharacterized	extr	99.000
	cgd3_630	84.6		Integral membrane protein	plas	0.3972
	Chro.30091	88.6		Proteoglycan	E.R.	0.3263
	cgd3_820	88.7		Uncharacterized	plas	0.8233
	cgd3_1073	52.6	FS	Synaptobrevin family protein	cyto	0.1464
	cgd3_1100	82.4		Nipped-B-like protein	cyto	1.0843
	cgd3_1150	70.2		Uncharacterized	extr	0.7599
	cgd3_1160	85.1		RNA polymerase-associated protein	plas	0.6439
	cgd3_1170	35.7	FS	Uncharacterized	extr	0.3388
	cgd3_1680	58.3	FS	Uncharacterized	plas	0.3542
	cgd3_1690	86.7		Uncharacterized	extr	0.5088
	cgd3_1710	85.7		Uncharacterized	extr	0.9842
	cgd3_1730	87.6		Uncharacterized	extr	1.2875
	cgd3_1740	85.0		Ubiquitin-like protein	mito	0.9198
	cgd3_1750	88.3		Inositol-phosphate phosphatase	extr	0.6631
	cgd3_1760	81.3		Uncharacterized	cyto	0.7370
	cgd3_1770	75.5		Uncharacterized	extr	0.8301
	cgd3_1780	82.2		Antigen	plas	1.1087
	Chro.30271	79.8	FS	Gaa1-like GPI transamidase component	plas	2.3317
	cgd3_2700	88.6	FS	Trafficking protein particle	extr	0.0586
	cgd3_2830	88.4	FS	Uncharacterized	mito	0.1788
	cgd3_4260	87.5		Insulinase like peptidase	plas	0.3161
	cgd3_4270	89.5		Insulinase like peptidase	plas	0.2621
	cgd3_4360	89.0		Uncharacterized	plas	0.4409

4	cgd4_10	86.3		Glutamate receptor	extr	0.6381
	cgd4_32	89.9	FS	Glycoprotein	cyto	0.3236
	cgd4_210	58.3	FS	Ubiquitin-conjugating enzyme 27	cysk	0.0010
	cgd4_770	88.1		Trichohyalin	cyto	0.1082
	cgd4_920	75.5	FS	Histidine phosphatase superfamily	plas	0.1992
	cgd4_1000	11.2		Cell wall anchor protein	nucl	99.000
	cgd4_1280	74.3	FS	Rtf2 RING-finger family protein	mito	0.5351
	cgd4_1300	58.0		Mucin	nucl	0.4326
	cgd4_2160	42.9	FS	Ribonuclease	extr_plas	0.3239
	cgd4_2450	86.6	FS	Tubulin-specific chaperone C	cyto	0.4707
	cgd4_2500	87.4	FS	Uncharacterized	extr	0.2272
	cgd4_2510	81.6	FS	Uncharacterized	extr	0.7576
	cgd4_2760	56.8	FS	Mitotic-spindle organizing protein	mito	0.3642
	cgd4_2830	87.6	FS	Mra1/NEP1 like protein	extr	0.2191
	cgd4_3060	60.8	FS	Uncharacterized	cyto	0.9623
	cgd4_3350	63.8	FS	Mob1/phocein family protein	extr	0.5179
	cgd4_3520	88.1		Proteophosphoglycan	nucl	0.2599
	cgd4_3550	85.4		Kazal-type serine protease inhibitor domain-containing protein	extr	0.2612
	cgd4_3630	65.7		Cross-beta structure silk protein 1	nucl	0.6389
	cgd4_3640	77.0	FS	Uncharacterized	cyto	0.3190
	cgd4_3650	55.9	FS	Uncharacterized	extr	1.3451
	cgd4_3660	37.8	FS	Uncharacterized	cyto	1.1696
	cgd4_3670	75.1		Collagen-like protein	nucl	0.4561
	cgd4_3680	87.8		Uncharacterized	cyto	0.5627
	cgd4_3690	70.2		Glycine-rich cell wall structural protein	plas	0.4939
	cgd4_3930	74.4	FS	Exosome complex component	mito	0.0233
	cgd4_3970	82.8		GPI-anchored protein	plas	0.2715
	cgd4_4070	30.8	FS	Uncharacterized	extr	0.6830
	cgd4_4210	56.5	FS	Antigen	plas	0.1748
	cgd4_4253	80.1	FS	Uncharacterized	cyto	0.3880
	cgd4_4390	70.6	FS	Uncharacterized	mito	0.0556
	cgd4_4470	88.4		Dentin sialophosphoprotein	plas	0.3753
	cgd4_4480	89.0		Uncharacterized	plas	0.4933
	cgd4_4500	73.4	FS	Proteophosphoglycan	nucl	0.7551
5	cgd6_5500	64.3	FS	Uncharacterized	cyto	0.2680
	cgd5_10	87.5		S-antigen protein	extr	0.6797
	cgd5_20	89.0		GPI-anchored adhesin-like	extr	0.5210
	Cgd5_40	71.8		Erythrocyte membrane protein	extr_plas	0.9587
	cgd5_50	82.1		Uncharacterized	extr_plas	1.0310
	cgd5_130	89.4		Ferlin like type II membrane associated protein	plas	0.0691
	cgd5_450	89.0		Putative RING zinc finger	nucl	0.1065
	cgd5_1090	87.9	FS	Uncharacterized	extr	0.6167
	cgd5_1580	84.8	FS	Uncharacterized	cyto	0.6614
	cgd5_1940	89.6		Viral A-type inclusion protein	nucl	0.3923
	cgd5_2180	81.8		Mucin 17-like protein	nucl	0.1433

	cgd5_2960	85.1	FS	Putative U5 small nuclear ribonucleoprotein 200 kDa helicase	plas	0.2925
	cgd5_3190	86.2	FS	Protein kinase domain protein	cyto	0.1631
	cgd5_3440	56.1	FS	Uncharacterized	extr	0.5517
	cgd5_3490	89.9		Biotin-protein ligase	extr	0.4294
	Chro.50010	44.5	FS	Proteophosphoglycan	plas	0.3678
6	cgd6_10	46.6		Proteophosphoglycan	extr	0.2680
	cgd6_40	72.5		Antigen	extr	0.6395
	cgd6_50	36.6	FS	Uncharacterized	extr	0.9261
	cgd6_60	88.1		Protease	nucl	0.3409
	cgd6_170	82.6	FS	Synaptobrevin-like protein	cyto	0.3476
	cgd6_260	57.7	FS	Diacylglycerol acyltransferase	plas	0.0922
	cgd6_340	63.1	FS	Uncharacterized	extr	0.6149
	cgd6_780	86.9	FS	Sporozoite cysteine-rich protein	plas	0.2063
	cgd6_920	48.8	FS	26S proteasome regulatory subunit 8	cyto	99.000
	cgd6_960	74.8	FS	Cysteinyl-tRNA synthetase	cyto_nucl	0.0802
	cgd6_1080	69.2		Glycoprotein	extr	0.5341
	cgd6_1170	89.7		Uncharacterized	cyto	0.5669
	cgd6_1620	89.2	FS	Uncharacterized	cyto	0.4667
	cgd6_2130	80.8	FS	RNA methyltransferase	plas	0.4113
	cgd6_2140	48.9	FS	Ion channel protein	cyto	0.2336
	cgd6_2270	47.6	FS	Membrane-associated protein	plas	0.1257
	cgd6_2500	77.8	FS	Rhoptry protein	plas	0.1178
	cgd6_2660	75.2	FS	DNA repair helicase	nucl	0.1060
	cgd6_2800	86.4	FS	Ras-related GTP-binding protein	cysk	0.1877
	cgd6_3050	81.8		Mucin	extr	0.7440
	cgd6_3360	71.8	FS	FYVE and coiled-coil domain-containing protein	extr	0.1301
	cgd6_3770	88.6	FS	Insulin-degrading enzyme	cyto	0.0661
	cgd6_3930	81.5		Glycoprotein	nucl	0.5344
	cgd6_3940	71.2		Glycoprotein	mito	1.8627
	cgd6_4100	45.5	FS	Uncharacterized	extr	0.2559
	cgd6_4230	89.0		Cement protein 3B	extr	0.5390
	cgd6_4670	56.2	FS	Splicing factor 3A subunit 3	cyto	0.1523
	cgd6_4740	84.1		Transmembrane protein 64	plas	0.4542
	cgd6_4980	46.5	FS	Uncharacterized	plas	99.000
	cgd6_5110	86.4	FS	Reticulocyte binding protein	plas	0.3134
	cgd6_5270	88.8		Uncharacterized	extr	0.3736
	cgd6_5400	70.5		Mucin	extr	0.1940
	cgd6_5410	85.7		Mucin	extr	0.3327
	cgd6_5430	86.6		GPI-anchored adhesin-like protein	plas	0.6845
7	cgd5_4530	23.2	FS	Uncharacterized	E.R._mito	0.1754
	cgd7_10	81.1	FS	Binding protein	plas	0.5083
	cgd7_1210	88.8		Integral membrane protein	extr	1.3153
	cgd7_1280	76.7		Glycoprotein	extr	0.6014
	cgd7_1370	89.1		Uncharacterized	extr	0.6406
	cgd7_1870	87.6	FS	Uncharacterized	extr	1.2958

	cgd7_2120	67.9	FS	Uncharacterized	extr	0.3770
	cgd7_2350	48.6	FS	Uncharacterized	plas	1.2958
	cgd7_2870	83.0	FS	Titin	nucl	0.5227
	cgd7_3420	30.7	FS	Uncharacterized	mito	0.1356
	cgd7_3440	68.2	FS	Uncharacterized	cyto	99.000
	cgd7_3800	82.1	FS	Uncharacterized	extr	0.1719
	cgd7_4020	88.8		Mucin	plas	0.0632
	cgd7_4260	89.4	FS	Uncharacterized	nucl	0.1824
	cgd7_4300	51.7	FS	Zinc finger, C2H2 type domain	cyto	0.1198
	cgd7_4310	82.7	FS	Cysteine-rich secretory protein	extr	0.1025
	cgd7_4430	83.3		Glycosyl transferase family	extr	0.6875
	cgd7_4500	81.9		Proteoglycan/glycoprotein	extr	0.6241
	cgd7_5400	85.1	FS	Uncharacterized	extr	0.0897
	cgd7_5510	89.1		Chromosome partition protein Smc	extr	0.9346
	cgd7_5520	82.5		Glycoprotein	mito	0.2623
8	cgd8_10	86.8		Uncharacterized	cyto	0.4493
	cgd8_20	86.8		Uncharacterized	plas	0.4105
	cgd8_30	87.3	FS	Uncharacterized	nucl	0.4544
	cgd8_40	79.3		Uncharacterized	plas	0.9231
	cgd8_50	89.6		Uncharacterized	plas	0.3737
	cgd8_60	71.3	FS	Uncharacterized	extr	0.7822
	cgd8_520	83.0		Histone H5	extr	0.6743
	cgd8_660	72.4	FS	Mucin	E.R.	0.6837
	cgd8_700	87.0		Mucin	plas	0.3744
	cgd8_1020	74.3	FS	N terminus of Rad21/Rec8 like protein	cyto	0.1515
	cgd8_1160	89.8		Mucin	plas	0.1597
	cgd8_1220	80.9	FS	Mucin	cyto	0.2073
	cgd8_1410	75.7	FS	DNA primase large subunit	cyto	0.0386
	cgd8_1570	71.6	FS	CCCH like finger domain nucleoporin	mito	0.2653
	cgd8_1750	89.7		Uncharacterized	extr	0.2194
	cgd8_1770	89.7		Proteophosphoglycan	plas	0.2210
	cgd8_1820	57.1	FS	Uncharacterized	extr	1.8643
	cgd8_2140	52.9	FS	Uncharacterized	plas	0.2520
	cgd8_2160	84.6		Poly(ADP-ribose) glycohydrolase	plas	0.6824
	cgd8_2220	85.8	FS	Male gamete fusion factor family	nucl	0.1903
	cgd8_2240	84.1	FS	Histidine phosphatase superfamily	cyto	4.2459
	cgd8_2590	58.3	FS	Uncharacterized	plas	0.0862
	cgd8_2800	63.5	FS	Mucin	plas	0.2877
	cgd8_3120	86.6	FS	Uncharacterized	extr	0.6230
	cgd8_3200	86.7	FS	Ubiquitin carboxyl-terminal hydrolase	cyto	0.2506
	cgd8_3540	89.9	FS	Uncharacterized	plas	0.5655
	cgd8_3550	38.5	FS	Uncharacterized	cyto	0.1517
	cgd8_3550	40.5	FS	Uncharacterized	mito	1.8383
	cgd8_3650	72.2	FS	Trafficking protein particle complex	cysk	0.3312
	cgd8_3670	85.3	FS	Uncharacterized	mito	0.1203

	cgd8_4190	73.7		Mucin	cyto	0.3961
	cgd8_4480	88.8		Type VI secretion system Vgr family	nucl	0.1217
	cgd8_4550	76.6	FS	Uncharacterized	cyto	0.3963
	cgd8_4740	66.2	FS	Phosphopantetheinyl transferase	cyto	0.1954
	cgd8_4820	23.6	FS	Transcription initiation factor IID	cyto	0.4455
	cgd8_4860	89.8	FS	Antigen	extr	0.3749
	cgd8_5050	70.7	FS	Palmitoyltransferase	plas	0.4310
	cgd8_5290	89.1		Glycoprotein	plas	0.3766
	cgd8_5360	26.9	FS	Glycoprotein	extr	0.7168
	cgd8_5370	64.6		Uncharacterized	extr	1.9001
	cgd8_5380	75.0		Rap guanine nucleotide exchange factor	extr	0.9647
	cgd8_5390	88.4		Uncharacterized	extr	1.2072
	cgd8_5420	24.7	FS	Uncharacterized	extr	0.5848

Supplementary Table 6

Description of hypervariable (<90.0% amino acid identities) protein-coding genes between *C. parvum parvum* UKP6 (IIaA15G2R1) and *C. parvum anthroponosum* UKP15 (IIcA5G3a).

Chromosome	CryptoDB ID[41]	% AA IDs[44]	InDel Frameshift[44]	Putative Protein Function[49]	Putative Localization[50]	KaKs[51]
1	cgd1_150	25.8	FS	Autophagy-related protein 11	plas	0.3287
	cgd1_470	80.3		Mucin	nucl	0.6767
2	cgd2_3140	85.4		Mucin	plas	0.1458
	cgd2_3530	87.8		Eukaryotic translation initiation factor	nucl	2.2698
3	cgd3_370	26.0	FS	Uncharacterized	extr	0.0010
	cgd3_1150	89.8		Uncharacterized	extr	0.7918
	cgd3_1160	38.2	FS	RNA polymerase-associated protein	plas	1.1871
	cgd3_1170	82.1	FS	Uncharacterized	extr	0.6709
	cgd3_1680	65.3	FS	Uncharacterized	plas	0.0010
4	cgd4_1280	74.3	FS	Rtf2 RING-finger family protein	mito	1.3363
	cgd4_1300	79.8		Mucin	nucl	0.5568
	cgd4_3690	44.2	FS	Glycine-rich cell wall structural protein	plas	1.2300
	cgd4_3660	40.1	FS	Uncharacterized	cyto	0.7244
	cgd4_3060	36.9	FS	Uncharacterized	cyto	0.0010
	cgd4_2830	89.7	FS	Mra1/NEP1 like protein	extr	0.0010
	cgd4_4070	31.2	FS	Uncharacterized	extr	1.6923
	cgd4_4390	71.3	FS	Uncharacterized	mito	0.2986
	cgd4_4470	29.4	FS	Dentin sialophosphoprotein	plas	0.3275
	cgd4_4500	67.8	FS	Proteophosphoglycan	nucl	0.8056
5	Cgd5_40	81.9		Erythrocyte membrane protein	extr_plas	0.4943
	cgd5_1670	84.4	FS	Lysine-rich arabinogalactan protein	mito	0.0010
	cgd5_2180	86.8		Mucin 17-like protein	nucl	0.2466
	Chro.50010	77.3		Proteophosphoglycan	plas	0.3196
6	cgd6_10	66.8		Proteophosphoglycan	extr	1.1110
	cgd6_40	89.1		Antigen	extr	0.7261
	cgd6_50	44.5	FS	Uncharacterized	extr	99.0000
	cgd6_170	89.9	FS	Synaptobrevin-like protein	cyto	99.0000
	cgd6_250	83.4	FS	TatD-like deoxyribonuclease	cyto	0.4839
	cgd6_340	60.9	FS	Uncharacterized	extr	0.7206
	cgd6_520	89.4		Ser/Thr protein kinase	cyto	0.1471
	cgd6_780	86.6	FS	Sporozoite cysteine-rich protein	plas	0.2870
	cgd6_1080	70.4		Glycoprotein	extr	0.6763
	cgd6_5270	79.4		Uncharacterized	extr	1.2639
7	cgd7_2120	63.6	FS	Uncharacterized	extr	1.1248
	cgd7_4310	83.4	FS	Cysteine-rich secretory protein	extr	0.4982
8	cgd8_10	75.2		Uncharacterized	cyto	0.5436
	cgd8_20	81.0		Uncharacterized	plas	0.9078
	cgd8_30	85.8		Uncharacterized	nucl	0.8820
	cgd8_40	89.4		Uncharacterized	plas	1.4514
	cgd8_1570	71.6		CCCH like finger domain nucleoporin	mito	1.3897

	cgd8_4190	87.0		Mucin	cyto	0.6159
	cgd8_4550	78.4	FS	Uncharacterized	cyto	0.5160
	cgd8_5190	85.9		BRCA2 family protein	plas	0.5171
	cgd8_5420	78.8	FS	Uncharacterized	extr	0.5314

Supplementary Table 7

Summary of RDP4[56] recombination results with position of breakpoints, and estimated dates of divergence (thousands of generations ago) between the sequences that are related to the sequences involved in the genetic exchange. The HybridCheck[65] algorithm was used to estimate the divergence time of the recombinant blocks identified by RDP4. The "major parent" is related to the greater part of the recombinant's sequence (i.e. it is generally the recipient). The "minor parent" is related to the sequences in the proposed recombinant region (i.e. the donor). For the analysis n=4: *C. p. parvum* subtypes IIaA15G2R1 (UKP6; IIa) and IIcA5G3j (UKP16; IIc-j), *C. p. anthroponosum* subtype IIcA5G3a (UKP15; IIc-a), and *C. hominis* subtype IbA10G2 (UKH1; Ib). Subtyping was based on gp60 genotyping. The p-value represents the probability that the identified recombination block is the result of the accumulation of mutations rather than by recombination. The critical value is Bonferroni corrected, α'=0.05/n, with n equal to the number of recombination events detected.

Breakpoints (bp)		Recombinant	Major parent	Minor Parent	RDP *p*-value	CDSs encoded within	Divergence Dating (TGA)
CHROMOSOME 1							
82251	**104422**	IIa	IIc-j	Unknown	8.28E-240	cgd1_370 - cgd1_490	NA
82251	**93181**	Ib	IIc-a/IIc-j	Unknown	4.46E-08	cgd1_370 - cgd1_430	NA
100170	**100278**	IIa	IIc-a/IIc-j	Unknown	3.07E-03	cgd1_470	NA
100631	**100831**	Ib	Unknown	IIa	5.25E-04		NA
109846	**110180**	Ib	Unknown	IIc-a	8.34E-13	Intergenic cgd1_510 - cgd1_520	NA
111232	**111726**	IIc-a	IIa/IIc-j	Ib	1.26E-13	cgd1_530	32358 (95% CI: 24014-42302)
115061	**116161**	IIc-j/IIa	IIc-a	Ib	2.78E-07	cgd1_550	8476 (95% CI: 5665-12074)
127173	**136648**	IIa	IIc-j	IIc-a	2.93E-72	cgd1_580 - cgd1_590	8234 (95% CI: 7177-9388)
136649	**140781**	IIc-j	IIa	IIc-a	1.18E-15	cgd1_590 - cgd1_600	6513 (95% CI: 5166-8073)
142478	**150610**	IIa	IIc-j	IIc-a	8.96E-16	cgd1_610 - cgd1_640	3738 (95% CI: 3006-4580)
376602	**386949**	IIc-a	Unknown	IIc-j/IIa	2.92E-02	cgd1_1580 - cgd1_1640	NA
734690	**744935**	IIa	IIc-j	IIc-a	4.18E-04	cgd1_3290 - cgd1_3340	1403 (95% CI: 1016-1878)
CHROMOSOME 2							
53785	**55454**	IIc-a	IIa/IIc-j	Ib	1.38E-14	cgd2_160	13159 (95%CI: 10150-16693)
57056	**57358**	IIc-a	IIa/IIc-j	Ib	1.54E-06	cgd2_140	28412 (95% CI: 18652-40881)
58483	**58812**	IIc-a	IIa/IIc-j	Unknown	3.16E-08	cgd2_120	NA
61997	**62206**	IIc-a	IIa/IIc-j	Ib	1.78E-09	Intergenic cgd2_110 - cgd2_100	39623 (95% CI: 26118-56711)
64582	**65341**	IIc-a	IIa/IIc-j	Ib	8.90E-11	cgd2_90	13627 (95% CI: 9320-19056)
67242	**67933**	IIc-a	IIa/IIc-j	Ib	2.23E-10		19260 (95% CI: 13841-25877)
71503	**72990**	IIc-a/IIa	IIc-j	Ib	2.82E-19	cgd2_80	13490 (95% CI: 10302-17256)
75512	**76343**	IIc-a	IIa/IIc-j	Ib	5.18E-08	cgd2_70	10614 (95% CI: 7015-15246)
79931	**80238**	IIc-a	IIa/IIc-j	Ib	4.95E-10	Intergenic cgd2_70 - cgd2_60	27117 (95% CI: 17794-39042)
294024	**294928**	IIa	IIc-j	Unknown	4.32E-10	cgd2_1370	NA
341750	**405045**	IIc-a	Unknown	IIc-j	4.28E-06	cgd2_1690 - cgd2_2040	NA
432700	**506795**	IIc-j	IIa	Unknown	6.55E-04	cgd2_2170 - cgd2_2560	NA
625528	**632428**	IIa	IIc-j	Unknown	1.93E-06	cgd2_3080 - cgd2_3110	NA
CHROMOSOME 3							
220866	**220932**	IIc-a	IIc-j/IIa	Unknown	4.15E-08	cgd3_720	NA
272798	**279815**	IIc-j	IIa	IIc-a	5.11E-04	cgd3_920 - cgd3_960	1335 (95% CI: 890-1907)
319189	**319570**	IIc-a	IIc-j/IIa	Unknown	9.41E-05	cgd3_1150	NA
321883	**322660**	IIc-j	IIa	IIc-a	1.75E-15	cgd3_1160	23555 (95% CI: 17832-30326)
797968	**799943**	IIc-a	IIc-j/IIa	Ib	1.43E-77	cgd3_3370	23504 (95% CI: 19790-27632)
995078	**1030425**	IIc-j	IIa	IIc-a	2.18E-17	cgd3_4190 - cgd3_4280	776 (95% CI: 616-961)

CHROMOSOME 4

3370	**5132**	IIa/IIc-j	IIc-a	Ib	1.70E-24	cgd4_20	21044 (95% CI: 17332-25227)
5137	**5788**	IIc-a	IIa/IIc-j	Ib	5.46E-41		106298 (95% CI: 93209-120222)
848234	**849840**	IIc-j	IIa	IIc-a	2.06E-13	cgd4_3630	10881 (95% CI: 7944-14449)
865724	**865737**	IIc-a	IIa/IIc-j	Ib	1.89E-06	cgd4_3690	34157 (95% CI: 27128-42213)
1054213	**1054636**	IIc-a	IIc-j/IIa	Ib	4.80E-14	cgd4_4480	40826 (95% CI: 30656-52820)
1057053	**1058582**	IIc-j/IIa	IIc-a	Ib	1.27E-54	cgd4_4490	35621 (95% CI: 30459-41290)
1058583	**1058932**	IIc-a	IIa/IIc-j	Ib	2.10E-13	Intergenic cgd4_4490 - cgd4_4500	37164 (95% CI: 26671-49871)
1059044	**1059293**	IIc-j/IIa	IIc-a	Ib	5.69E-12	cgd4_4500	55261 (95% CI: 40390-72926)
1059418	**1060146**	IIc-a	IIa/IIc-j	Unknown	3.67E-47		NA
1060336	**1060469**	IIa/IIc-j	IIc-a	Ib	3.22E-07		62039 (95% CI: 41237-87846)
1060678	**1060737**	IIc-a	IIa/IIc-j	Ib	1.45E-06		146415 (95% CI: 101853-197080)
1061059	**1061153**	IIc-a	IIa/IIc-j	Ib	2.89E-04		79429 (95% CI: 51637-113811)
1061156	**1061888**	IIc-j/IIa	IIc-a	Ib	1.65E-77	Intergenic cgd4_4500 - 3' telomere	43165 (95% CI: 35048-52324)
1061941	**1062512**	IIc-a	IIc-j/IIa	Unknown	5.83E-27		NA
1062847	**1063606**	IIc-j/IIa	IIc-a	Ib	1.97E-61		75905 (95% CI: 65591-87050)

CHROMOSOME 5

3694	**6176**	IIa	IIc-j	IIc-a	1.03E-23	Chro.50010	9774 (95% CI: 7624-12287)
585260	**586337**	IIc-j	IIc-a/IIa	Ib	3.76E-28	cgd5_2180	81221 (95% CI: 71131-92033)
649071	**649362**	Ib	Unknown	IIa/IIc-j	8.09E-51	cgd5_1940	NA
1031972	**1033136**	IIc-j/IIa	IIc-a	Ib	2.15E-45	cgd5_40	62872 (95% CI: 52872-73873)

CHROMOSOME 6

49	**140**	IIc-j	IIa	Unknown	3.56E-05	Chro.60010	NA
146	**1792**	IIc-j	IIa	Unknown	2.59E-164		NA
1793	**1905**	IIc-j	IIa	Ib	1.43E-12	Intergenic Chro.60010 - cgd6_10	100894 (95% CI: 72306-134229)
1986	**2351**	IIc-j	IIa	Unknown	7.34E-43		NA
2352	**2537**	IIc-j	IIa	Ib	8.61E-23		108012 (95% CI: 81123-138605)
2538	**2963**	IIc-j	IIa	Unknown	3.91E-09		NA
3510	**3670**	IIc-a	IIa	Ib	3.10E-02		33270 (95% CI: 19516-51854)
4026	**6334**	IIa	IIc-j	IIc-a	6.78E-144	cgd6_10	6820 (95% CI: 4794-9277)
7166	**7713**	IIc-j/IIa	IIc-a	Ib	5.31E-18	Intergenic cgd6_10 - cgd6_20	51444 (95% CI: 41285-62947)
7784	**7896**	IIa/IIc-j	IIc-a	Ib	2.21E-04		73535 (95% CI: 49092-103541)
8033	**8972**	Ib	Unknown	IIc-a	2.03E-14	cgd6_20	NA
9758	**9992**	IIc-a	IIa	Ib	1.12E-03		25112 (95% CI: 15057-38631)
10386	**12685**	IIc-j	IIa	IIc-a	3.16E-32	cgd6_30 - cgd6_40	8573 (95% CI: 6516-11016)
13148	**13482**	IIc-a	IIc-j/IIa	Unknown	7.84E-06	Intergenic cgd6_40 - cgd6_50	NA
14883	**18178**	IIa	IIc-j	IIc-a	2.09E-24	cgd6_50	5770 (95% CI: 4366-7444)
20061	**20401**	IIc-a	IIa/IIc-j	Unknown	3.04E-19	cgd6_60	NA
20936	**21391**	IIc-j/IIa	IIc-a	Ib	7.46E-14		54420 (95% CI: 43213-67169)
186255	**187077**	IIa/IIc-j	IIc-a	Ib	1.44E-06	cgd6_800	10123 (95% CI: 6602-14690)
240190	**240717**	IIa/IIc-j	IIc-a	Ib	3.02E-06	cgd6_1020	16881 (95% CI: 11181-24180)
245902	**247871**	IIc-a	Unknown	IIa	2.52E-31	cgd6_1060	NA
247872	**256568**	IIa	IIc-j	Unknown	2.04E-228	cgd6_1060 - cgd6_1100	NA
1225101	**1225478**	IIc-a	IIa/IIc-j	Ib	1.49E-04	cgd6_5260	19386 (95% CI: 12309-28648)
1226191	**1226342**	IIc-a	IIc-j/IIa	Unknown	1.28E-15	cgd6_5260 - cgd6_5270	NA
1226343	**1226614**	IIc-a	IIa/IIc-j	Ib	8.40E-13	cgd6_5270	45290 (95% CI: 32355-60923)

1276817	**1278061**	IIc-a	IIa	Ib	7.88E-28	cgd6_5450	22826 (95% CI: 18282-28028)
1278062	**1278345**	IIc-a	IIa	Unknown	5.11E-07	Intergenic cgd6_5450 - cgd6_5500	NA
1278346	**1280578**	IIc-a	IIa	Ib	4.13E-90	cgd6_5500	109055 (95% CI: 95467-123511)

CHROMOSOME 7

285016	**292270**	IIc-j	IIa	Unknown	3.04E-06	cgd7_1150 - cgd7_1170	NA
317243	**319588**	IIc-j	IIa	IIc-a	3.97E-46	cgd7_1270	15486 (95% CI: 12724-18608)
878265	**897621**	IIc-j	IIa	IIc-a	2.69E-08	cgd7_3910 - cgd7_4020	820 (95% CI: 603-1083)
897622	**898690**	IIc-j	IIa	IIc-a	7.11E-24	cgd7_4020	7422 (95% CI: 3670-13102)
897728	**898242**	IIa	IIc-j	Unknown	1.07E-62		NA
898691	**899005**	Ib	Unknown	IIc-j/IIa	2.16E-14		NA
899011	**935740**	IIc-j	IIa	IIc-a	5.23E-25	cgd7_4020 - cgd7_4220	1241 (95% CI: 1039-1467)
1055570	**1063864**	IIc-j	IIa	IIc-a	1.45E-03	cgd7_4710 - cgd7_4750	887 (95% CI: 560-1323)

CHROMOSOME 8

80	**1150**	IIc-a	IIa/IIc-j	Unknown	9.69E-75	cgd8_10	NA
1334	**1408**	IIc-a	IIa/IIc-j	Unknown	2.72E-08	Intergenic cgd8_10 - cgd8_20	NA
1409	**1526**	IIc-a	IIa/IIc-j	Ib	6.92E-07		79550 (95% CI: 54516-109778)
3201	**3369**	IIc-a	IIa/IIc-j	Unknown	1.18E-08	cgd8_20	NA
3623	**5676**	IIc-j/IIa	IIc-a	Ib	1.29E-110		57629 (95% CI: 51933-63671)
5677	**5972**	IIc-a	IIc-j/IIa	Ib	7.91E-06		33185 (95% CI: 22613-46384)
6026	**7033**	IIc-j/IIa	IIc-a	Ib	6.26E-40	cgd8_30	43724 (95% CI: 36564-51671)
7274	**9938**	IIc-a/IIc-j	IIa	Unknown	2.95E-78		NA
10005	**11970**	IIc-j	IIa	IIc-a	6.41E-07	cgd8_40	7579 (95% CI: 5497-10123)
12805	**14933**	IIc-j	IIa	IIc-a	9.67E-17	cgd8_40 - cgd8_50	20625 (95% CI: 17271-24367)
15040	**26389**	IIa	IIc-j	IIc-a	4.29E-19	cgd8_50 - cgd8_100	3927 (95% CI: 3283-4650)
42714	**48676**	IIa	IIc-j	IIc-a	6.45E-14	cgd8_170 - cgd8_180	2321 (95% CI: 1671-3121)
75004	**84938**	IIc-j	IIa	Unknown	4.04E-06	cgd8_300 - cgd8_350	NA
547697	**563658**	IIa	IIc-j	IIc-a	5.21E-33	cgd8_2090 - cgd8_2150	3327 (95% CI: 2824-3886)
563659	**564762**	IIc-j	IIa	IIc-a	2.16E-26	cgd8_2160	20224 (95% CI: 15722-25475)
564902	**618348**	IIa	IIc-j	IIc-a	1.76E-115	cgd8_2160 - cgd8_2400	3106 (95% CI: 2834-3395)
584382	**584669**	Ib	Unknown	IIc-j	5.59E-04	cgd8_2260	NA
618349	**628553**	IIa	IIc-j	Unknown	2.73E-08	cgd8_2400 - cgd8_2440	NA
1085940	**1086106**	IIc-a	IIa/IIc-j	Unknown	9.40E-32	cgd8_4480	NA

Supplementary Table 8

Whole genome comparison of two outbreak strain WGS reveals estimated mutation accumulation rates per generation for *Cryptosporidium spp*.

UKP4 v UKP6 Whole Genome Comparison	
Sampling separation	7 days
No. of sites in WGA (bp)	9086411
No. of SNPs	10
Nucleotide diversity	0.0000011
No. of indel sites	78
No. of indel events	35
Total no. of polymorphisms (SNPs + indel Events)	45
Per base SNP mutation rate per generation (μ)	9.50E-08
Per base indel rate per generation (μ)	3.32E-07
Combined mutation rate per generation (μ)	4.27E-07

Supplementary Table 9

Oocyst infectivity and intensity rates in human volunteers summarized from peer-reviewed publications.

Reference	Challenge organism	Challenge dose	Onset of Excretion (days)	Duration of Excretion (days)	Total no. of oocysts excreted	Estimated no. of oocyst generations	Estimated no. of days/generation
24	*C. parvum*	100	7.5	3.5	1.8×10^6	4-5	2-4
	C. parvum	300	5	3	3.5×10^6	3-4	2-3
	C. parvum	1,000	4	11	3.1×10^8	4-5	3-4
	C. parvum	3,000	5	6	2.1×10^7	~3	3-4
25	*C. meleagridis*	10,000	8	3	4.5×10^8	~3	3-4

Supplementary Table 10

Description of neutrally-evolving (Ka/Ks = 0.2-0.6; 93.0-98.0% nucleotide IDs) protein-coding genes between *C. parvum parvum* UKP6 and *C. hominis* UKH4 used in the concatenated phylogeny.

Chromosome	CryptoDB ID (*C. hominis*)	CryptoDB ID (*C. parvum*)	Ka/Ks	% Nuc Ids
1	Chro.10076	cgd1_640	0.319577	96.6
	Chro.10167	cgd1_1450	0.438804	96.73
	Chro.10199	cgd1_1730	0.569446	95.58
	Chro.10229	cgd1_2000	0.564612	96.80
	Chro.10411	cgd1_3650	0.497442	95.94
	Chro.10424	cgd1_3780	0.511207	95.87
	Chro.10425	cgd1_3790	0.346492	96.8
2	Chro.20024	cgd2_180	0.4812	96.2
	Chro.20105	cgd2_940	0.382475	96.0
	Chro.20262	cgd2_2470	0.314043	95.7
	Chro.20223	cgd2_2060	0.361484	97.6
	Chro.20326	cgd2_3110	0.31982	95.32
	Chro.20388	cgd2_3630	0.586577	96.30
	Chro.20406	cgd2_3810	0.33444	97.90
3	Chro.30055	cgd3_380	0.386803	96.18
	Chro.30132	cgd3_1010	0.390058	96.03
	Chro.30206	cgd3_1720	0.407783	96.09
	Chro.30299	cgd3_2600	0.366692	97.25
	Chro.30349	cgd3_3070	0.326581	96.76
	Chro.30377	cgd3_3310	0.511435	95.60
	Chro.30413	cgd3_3650	0.262038	97.25
	Chro.30476	cgd3_4230	0.333963	96.12
4	Chro.40051	cgd4_370	0.111926	97.55
	Chro.40248	cgd4_2180	0.387906	97.82
	Chro.40252	cgd4_2210	0.217421	97.63
	Chro.40294	cgd4_2620	0.466828	96.92
	Chro.40317	cgd4_2820	0.504021	96.98
	Chro.40433	cgd4_3800	0.509732	97.39
	Chro.40495	cgd4_4360	0.341557	96.46
	Chro.40503	cgd4_4440	0.350652	97.20
5	Chro.50012	cgd5_3600	0.292362	96.80
	Chro.50084	cgd5_2890	0.425943	96.54
	Chro.50103	cgd5_2730	0.410499	97.23
	Chro.50107	cgd5_2700	0.527435	96.40
	Chro.50155	cgd5_2250	0.249098	96.80
	Chro.50195	cgd5_1860	0.389703	96.68
	Chro.50250	cgd5_1340	0.416003	97.1
	Chro.50420	cgd5_4240	0.322667	96.63
6	Chro.60245	cgd6_2100	0.313076	97.4
	Chro.60295	cgd6_2560	0.382122	96.83
	Chro.60314	cgd6_2720	0.462682	96.05
	Chro.60470	cgd6_4090	0.36524	96.51
	Chro.60490	cgd6_4280	0.366079	96.13
	Chro.60610	cgd6_5300	0.4904	97.43
	Chro.60619	cgd6_5370	0.441644	96.72
7	Chro.70047	cgd7_340	0.333681	96.19
	Chro.70111	cgd7_890	0.318737	96.0
	Chro.70152	cgd7_1270	0.484978	95.8
	Chro.70160	cgd7_1330	0.419706	96.1
	Chro.70211	cgd7_1810	0.292609	96.72
	Chro.70267	cgd7_2340	0.297261	96.8
	Chro.70296	cgd7_2600	0.318605	96.4
	Chro.70395	cgd7_3550	0.500737	96.76
8	Chro.80024	cgd8_140	0.366147	97.36
	Chro.80102	cgd8_830	0.505411	96.50
	Chro.80229	cgd8_1960	0.378299	96.38
	Chro.80245	cgd8_2080	0.435382	96.39
	Chro.80332	cgd8_2850	0.437901	96.7
	Chro.80353	cgd8_3030	0.287705	96.45
	Chro.80409	cgd8_3560	0.438279	96.96
	Chro.80605	cgd8_5310	0.470142	96.32

Supplementary Figure 1

Host ranges for human-infective *Cryptosporidium spp*. gp60 subtype families from GenBank-submitted gp60 sequences. Host ranges were determined for *C. hominis* gp60 subtypes Ia (N=327) and Ib (N=1752), *C. p. anthroponosum* IIc-a (N=111), and *C. p. parvum* subtypes IIa (N=843) and IId (N=377). Host types were characterised as equine, human, marsupial, mollusc, rodent, ruminant, primate, and other.

Supplementary Figure 1.eps

Supplementary Figure 2

Concatenated phylogeny of 21 human-infective *Cryptosporidium spp*. The maximum likelihood (ML) phylogeny based on a 153,421 bp alignment of 61 loci is shown. Included sequence targets exhibited neutral evolution between *C. p. parvum* UKP6 and *C. hominis* UKH4 (Ka/Ks 0.2-0.6, 93.0-98.0% nucleotide identities). Confidence values on the phylogeny reflect 2,000 bootstrap replications.

Supplementary Figure 2.eps

Supplementary Figure 3

Gene-by-gene signatures of selection (Ka/Ks) and nucleotide diversity (π) between human-infective *Cryptosporidium spp*. WGS across chromosomes 1-8. The nucleotide diversity is highest in *C. hominis* UKH4, whereas the signature of positive selection is most pronounced for *C. p. anthroponosum* UKP15.

Supplementary Figure 3.eps

Supplementary Figure 4

Mean (±SE) nucleotide diversity (π) and signature of selection (Ka/((Ks+1)/S)) of genes in the non-telomeric (green, n=2827 CDSs), subtelomeric (yellow, n=326 CDSs) and peri-telomeric (red, n=312 CDSs) regions. Genes near the telomeres are the fastest evolving.

Supplementary Figure 4.eps

Supplementary Figure 5

(A) Predicted proportion of protein localization types for genome-wide CDSs and CDSs exhibiting significantly positive Ka/Ks values (>1.0), as compared between *C. p. parvum* UKP6 and *C. p. anthroponosum* UKP15. Protein localizations were categorised as cytoskeleton (Cysk), cytoplasm (Cyto), endoplasmic reticulum (E.R.), mitochondrion (Mito), nucleus (Nuc), peroxisome (Pero) and plasma membrane (Plas). (B) Comparative selective pressure (Ka/(Ka+Ks)) and nucleotide diversity (π) between CDSs annotated as having a cytoplasmic versus extracellular protein localization. Extracellular CDSs have a significantly faster rate of evolution (higher π) that is driven by positive selection (significantly higher Ka/(Ka+Ks) (two-tailed Mann-Whitney test n=3465 CDSs: Cytoplasmic n=1152 (Min=0.0000000, Median=0.0009709, Max=0.0375539), Extracellular n=333 (Min=0.0000000, Median=0.001311, Max=0.837771)). Exact p-value Mann-Whitney Ka/(Ka+Ks): p=0.0013. Exact p-value Mann-Whitney nucleotide diversity (π): p=1.233E-07.

Supplementary Figure 5.eps

Supplementary Figure 6

Mean and 5-95% confidence intervals of the expected number of recombination events per chromosome (based on chromosome size expressed as nucleotides) compared to observed number of recombination events in the RDP4 analysis (see Supplementary Table 2). The number of recombination events (n=104) are not homogeneously distributed across chromosomes, and chromosome 6 shows a significantly elevated number of events.

Supplementary Figure 7

Incongruence between concatenated (A) and GP60-based (B) phylogenies of WGS used in this study. Zoomed sections illustrate phylogenies constructed using the same sequence alignments, but including only *C. parvum* WGS. This illustrates that the taxonomic relationships of the isolates based on the commonly used GP60 locus differs from that obtained by WGS, and that the GP60 locus alone cannot effectively resolve the evolutionary relationships between species. Trees were generated using the automated ClustalW alignment algorithm and Maximum Likelihood phylogeny builder, using 1000 bootstrap replications, in Mega 7.0.[48]

Supplementary Figure 7.eps

Supplementary Figure 8

Stacked bar graph of the number of calls of bases from the reads of the four isolates that were studied in the genetic introgression analysis (UKH1, UKP6, UKP15 and UKP16). Note that the Y-axis is $\log_{10}$-transformed, and that the vast majority (>99.85%) of the calls are single bases (AC=0), which gives confidence that each of these four samples represent a single isolate.

Method
In order to examine whether the nucleotide bases that have been called were derived from a single genotype, or whether there might be multiple diverged genotypes present in a reads (e.g. due to mixed infections), we counted the AC values of all bases in the four isolates that were studied in the genetic introgression analysis (UKH1, UKP6, UKP15 and UKP16). In this analysis, AC=0 represent "single called" bases for which there is no evidence of alternative calls. AC=1 indicates an ambiguous call, and AC=2 indicates a true alternative call. Such ambiguous and alternative calls are evidence of polymorphisms, which for this haploid species suggests either: (1) contamination from e.g. mixed infections, (2) polymorphisms arising due to novel mutations in the genome of parasite population accumulated whilst in the host, or (3) sequencing errors. To produce the graph in Fig. S8, poor quality bases, adaptor sequences and reads less than 36 base pairs (bp) long were removed using Trimmomatic [1]. The reads were then mapped to the respective genome assemblies for each isolate with Bowtie2 [2] using the "--sensitive-local" mapping parameters, and a 5 bp trim applied to each end. Pilon [3] was then run with default parameters, to fix the SNPs and indels only (i.e. the "--fix bases" option) and to output a VCF file of the sequence variants. For all four isolates examined, the fast majority of bases (>99.85%) were reliable assessed as "single calls" (i.e. AC=0), (Fig. S8). The UKP6 isolate had 0.134% of its bases called ambiguously (AC=1), and 0.009% bases called with an alternative base (AC=2). This represents a very small fraction of the genome in total, which gives confidence that each of these four samples represent a single isolate.

References
[1] Bolger, A. M., Lohse, M., & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120 (2014).
[2] Langmead, B., & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature methods* **9**, 357 (2012)
[3] Walker, B. J. *et al*. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PloS one* **9**, e112963 (2014).


Supplementary figure 8.eps

Supplementary Figure 9

Illustration of a *Cryptosporidium* generation[79,80]
Schematic illustrates the required rounds of DNA replication to complete the *Cryptosporidium* life-cycle. Oocysts in the environment contain four haploid sporozoites which are released from thick-walled oocysts in the host after ingestion. Each sporozoite is infective, forming a trophozoite following infection and invasion of an intestinal epithelial cell. Three rounds of DNA replication – merogony – follow, forming a type 1 meront which releases 8 type I merozoites. Each type 1 merozoite is able to independently infect an additional epithelial cell and two further rounds of DNA replication follow to form a type 2 meront which releases 4 type II merozoites. Alternatively, type 1 merozoites can produce further type 1 meronts. Type 2 merozoites are able to undergo gametocytogenesis producing either single haploid macrogametocyte or (following four rounds of DNA replication) 16 haploid microgametes. The cycle is completed when fusion of a microgamete with a macrogametocyte produce a diploid zygote and the ensuing meiosis gives rise to oocysts with 4 haploid sporozoites. Oocysts are either thick-walled environmentally resistant forms or thin walled forms that lead to autoinfection. (n = one haploid genome. The proportions/numbers of parasites shown progressing through the life-cycle are approximated for illustrative purposes).

Supplementary Figure 9.eps