incubated with 20 μl Sepharose-immobilized monoclonal anti-HA antibodies (Covance). Beads were washed with buffer T without BSA before elution. DSP-induced crosslinks in eluted proteins were thiol-cleaved before separation by 8–16% SDS–PAGE and detection by Sypro Ruby (Bio-Rad).

Quinone analyses

Lipid extractions and quinone detection were performed as described[19]. Hydrogenosomes (5.5 mg protein) were extracted and resuspended in 150 μl 9:1 methanol/ethanol, of which 50 μl was injected onto an HPLC system linked to an ECD.

Sequence analyses

Accession numbers for sequences used to reconstruct NuoF and NuoE phylogenies are listed in Supplementary Tables 2 and 3. NuoF sequences were aligned with CLUSTALX. NuoE sequences were aligned with Wisconsin Package Version 10.2 programs (Genetics Computer Group). A profile hidden Markov model (HMM) was built from *Escherichia coli*, *Neurospora crassa*, *Bos taurus*, *Paracoccus denitrificans* and *Thermus thermophilus* sequences with HMMBUILD. Additional sequences were aligned to the profile with HMMALIGN. Both alignments were edited to remove C- and N-terminal extensions. Analyses of NuoF and NuoE evolution were performed with MRBAYES[30] with the JTT amino-acid substitution model and with two Markov chains Monte Carlo. Chains were run for 100,000 generations, with sampling every 50 generations. The first 5,000 generations were discarded as burn-in. Consensus trees satisfying the more than 50 majority rule were drawn with Treeview, and probabilities of branch partitions were calculated.

Received 25 June; accepted 2 September 2004; doi:10.1038/nature02990.

1. Müller, M. The hydrogenosome. *J. Gen. Microbiol.* **139,** 2879–2889 (1993).
2. Horner, D. S., Hirt, R. P. & Embley, T. M. A single eubacterial origin of eukaryotic pyruvate: ferredoxin oxidoreductase genes: implications for the evolution of anaerobic eukaryotes. *Mol. Biol. Evol.* **16,** 1280–1291 (1999).
3. Horner, D. S., Foster, P. G. & Embley, T. M. Iron hydrogenases and the evolution of anaerobic eukaryotes. *Mol. Biol. Evol.* **17,** 1695–1709 (2000).
4. Embley, T. M., van der Giezen, M., Horner, D. S., Dyal, P. L. & Foster, P. Mitochondria and hydrogenosomes are two forms of the same fundamental organelle. *Phil. Trans. R. Soc. Lond. B* **358,** 191–201 (2003).
5. Gray, M. W., Burger, G. & Lang, B. F. Mitochondrial evolution. *Science* **283,** 1476–1481 (1999).
6. Clemens, D. L. & Johnson, P. J. Failure to detect DNA in hydrogenosomes of *Trichomonas vaginalis* by nick translation and immunomicroscopy. *Mol. Biochem. Parasitol.* **106,** 307–313 (2000).
7. Dyall, S. D., Brown, M. T. & Johnson, P. J. Ancient invasions: from endosymbionts to organelles. *Science* **304,** 253–257 (2004).
8. Bui, E. T. N., Bradley, P. J. & Johnson, P. J. A common evolutionary origin for mitochondria and hydrogenosomes. *Proc. Natl Acad. Sci. USA* **93,** 9651–9656 (1996).
9. Horner, D. S., Hirt, R. P., Kilvington, S., Lloyd, D. & Embley, T. M. Molecular data suggest an early acquisition of the mitochondrion endosymbiont. *Proc. R. Soc. Lond. B* **263,** 1053–1059 (1996).
10. Roger, A. J., Clark, C. G. & Doolittle, W. F. A possible mitochondrial gene in the early-branching amitochondriate protist *Trichomonas vaginalis*. *Proc. Natl Acad. Sci. USA* **93,** 14618–14622 (1996).
11. Dyall, S. D. *et al.* Presence of a member of the mitochondrial carrier family in hydrogenosomes: conservation of membrane-targeting pathways between hydrogenosomes and mitochondria. *Mol. Cell. Biol.* **20,** 2488–2497 (2000).
12. Friedrich, T. & Bottcher, B. The gross structure of the respiratory complex I: a Lego System. *Biochim. Biophys. Acta* **1608,** 1–9 (2004).
13. Fearnley, I. M. & Walker, J. E. Conservation of sequences of subunits of mitochondrial complex I and their relationships with other proteins. *Biochim. Biophys. Acta* **1140,** 105–134 (1992).
14. Yano, T., Sled, V. D., Ohnishi, T. & Yagi, T. Identification of amino acid residues associated with the [2Fe–2S] cluster of the 25 kDa (NQO2) subunit of the proton-translocating NADH-quinone oxidoreductase of *Paracoccus denitrificans*. *FEBS Lett.* **354,** 160–164 (1994).
15. Yano, T., Sled, V. D., Ohnishi, T. & Yagi, T. Expression and characterization of the flavoprotein subcomplex composed of 50-kDa (NQO1) and 25-kDa (NQO2) subunits of the proton-translocating NADH-quinone oxidoreductase of *Paracoccus denitrificans*. *J. Biol. Chem.* **271,** 5907–5913 (1996).
16. Hrdy, I. & Müller, M. Primary structure of the hydrogenosomal malic enzyme of *Trichomonas vaginalis* and its relationship to homologous enzymes. *J. Eukaryot. Microbiol.* **2,** 593–603 (1995).
17. Hrdy, I. & Müller, M. Primary structure and eubacterial relationships of the pyruvate:ferredoxin oxidoreductase of the amitochondriate eukaryote *Trichomonas vaginalis*. *J. Mol. Evol.* **41,** 388–396 (1995).
18. Dolezal, P., Vanacova, S., Tachezy, J. & Hrdy, I. Malic enzymes of *Trichomonas vaginalis*: two enzyme families, two distinct origins. *Gene* **329,** 81–92 (2004).
19. Jonassen, T., Davis, D. E., Larsen, P. L. & Clarke, C. F. Reproductive fitness and quinone content of *Caenorhabditis elegans clk-1* mutants fed coenzyme Q isoforms of varying length. *J. Biol. Chem.* **278,** 51735–51742 (2003).
20. Meganathan, R. Ubiquinone biosynthesis in microorganisms. *FEMS Microbiol. Lett.* **203,** 131–139 (2001).
21. Massanz, C., Schmidt, S. & Friedrich, B. Subforms and *in vitro* reconstitution of the NAD-reducing hydrogenase of *Alcaligenes eutrophus*. *J. Bacteriol.* **180,** 1023–1029 (1998).
22. Land, K. M. *et al.* Targeted gene replacement of a ferredoxin gene in *Trichomonas vaginalis* does not lead to metronidazole resistance. *Mol. Microbiol.* **51,** 115–122 (2004).
23. Friedrich, T. & Scheide, D. The respiratory complex I of bacteria, archaea and eukarya and its module common with membrane-bound multisubunit hydrogenases. *FEBS Lett.* **479,** 1–5 (2000).
24. Lang, B. F. *et al.* An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. *Nature* **387,** 493–497 (1997).
25. Andersson, S. G. *et al.* The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. *Nature* **396,** 133–140 (1998).
26. Akhmanova, A. *et al.* A hydrogenosome with a genome. *Nature* **396,** 527–528 (1998).
27. Doolittle, W. F. *et al.* How big is the iceberg of which organellar genes in nuclear genomes are but the tip? *Phil. Trans. R. Soc. Lond. B* **358,** 39–57 (2003).
28. Martin, W. & Müller, M. The hydrogen hypothesis for the first eukaryote. *Nature* **392,** 37–41 (1998).
29. Doolittle, W. F. You are what you eat: a gene transfer ratchet could account for bacterial genes in eukaryotic nuclear genomes. *Trends Genet.* **14,** 307–311 (1998).
30. Huelsenbeck, J. P. & Ronquist, F. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17,** 754–755 (2001).

Supplementary Information accompanies the paper on **www.nature.com/nature**.

Acknowledgements We thank J. L. Kerwin for mass spectrometric analyses. Preliminary *T. vaginalis* genome sequence data were obtained from TIGR through the website at http://www.tigr.org. Sequencing of the *T. vaginalis* genome was accomplished with support from the NIH. This work was supported by National Institute of Health (NIH) grants (P.J.J. and C.F.C.) and a National Aeronautics and Space Administration Astrobiology grant to UCLA. P.J.J. is a Burroughs Wellcome Scholar in Molecular Parasitology. The UCLA Mass Spectrometry and Proteomics Technology Center (J.A.L.) was established with a grant from the W. M. Keck Foundation.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to P.J.J. (johnsonp@ucla.edu).

The genome of *Cryptosporidium hominis*

Ping Xu[1,2,3]*, Giovanni Widmer[4]*, Yingping Wang[1,2], Luiz S. Ozaki[1,2], Joao M. Alves[1,2], Myrna G. Serrano[1,2], Daniela Puiu[1], Patricio Manque[1,2], Donna Akiyoshi[4], Aaron J. Mackey[5]†, William R. Pearson[6], Paul H. Dear[7], Alan T. Bankier[7], Darrell L. Peterson[8], Mitchell S. Abrahamsen[9,10], Vivek Kapur[10,11], Saul Tzipori[4] & Gregory A. Buck[1,2]

[1]*Center for the Study of Biological Complexity, Virginia Commonwealth University, Richmond, Virginia 23284-2030, USA*
[2]*Department of Microbiology and Immunology, Virginia Commonwealth University, Richmond, Virginia 23298-0678, USA*
[3]*Philips Institute for Oral and Craniofacial Molecular Biology, Virginia Commonwealth University, Richmond, Virginia 23298-0566, USA*
[4]*Tufts University School of Veterinary Medicine, North Grafton, Massachusetts 01536, USA*
[5]*Department of Microbiology, University of Virginia, Charlottesville, Virginia 22908, USA*
[6]*Department of Biochemistry and Molecular Genetics, University of Virginia, Charlottesville, Virginia 22908, USA*
[7]*MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK*
[8]*Department of Biochemistry and Molecular Biophysics, Virginia Commonwealth University, Richmond, Virginia 23298-0614, USA*
[9]*Department of Veterinary and Biomedical Sciences, University of Minnesota, St Paul, Minnesota 55108, USA*
[10]*Biomedical Genomics Center, University of Minnesota, St Paul, Minnesota 55108, USA*
[11]*Department of Microbiology, University of Minnesota, St Paul, Minnesota 55108, USA*

* These authors contributed equally to this work
† Present address: Department of Biology, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

Cryptosporidium **species cause acute gastroenteritis and diarrhoea worldwide. They are members of the Apicomplexa—protozoan pathogens that invade host cells by using a specialized apical complex and are usually transmitted by an invertebrate vector or intermediate host. In contrast to other Apicomplexans, *Cryptosporidium* is transmitted by ingestion of oocysts and completes its life cycle in a single host. No therapy is available, and control focuses on eliminating oocysts in water supplies[1]. Two species, *C. hominis* and *C. parvum*, which differ in host range, genotype and pathogenicity, are most relevant to humans[1–3]. *C. hominis* is restricted to humans, whereas**

***C. parvum* also infects other mammals[2]. Here we describe the eight-chromosome ~9.2-million-base genome of *C. hominis*[2]. The complement of *C. hominis* protein-coding genes shows a striking concordance with the requirements imposed by the environmental niches the parasite inhabits. Energy metabolism is largely from glycolysis. Both aerobic and anaerobic metabolisms are available, the former requiring an alternative electron transport system in a simplified mitochondrion. Biosynthesis capabilities are limited, explaining an extensive array of transporters. Evidence of an apicoplast is absent, but genes associated with apical complex organelles are present. *C. hominis* and *C. parvum* exhibit very similar gene complements, and phenotypic differences between these parasites must be due to subtle sequence divergence.**

We generated a ~12-fold sequence and ~8-fold bacterial artificial chromosome (BAC) clone coverage of the genome of *C. hominis* isolate TU502 (ref. 3, Fig. 1, Supplementary Figs 1–8, Supplementary Tables 1 and 2). Alignment of the ~9.2-million-base (Mb) final sequence with the HAPPY map[4] and chromosomes of the *C. parvum* genome[5] covered ~9.1 Mb. The eight chromosomes range from ~0.9 to ~1.4 Mb and exhibit 31.7% GC content (compare with 30.3% and 19.4% for *C. parvum* and *P. falciparum*[6], respectively). The density of 2–50-base-pair (bp) repeats was about 1 per 2,800 bp. The distribution of repeats is biased towards chromosome ends because over 85% are in the telomere-proximal thirds of five of the chromosomes (Supplementary Fig. 9). Two octamers, TGGCGCCA and TGCATGCA, over-represented in other apicomplexans[4], are ~40-fold and 15-fold over-represented in *C. hominis* (Supplementary Table 3). More than 80% of these are in non-coding sequences, indicating possible regulatory or other conserved function. Forty-five tRNAs, four or five rRNA operons—at least one of each of the two known types (Supplementary Table 4)—and two clusters of three tandem 5S rRNA genes are present. As in *P. falciparum*[6], two $tRNA^{Met}$ genes are present, suggesting discrete roles in initiation and extension. We estimate that there are ~3,994 genes in *C. hominis*, in comparison with 3,952 genes in *C. parvum* and 5,268 in *P. falciparum*[6] (Table 1). About 60% exhibit similarity to known genes. The distribution of GO annotations for *Cryptosporidium*, *Plasmodium* and *Saccharomyces*, is remarkably similar (Supplementary Fig. 10), indicating that their phenotypic differences are a reflection of non-conserved or previously unreported gene families of unknown function rather than to the functional specialization of conserved gene families. We estimate that 5–20% of *C. hominis* genes have introns.

Table 1 ***Cryptosporidium hominis* genome summary**

(a) The genome	*C. hominis*	*C. parvum*	*P. falciparum*
Size (Mb)	9.16	9.11	22.85
No. of physical gaps	246	5	93
No. of contigs	1413*	n.a.	n.a.
(G+C) content (%)	31.7	30.3	19.4
Coding regions†			
Coding size (Mb)	6.29	6.80	12.03
Percentage coding	69	74	53
(G + C) content (%)	32.3	31.9	23.7
No. of genes	3,994	3,952	5,268
Mean gene length (bp)	1,576	1,720	2,283
Gene density (bp per gene)	2,293	2,305	4,338
Genes with introns (%)‡	5–20%	5%	54%
Hits nr§	2,331	2,483	n.d.
Percentage hits nr§	58	63	n.d.
Intergenic regions			
Non-coding size (Mb)	2.87	2.32	10.83
Percentage not coding	31	25	47
(G+C) content (%)	30.3	25.6	14.6
No. of intergenic regions	4,003	3,960	6,392
Mean length (bp)	716	585	1,694
RNAs			
No. of tRNA genes	45	45	43
No. of 5S rRNA genes	6	6	3
No. of 5.8S,18S and 28S	5	5	7
(b) The proteome			
Total predicted proteins	3,994	3,952	5,268
Hypothetical proteins	2,779	2,567	3,208
Gene ontology			
Biological process	1,239	n.d.	1,613
Cellular component	1,265	n.d.	1,586
Molecular function	1,235	n.d.	1,625
Structural features			
Transmembrane domain	786	n.d.	1,631
Signal peptide	421	n.d.	544
Signal anchor	221	n.d.	367

*An additional 673 very short contigs are not assembled and probably include contaminant sequences.
†Excluding introns.
‡Estimated intron content from expressed sequence tags.
§Hits, or putative homologous proteins in the non-redundant protein database.
Hypothetical proteins, proteins without sufficient similarity to any other gene to permit functional assignment; n.a., not applicable; n.d., not determined; physical gaps, those that no existing clone closes; transmembrane domains, TMHMM, Trans Membrane Hidden Markov Model (for prediction of transmembrane helices in proteins); signal peptide and signal anchor, SignalP-2.0. *C. parvum* and *C. hominis* genomes were annotated with identical strategies to permit comparison.

Analysis of the *C. hominis* genome shows that the parasite possesses a highly tailored glycolysis-based metabolism, is dependent on the host for nutrients, and is exquisitely adapted for its life cycle (Fig. 2, Supplementary Tables 5 and 6). Glycolysis seems functional, unlike the tricarboxylic acid (TCA) cycle and oxidative phosphorylation. Both an anaerobic pathway using pyruvate:$NADP^+$ oxidoreductase (PNO) and an aerobic pathway using an alternative oxidase (AOX) are available for recycling NAD^+ to NADH. In the former, pyruvate is fermented to acetyl coenzyme A (acetyl-CoA) producing NADPH, which is then reduced to $NADP^+$, releasing hydrogen, by a Narf-like [Fe]-hydrogenase, as in *Trichomonas*[7]. Acetyl-CoA is processed by acetate CoA synthase to produce acetate and ATP, as in *Giardia*[8], yielding four ATP per glucose. Acetyl-CoA can also be processed to ethanol yielding no additional ATP. Under glucose-limited conditions, conversion of acetyl-CoA to acetate, generating two extra ATP per glucose, might be favoured. When glucose is in excess, pyruvate can be converted to lactate or ethanol to regenerate NAD^+ but no additional ATP. *C. hominis* can also generate ATP by metabolism of glycerol using glycerol-3-phosphate dehydrogenase and triose phosphate isomerase.

C. hominis can convert pyruvate to malate and subsequently to oxaloacetate (OAA), regenerating NAD^+. However, malate shuttle enzymes—for example, aspartate amino transferase—which process OAA to aspartic acid for export from the mitochondrion, are absent. Cytoplasmic malate could be converted to OAA by a mitochondrial membrane-bound malate dehydrogenase, like the lactate shuttle of *Euglena gracilis*[9], passing electrons from malate to an electron transport system composed of elements of Complexes I and III and an alternative oxidase system with O_2 as electron acceptor and producing no additional ATP.

Enzymes for metabolism of glycogen, starch and amylopectin are present, which is consistent with suggestions that amylopectin represents an energy reserve for sporozoites[10]. Lack of glucose-6-phosphate 1-dehydrogenase and other enzymes of the pentose phosphate pathway suggests that, unlike *P. falciparum* and other apicomplexans[6], *C. hominis* cannot metabolize five-carbon sugars or nucleotides. Components of β-oxidation, for example enoyl-CoA hydratase and acetyl-CoA C-acyltransferase, are also absent, precluding ATP generation from fatty acids. Enzymes for the catabolism of proteins are also absent.

Major TCA-cycle enzymes—isocitrate dehydrogenase, succinyl-CoA synthase and succinate dehydrogenase—are absent in *C. hominis*. Despite the presence of ubiquinol-cytochrome *c* reductase, NADH dehydrogenase (ubiquinone), H^+-transporting ATPase and iron–sulphur cluster-like proteins, among others, key components of Complexes II and IV are absent, precluding ATP generation by oxidative phosphorylation. Components of oxidative phosphorylation that are present (parts of Complexes I and III) probably reoxidize NADH in a simplified electron-transport chain,

as in some plants and protozoa.

Consistent with previous suggestions[11] is the observation that *Cryptosporidium* lacks enzymes for the synthesis of key biochemical building blocks—simple sugars, amino acids and nucleotides. However, starch, amylopectin and fatty acids can be generated from precursors. Interestingly, these *C. hominis* enzymes have minimal similarity to the known biosynthetic enzymes and are potential therapeutic targets.

Enzymes of the TCA, urea and nitrogen cycles and of the shikimate pathway are absent, indicating that *Cryptosporidium* is an amino-acid auxotroph. The shikimate pathway has been proposed as a potential target for glyphosate-based chemotherapy in other parasites including *Cryptosporidium*. We found no evidence to support this hypothesis. Enzymes that interconvert amino acids are encoded in *C. hominis*, and, unlike *P. falciparum*[6], *C. hominis* has a large complement of amino acid transporters.

C. hominis lacks enzymes to synthesize bases or nucleosides, but encodes enzymes that convert nucleosides into nucleotides and

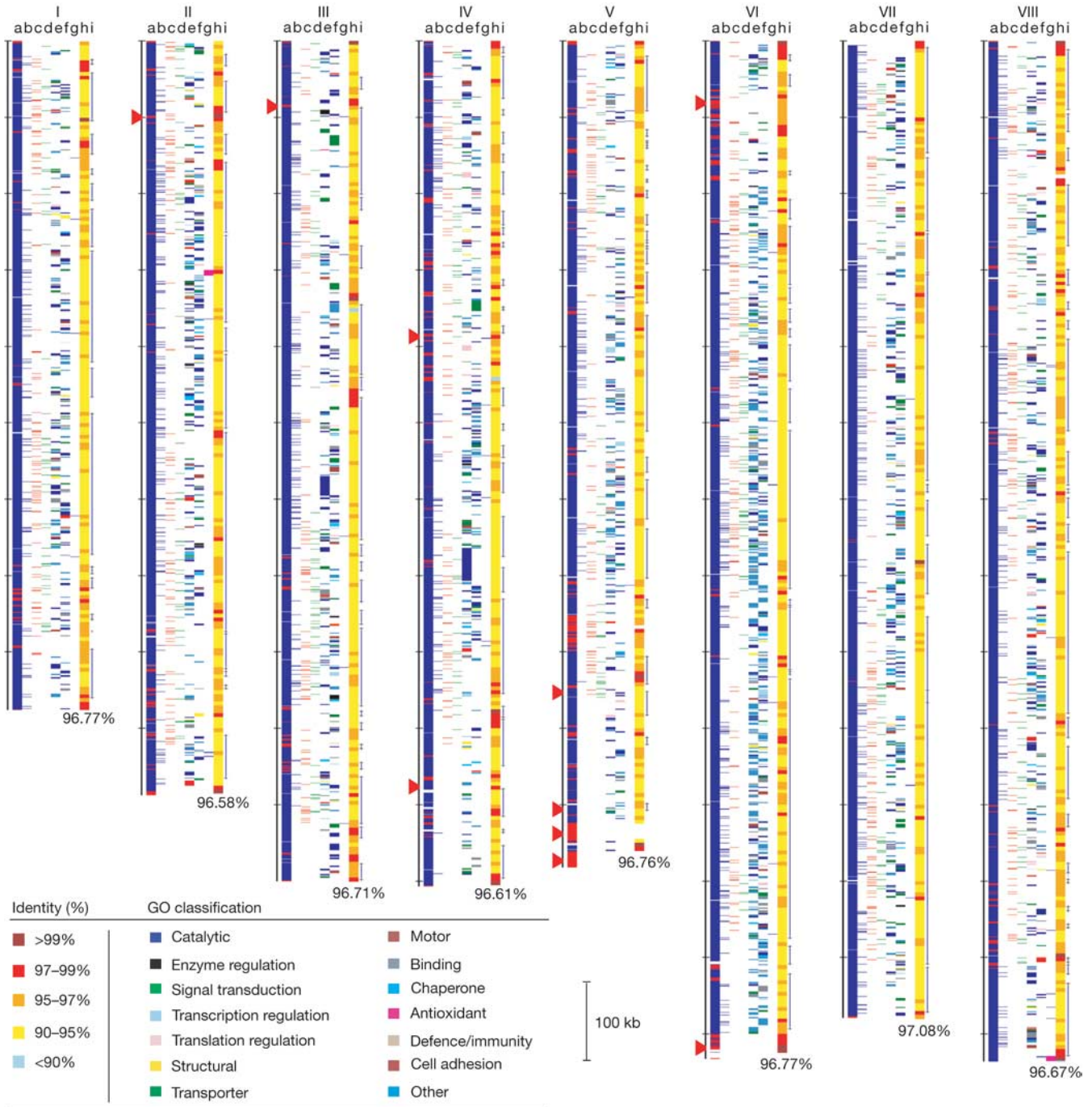

Figure 1 Schematic representation of the *C. hominis* chromosomes. Tracks indicate *C. hominis* contigs (blue), sequence gaps (white) and physical gaps (red) (**a**); HAPPY[4] markers (**b**); positions of the octamers TGGCGCCA (**c**) and TGCATGCA (**d**); Gene Ontology (GO) of molecular function for the predicted genes shown by strand (see key, left (**e**) and right (**f**)); tRNAs (blue) and rRNAs (magenta) (**g**); percentage identity to *C. parvum* in 5-kb windows (see key; average identities are shown at the foot of each chromosome) (**h**); BAC clone coverage (overlapping clones collapsed to a single line) (**i**). The scale to the left of each chromosome represents *C. parvum* sequences (red triangles show sequence gaps), with the first base at the top.

interconvert nucleotides. As in other parasites, thymidylate synthase and dihydrofolate reductase of *C. hominis* are encoded as a bifunctional polypeptide, and novel polymorphisms at crucial sites have been proposed to explain *Cryptosporidium*'s resistance to antifolates[12]. As previously suggested[11], several nucleotide conversion enzymes seem to have a prokaryotic origin.

Fatty-acid biosynthesis in apicomplexans occurs in the apicoplast by means of a type II system including fatty-acid synthase (FAS). However, consistent with the absence of an apicoplast in *Cryptosporidium*[13] is the observation that *C. hominis* encodes large FAS and polyketide synthase (PKS) enzymes, indicating a type I mechanism. The type I FAS and PKS enzymes of *C. hominis* also have prokaryotic characteristics[14,15].

Glycerolipid and phospholipid metabolic pathways for phosphatidylinositol biosynthesis are available in *C. hominis*. 1,2-Diacylglycerol is a precursor for glycosylphosphatidylinositol anchor synthesis. All enzymes required for synthesis of these anchors are apparently present[16].

Polyamines like putrescine, spermine and spermidine are critical for cellular viability, and enzymes required for their synthesis are attractive therapeutic targets. *Cryptosporidium* can synthesize polyamines using arginine decarboxylase rather than ornithine decarboxylase[17]. The putative arginine decarboxylase, spermidine synthase and other relevant enzymes encoded by *C. hominis* have diverged significantly from their homologues and are potential therapeutic targets.

C. hominis encodes adenylate cyclase, cyclic-AMP phosphodiesterase and protein kinase A, indicating the presence of the cAMP-mediated signalling pathway (Supplementary Table 7). Trimeric G protein, often involved in the activation of cAMP-mediated signalling, was not found in *C. hominis*, indicating that, as in Kinetoplastida[18] and reminiscent of plants, this pathway is independent of this complex in *C. hominis*. The presence of phosphatidylinositol 3-kinase and phospholipase C indicates that *C. hominis* utilizes phosphatidylinositol phosphate and Ca^{2+}-mediated regulatory mechanisms. The presence of putative Ca^{2+} transporters,

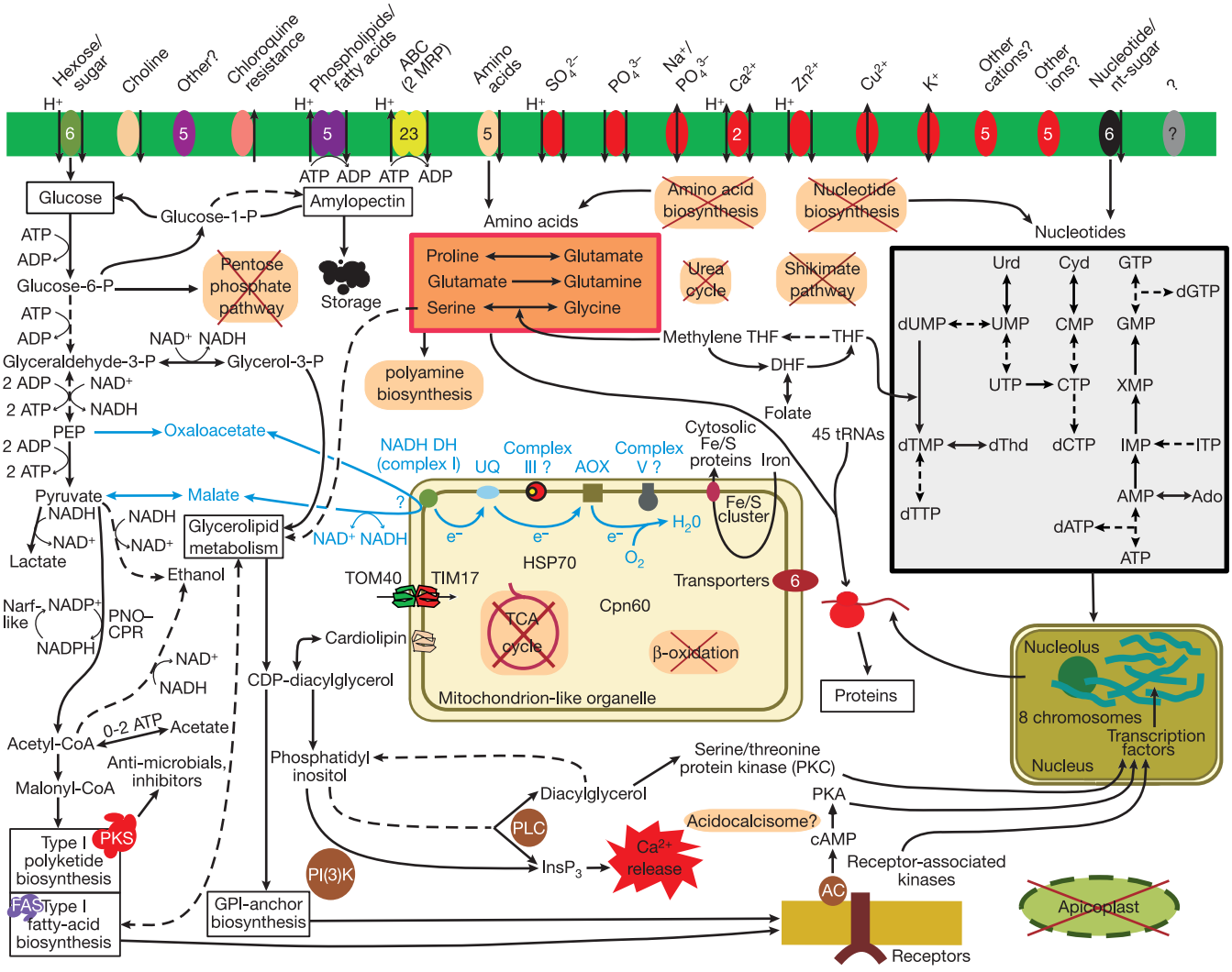

Figure 2 Schematic representation of selective *C. hominis* proteins, enzymes and pathways. The green strip represents the cellular membrane with putative transporters; numbers indicate the number of genes for a given class of transporter. Solid arrows indicate pathways that are present; multistep pathways are indicated with dashed arrows. Components or pathways that are absent are crossed out. Steps or components whose exact nature is questionable are shown with question marks. Blue arrows and names indicate proposed aerobic parts of the metabolism. Abbreviations: ABC, ATP-binding cassette; AC, adenylyl cyclase; Ado, adenosine; AOX, alternative oxidase; Cpn60, chaperone 60; Cyd, cytidine; DHF, dihydrofolate; dThd, deoxythymidine; GPI, glycosylphosphatidylinositol; Hsp70, heat-shock protein 70; $InsP_3$, inositol phosphate; MRP, multiple-drug-resistance protein; NADH DH, NADH dehydrogenase; Narf-like, nuclear prelamin A recognition factor-like protein; PEP, phosphoenolpyruvate; PI(3)K, phosphatidylinositol 3-kinase; PKA, protein kinase A; PLC, phospholipase C; PKC, protein kinase C; PNO–CPR, pyruvate:$NADP^+$ oxidoreductase fused to cytochrome P450 reductase domain; THF, tetrahydrofolate; TIM17, translocase of the inner mitochondrial membrane 17; TOM40, translocase of the outer mitochondrial membrane 40; UQ, ubiquinone; Urd, uridine.

enzymes associated with acidocalcisomes, and calmodulin imply that Ca^{2+} transport and sequestering are functional. Protein kinase C receptors indicate that *C. hominis* has the ability to signal by activation of soluble cytoplasmic receptor-associated kinases.

No mitochondrial DNA sequences were found in *C. hominis*, and the TCA cycle and oxidative phosphorylation are absent (Supplementary Tables 5, 6 and 8). However, a double-membrane-bound organelle generates a proton gradient using cardiolipin and performs some related mitochondrial functions, and mitochondrial marker chaperonin 60 was localized to this structure[19]. Core enzymes of [Fe–S] cluster biosynthesis, namely CpFd1, IscU, IscS, mt-HSP70, mtFNR and frataxin, have been reported in *Cryptosporidium*[20], and we were not surprised to observe proteins involved in electron transport. We used CDART[21] to identify [Fe–S] domains in HscB (JAC) and ATM1, which are possibly involved in chaperonin activity of Hsp40/DnaJ type and ABC transport. Thus, *C. hominis*, like the microsporidian *Encephalitozoon cuniculi*[22], another obligate intracellular parasite, contains a minimal set of these proteins. These results imply significant mitochondrial function in *C. hominis* and indicate that the previously reported organelle[19] is an atypical mitochondrion.

Cryptosporidium apparently lacks an apicoplast[13,23], and searches of the *C. hominis* genome identified no apicoplast-encoded genes (Supplementary Table 9). Some putative nuclear-encoded apicoplast genes, for example acetyl-CoA carboxylase 1 precursor[24] and adenylyl cyclase[25], are present. Others, such as the apicoplast 50S ribosomal protein L33 and the ribosomal L28 and S9 precursor proteins, were not found. The data indicate that *Cryptosporidium* lost an ancestral apicoplast. The presence of D-glucose-6-phosphate ketol-isomerase and 2-phospho-D-glycerate hydrolase, which are similar to plant genes and may be derived from ancient algal endosymbionts, is also indicative that engulfment of the alga that gave rise to the apicoplast preceded the divergence of *Cryptosporidium* from other apicomplexans. One hypothesis is that the acquisition of the type 1 FAS by a progenitor organism obviated the fatty-acid synthesis capabilities of the apicoplast[14,15].

The *C. hominis* genome encodes multiple proteins specific for components of the apical complex including micronemes and rhoptries (Supplementary Table 9). No specific dense granule-associated proteins were observed, probably because these proteins diverge rapidly[26]. However, proteins implicated in the regulation of transport and enhancement of the release of dense granule proteins[27] are present. As for *Plasmodium*, a typical Golgi structure is not apparent in *C. hominis*[23]. However, the presence of secretory organelles implies the existence of a functional endoplasmic reticulum and Golgi, and *C. hominis* encodes proteins similar to many related components, including the NSF/SNAP/SNARE/Rab machinery, which participates in dense granule release[28], and the rhoptry biogenesis mediator activator protein 1, involved in endoplasmic-reticulum–Golgi-organelle protein traffic[29]. The endoplasmic-reticulum–Golgi-organelle machinery of *C. hominis* therefore seems similar to that of other apicomplexans.

As described above, *C. hominis* exhibits limited biosynthetic capabilities and is apparently dependent on its ability to import essential nutrients such as amino acids, nucleotides and simple sugars. The genome encodes more than 80 genes with strong similarity to known transporters and several hundred genes with transporter-like properties. At least 12 sugar or nucleotide-sugar transporters, five putative amino-acid transporters, three fatty-acid transporters, 23 ABC family transporters including possible multiple-drug-resistance proteins, and several putative mitochondrial transporters are present. Other putative transporters for choline uptake, aminophospholipid transport, ATP/ADP, and others with unclear function, were also identified. These transporters are ideal therapeutic targets (Supplementary Table 10).

Comparison of the genomes of *C. hominis* and *C. parvum* (Fig. 1, Supplementary Table 11) showed that the two genomes are very similar, exhibiting only 3–5% sequence divergence with no large insertions, deletions (Supplementary Fig. 11) or rearrangements evident. In fact, the gene complements of the two species are essentially identical because the few *C. parvum* genes not found in *C. hominis* are proximal to known sequence gaps (Supplementary Table 1). We therefore conclude that the significant phenotypic differences between these parasites are due to functionally significant polymorphisms in relevant protein-coding genes and to subtle gene regulatory differences.

A striking feature of the *C. hominis* genome is the concordance between its gene complement and the metabolic requirements in the environmental niches of its two primary life-cycle stages—the quiescent oocyst in the nutrient-poor aerobic environment of contaminated water, and the vegetative parasites in the nutrient-rich anaerobic or microaerophilic environment of the host. Oocysts probably persist by aerobically metabolizing stores of complex carbohydrates by means of glycolysis and the alternative electron transport system in the unconventional mitochondrion. Consistent with the lack of the energy-generating TCA cycle, oxidative phosphorylation, β-oxidation and the pentose phosphate pathways is the observation that oocysts are relatively inactive, and the two ATP per glucose from glycolysis can provide sufficient energy. In the host, the parasite can import sugars to fuel glycolysis directly, netting two ATP per hexose. In limiting glucose, an additional two ATP per hexose can be generated either by converting acetyl-CoA to acetate or by means of glycerol metabolism. The residual mitochondrion lacks the TCA cycle and oxidative phosphorylation as expected in an organism that replicates in anaerobic or microaerophilic environments, and a simplified electron transport system for regenerating reducing power is available. Thus, a glycolysis-based metabolism is sufficient to support *Cryptosporidium* in all life-cycle stages.

As previously noted, our analysis shows that *Cryptosporidium* is a mosaic of sequences from diverse progenitors, including the hypothetical endosymbiont alga that formed the apicoplast, the mitochondrion and numerous genes acquired from prokaryotes by lateral transfer. *Cryptosporidium* also exhibits modular gene loss. We assume, on the basis of inference from other apicomplexans and earlier diverging groups such as the Euglenozoa, the Heterolobosea and the jakobids[30], that *Cryptosporidium* progenitors exhibited the TCA cycle, β-oxidation, oxidative phosphorylation, amino acid, nucleotide and sugar biosynthesis, fully competent mitochondria, and a functional apicoplast. Genes associated with these functions are dispersed throughout the genome in *Plasmodium* and, we assume, in the progenitor. However, these systems seem to have been deleted cleanly in *Cryptosporidium*, leaving few identifiable residual genes or pseudogenes. Thus, the *Cryptosporidium* genome is a mosaic resulting from multiple lateral gene transfers and selective gene deletion.

The tailored physiology of *C. hominis* indicates attractive therapeutic targets (Supplementary Table 10), for example: essential transport systems; components of glycolysis; the unique prokaryotic FAS1 and PKS1; starch and amylopectin metabolism; nucleic-acid or amino-acid metabolism; the AOX electron transport system; the bifunctional thymidylate synthase–dihydrofolate reductase; and the diverged polyamine synthesis enzymes. Finally, many potential vaccine targets were identified in the *C. hominis* genome (not shown), and, in contrast with other protozoan parasites, no extensive arrays of potentially variant surface proteins were observed, indicating a possible role for immunoprophylaxis for cryptosporidiosis.

The availability of the genome sequence of the human pathogen *C. hominis* is a crucial step forward in our understanding of the biology of this parasite. The gene complement provides very significant insight into its physiology and metabolism, validating previous hypotheses and indicating the possibility of others. New obvious targets for chemotherapy and immunotherapy are already apparent. In short, we expect that the availability of the sequence of

C. hominis will stimulate progress in research on this organism and its pathogenicity, and strategies for intervention in the diseases it causes. □

Methods

A modified whole-genome shotgun strategy was used to sequence the ~9.2-Mb genome of *C. hominis* isolate TU502, which was derived from an infected child from Uganda. DNA was purified from surface-sterilized oocysts, shotgun and BAC clones were constructed, and end sequences were generated. About 220,000 sequence reads from small insert clones, and end sequences from ~2,000 BAC clones averaging ~35 kbp in size, were generated. The data represents a ~12-fold shotgun clone coverage of the genome with a quality score of Phred 20, and a 7–8-fold coverage with BAC clones. The sequences were assembled with Phrap, yielding a ~9.16-Mb assembly, which was structurally and functionally analysed with a variety of available software programs and in-house scripts (see Supplementary Text 1 and 2 for further details and references).

Received 4 May; accepted 6 August 2004; doi:10.1038/nature02977.

1. Tzipori, S. & Ward, H. Cryptosporidiosis: biology, pathogenesis and disease. *Microbes Infect.* **4,** 1047–1058 (2002).
2. Morgan-Ryan, U. M. *et al. Cryptosporidium hominis* n. sp. (Apicomplexa: Cryptosporidiidae) from *Homo sapiens*. *J. Eukaryot. Microbiol.* **49,** 433–440 (2002).
3. Akiyoshi, D. E., Feng, X., Buckholt, M. A., Widmer, D. & Tzipori, S. Genetic analysis of *Cryptosporidium parvum*: human genotype 1 isolate passaged through different host species. *Infect. Immun.* **70,** 5670–5675 (2002).
4. Bankier, A. T. *et al.* Integrated mapping, chromosomal sequencing and sequence analysis of *Cryptosporidium parvum*. *Genome Res.* **13,** 1787–1799 (2003).
5. Abrahamsen, M. S. *et al.* Complete genome sequence of the apicomplexan, *Cryptosporidium parvum*. *Science* **304,** 441–445 (2004).
6. Gardner, M. J. *et al.* Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* **419,** 498–511 (2002).
7. Muller, M. in *Evolutionary Relationships among Protozoa* (eds Coombs, G. H., Vickerman, K., Sleigh, M. A. & Warren, A.) 109–131 (Kluwer Academic, Dordrecht, 1998).
8. Sanchez, L. B., Galperin, M. Y. & Muller, M. Acetyl-CoA synthetase from the amitochondriate eukaryote *Giardia lamblia* belongs to the newly recognized superfamily of acyl-CoA synthetases (Nucleoside diphosphate-forming). *J. Biol. Chem.* **275,** 5794–5803 (2000).
9. Jasso-Chavez, R. & Moreno-Sanchez, R. Cytosol-mitochondria transfer of reducing equivalents by a lactate shuttle in heterotrophic *Euglena*. *Eur. J. Biochem.* **270,** 4942–4951 (2003).
10. Petry, F. & Harris, J. R. Ultrastructure, fractionation and biochemical analysis of *Cryptosporidium parvum* sporozoites. *Int. J. Parasitol.* **29,** 1249–1260 (1999).
11. Striepen, B. *et al.* Gene transfer in the evolution of parasite nucleotide biosynthesis. *Proc. Natl Acad. Sci. USA* **101,** 3154–3159 (2004).
12. Atreya, C. E. & Anderson, K. S. Kinetic characterization of bifunctional thymidylate synthase–dihydrofolate reductase (TS–DHFR) from Cryptosporidium hominis: A paradigm shift for TS activity and channeling behavior. *J. Biol. Chem.* **279,** 18314–18322 (2004).
13. Zhu, G., Marchewka, M. J. & Keithly, J. S. *Cryptosporidium parvum* appears to lack a plastid genome. *Microbiol.* **146,** 315–321 (2000).
14. Zhu, G. *et al.* Expression and functional characterization of a giant Type I fatty acid synthase (CpFAS1) gene from *Cryptosporidium parvum*. *Mol. Biochem. Parasitol.* **134,** 127–135 (2004).
15. Zhu, G. *et al. Cryptosporidium parvum*: the first protist known to encode a putative polyketide synthase. *Gene* **298,** 79–89 (2002).
16. Priest, J. W., Xie, L. T., Arrowood, M. J. & Lammie, P. J. The immunodominant 17-kDa antigen from *Cryptosporidium parvum* is glycosylphosphatidylinositol-anchored. *Mol. Biochem. Parasitol.* **113,** 117–126 (2001).
17. Keithly, J. S. *et al.* Polyamine biosynthesis in *Cryptosporidium parvum* and its implications for chemotherapy. *Mol. Biochem. Parasitol.* **88,** 35–42 (1997).
18. Parsons, M. & Ruben, L. Pathways involved in environmental sensing in trypanosomatids. *Parasitol. Today* **16,** 56–62 (2000).
19. Riordan, C. E., Ault, J. G., Langreth, S. G. & Keithly, J. S. *Cryptosporidium parvum* Cpn60 targets a relict organelle. *Curr. Genet.* **44,** 138–147 (2003).
20. LaGier, M. J., Tachezy, J., Stejskal, F., Kutisova, K. & Keithly, J. S. Mitochondrial-type iron-sulfur cluster biosynthesis genes (IscS and IscU) in the apicomplexan *Cryptosporidium parvum*. *Microbiol.* **149,** 3519–3530 (2003).
21. Geer, L. Y., Domrachev, M., Lipman, D. J. & Bryant, S. H. CDART: protein homology by domain architecture. *Genome Res.* **12,** 1619–1623 (2002).
22. Katinka, M. D. *et al.* Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*. *Nature* **414,** 450–453 (2001).
23. Tetley, L., Brown, S. M., McDonald, V. & Coombs, G. H. Ultrastructural analysis of the sporozoite of *Cryptosporidium parvum*. *Microbiol.* **144,** 3249–3255 (1998).
24. Zuther, E., Johnson, J. J., Haselkorn, R., McLeod, R. & Gornicki, P. Growth of *Toxoplasma gondii* is inhibited by aryloxyphenoxypropionate herbicides targeting acetyl-CoA carboxylase. *Proc. Natl Acad. Sci. USA* **96,** 13387–13392 (1999).
25. Muhia, D. K. *et al.* Multiple splice variants encode a novel adenylyl cyclase of possible plastid origin expressed in the sexual stage of the malaria parasite *Plasmodium falciparum*. *J. Biol. Chem.* **278,** 22014–22022 (2003).
26. Beyer, T. V., Svezhova, N. V., Radchenko, A. I. & Sidorenko, N. V. Parasitophorous vacuole: morphofunctional diversity in different coccidian genera (a short insight into the problem). *Cell Biol. Int.* **26,** 861–871 (2002).
27. Stedman, T. T., Sussmann, A. R. & Joiner, K. A. *Toxoplasma gondii* Rab6 mediates a retrograde pathway for sorting of constitutively secreted proteins to the Golgi complex. *J. Biol. Chem.* **278,** 5433–5443 (2003).
28. Chaturvedi, S. *et al.* Constitutive calcium-independent release of *Toxoplasma gondii* dense granules occurs through the NSF/SNAP/SNARE/Rab machinery. *J. Biol. Chem.* **274,** 2424–2431 (1999).
29. Ngo, H. M. *et al.* AP-1 in *Toxoplasma gondii* mediates biogenesis of the rhoptry secretory organelle from a post-Golgi compartment. *J. Biol. Chem.* **278,** 5343–5352 (2003).
30. Simpson, A. G. *et al.* Evolutionary history of 'early-diverging' eukaryotes: the excavate taxon Carpediemonas is a close relative of *Giardia*. *Mol. Biol. Evol.* **19,** 1782–1791 (2002).

Supplementary Information accompanies the paper on **www.nature.com/nature**.

Acknowledgements We thank S. Hendricks, M. R. C. Carvalho, S. Tula, G. Kazanina, J. Power, A. Holzgrefe, N. Ebashi, E. Butt, B. Sutton, S. Millett, W. Vogel and B. Constance for their technical contributions to this project and J. Elhai, D. Mallonee, L. M. Wen, T. Zwierzynski and Z. Chen for their contributions to the planning and performance of this project. This research was supported by grants from the National Institutes of Health.

Author contributions S.T. and G.A.B. contributed equally.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to G.A.B. (buck@mail2.vcu.edu) or S.T. (saul.tzipori@tufts.edu). The sequences reported in this paper have been deposited in GenBank under the project accession number AAEL000000. Further details of the accession numbers are available in the Supplementary Information.

MYC inactivation uncovers pluripotent differentiation and tumour dormancy in hepatocellular cancer

Catherine M. Shachaf[1], Andrew M. Kopelman[1], Constadina Arvanitis[1], Åsa Karlsson[1], Shelly Beer[1], Stefanie Mandl[2], Michael H. Bachmann[2], Alexander D. Borowsky[3], Boris Ruebner[3], Robert D. Cardiff[3], Qiwei Yang[1], J. Michael Bishop[4], Christopher H. Contag[2] & Dean W. Felsher[1]

[1]*Division of Medical Oncology, Departments of Medicine and Pathology,*
[2]*Department of Pediatrics, Stanford University, California 94305, USA*
[3]*Department of Pathology, University of California Davis Medical Center, Davis, California 95616, USA*
[4]*G. W. Hooper Foundation, University of California, San Francisco, California 94143, USA*

Hepatocellular carcinoma is generally refractory to clinical treatment[1]. Here, we report that inactivation of the MYC oncogene is sufficient to induce sustained regression of invasive liver cancers. MYC inactivation resulted *en masse* in tumour cells differentiating into hepatocytes and biliary cells forming bile duct structures, and this was associated with rapid loss of expression of the tumour marker α-fetoprotein, the increase in expression of liver cell markers cytokeratin 8 and carcinoembryonic antigen, and in some cells the liver stem cell marker cytokeratin 19. Using *in vivo* bioluminescence imaging we found that many of these tumour cells remained dormant as long as MYC remain inactivated; however, MYC reactivation immediately restored their neoplastic features. Using array comparative genomic hybridization we confirmed that these dormant liver cells and the restored tumour retained the identical molecular signature and hence were clonally derived from the tumour cells. Our results show how oncogene inactivation may reverse tumorigenesis in the most clinically difficult cancers. Oncogene inactivation uncovers the pluripotent capacity of tumours to differentiate into normal cellular lineages and tissue structures, while retaining their latent potential to become cancerous, and hence existing in a state of tumour dormancy.

Cancer is largely caused by genomic catastrophes that result in the activation of proto-oncogenes and/or inactivation of tumour-suppressor genes[2]. Even brief inactivation of a single oncogene can

Data collection and processing

A single wavelength anomalous diffraction experiment was carried out at the CuKα wavelength on crystals cryoprotected in mother liquor plus 25% 2-methyl-2,4-pentanediol; clear diffraction was observed to at least 1.8 Å resolution. We indexed, integrated and scaled the data using D*TREK[25], identified heavy atom sites with SOLVE[26], and carried out refinement with CNS[27] to obtain a model with final R_{xtal} and R_{free} values of 17.8% and 22.8%, respectively. The model contains all RNA atoms except A82, for which electron density was not observed, along with the hypoxanthine ligand. For additional experimental details and references, see Supplementary Information.

Received 28 July; accepted 13 September 2004; doi:10.1038/nature03037.

1. Vitreschak, A. G., Rodionov, D. A., Mironov, A. A. & Gelfand, M. S. Riboswitches: the oldest mechanism for the regulation of gene expression? *Trends Genet.* **20,** 44–50 (2004).
2. Mandal, M. & Breaker, R. R. Gene regulation by riboswitches. *Nature Rev. Mol. Cell. Biol.* **5,** 451–463 (2004).
3. Mandal, M., Boese, B., Barrick, J. E., Winkler, W. C. & Breaker, R. R. Riboswitches control fundamental biochemical pathways in *Bacillus subtilis* and other bacteria. *Cell* **113,** 577–586 (2003).
4. Johansen, L. E., Nygaard, P., Lassen, C., Agerso, Y. & Saxild, H. H. Definition of a second *Bacillus subtilis pur* regulon comprising the *pur* and *xpt–pbuX* operons plus *pbuG*, *nupG* (*yxjA*), and *pbuE* (*ydhL*). *J. Bacteriol.* **185,** 5200–5209 (2003).
5. Ellington, A. D. & Szostak, J. W. *In vitro* selection of RNA molecules that bind specific ligands. *Nature* **346,** 818–822 (1990).
6. Gold, L., Polisky, B., Uhlenbeck, O. & Yarus, M. Diversity of oligonucleotide functions. *Annu. Rev. Biochem.* **64,** 763–797 (1995).
7. Silverman, S. K. Rube Goldberg goes (ribo)nuclear? Molecular switches and sensors made from RNA. *RNA* **9,** 377–383 (2003).
8. Seetharaman, S., Zivarts, M., Sudarsan, N. & Breaker, R. R. Immobilized RNA switches for the analysis of complex chemical and biological mixtures. *Nature Biotechnol.* **19,** 336–341 (2001).
9. Zimmermann, G. R., Jenison, R. D., Wick, C. L., Simmorre, J.-P. & Pardi, A. Interlocking structural motifs mediate molecular discrimination by a theophylline-binding RNA. *Nature Struct. Biol.* **4,** 644–649 (1997).
10. Fan, P., Suri, A. K., Fiala, R., Live, D. & Patel, D. J. Molecular recognition in the FMN–RNA aptamer complex. *J. Mol. Biol.* **258,** 480–500 (1996).
11. Baugh, C., Grate, D. & Wilson, C. 2.8 Å crystal structure of the malachite green aptamer. *J. Mol. Biol.* **301,** 117–128 (2000).
12. Koizumi, M., Soukup, G. A., Kerr, J. N. & Breaker, R. R. Allosteric selection of ribozymes that respond to the second messengers cGMP and cAMP. *Nature Struct. Biol.* **6,** 1062–1071 (1999).
13. Leulliot, N. & Varani, G. Current topics in RNA–protein recognition: control of specificity and biological function through induced fit and conformational capture. *Biochemistry* **40,** 7947–7956 (2001).
14. Williamson, J. R. Induced fit in RNA–protein recognition. *Nature Struct. Biol.* **7,** 834–837 (2000).
15. Mandal, M. & Breaker, R. R. Adenine riboswitches and gene activation by disruption of a transcription terminator. *Nature Struct. Mol. Biol.* **11,** 29–35 (2004).
16. Doherty, E. A., Batey, R. T., Masquida, B. & Doudna, J. A. A universal mode of helix packing in RNA. *Nature Struct. Biol.* **8,** 339–343 (2001).
17. Nissen, P., Ippolito, J. A., Ban, N., Moore, P. B. & Steitz, T. A. RNA tertiary interactions in the large ribosomal subunit: the A-minor motif. *Proc. Natl Acad. Sci. USA* **98,** 4899–4903 (2001).
18. Cate, J. H. *et al.* RNA tertiary structure mediation by adenosine platforms. *Science* **273,** 1696–1699 (1996).
19. Correll, C. C., Beneken, J., Plantinga, M. J., Lubbers, M. & Chan, Y. L. The common and the distinctive features of the bulged-G motif based on a 1.04 Å resolution RNA structure. *Nucleic Acids Res.* **31,** 6806–6818 (2003).
20. Molinaro, M. & Tinoco, I. Jr Use of ultra stable UNCG tetraloop hairpins to fold RNA structures: thermodynamic and spectroscopic applications. *Nucleic Acids Res.* **23,** 3056–3063 (1995).
21. De la Pena, M., Gago, S. & Flores, R. Peripheral regions of natural hammerhead ribozymes greatly increase their self-cleavage activity. *EMBO J.* **22,** 5561–5570 (2003).
22. Khvorova, A., Lescoute, A., Westhof, E. & Jayasena, S. D. Sequence elements outside the hammerhead ribozyme catalytic core enable intracellular activity. *Nature Struct. Biol.* **10,** 708–712 (2003).
23. Choi, K. Y. & Zalkin, H. Structural characterization and corepressor binding of the *Escherichia coli* purine repressor. *J. Bacteriol.* **174,** 6207–6214 (1992).
24. Kieft, J. S. & Batey, R. T. A general method for rapid and nondenaturing purification of RNAs. *RNA* **10,** 988–995 (2004).
25. Pflugrath, J. W. The finer things in X-ray diffraction data collection. *Acta Crystallogr. D* **55,** 1718–1725 (1999).
26. Terwilliger, T. SOLVE and RESOLVE: automated structure solution, density modification and model building. *J. Synchrotron Radiat.* **11,** 49–52 (2004).
27. Brunger, A. T. *et al.* Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54,** 905–921 (1998).
28. Schumacher, M. A., Choi, K. Y., Zalkin, H. & Brennan, R. G. Crystal structure of LacI member, PurR, bound to DNA: minor groove binding by α helices. *Science* **266,** 763–770 (1994).

Supplementary Information accompanies the paper on **www.nature.com/nature**.

Acknowledgements We thank S. Edwards for maintaining and managing the Biochemistry Division X-ray Crystallography facility; and T. Cech, A. Pardi, D. Wuttke, J. Kieft and R. Rambo for discussions and comments on the manuscript. This work was funded in part from a grant from the Research Corporation and the University of Colorado Butcher Biotechnology Initiative. S.D.G. was supported in part by a NIH predoctoral training grant.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to R.T.B. (robert.batey@colorado.edu). The atomic coordinates and structure factors have been deposited in the RCSB Protein Data Bank under accession number 1U8D.

corrigendum

The genome of *Cryptosporidium hominis*

Ping Xu, Giovanni Widmer, Yingping Wang, Luiz S. Ozaki, Joao M. Alves, Myrna G. Serrano, Daniela Puiu, Patricio Manque, Donna Akiyoshi, Aaron J. Mackey, William R. Pearson, Paul H. Dear, Alan T. Bankier, Darrell L. Peterson, Mitchell S. Abrahamsen, Vivek Kapur, Saul Tzipori & Gregory A. Buck

Nature **431,** 1107–1112 (2004).

The GenBank accession number was supplied incorrectly as AAEL000000. The sentence should have read: 'The sequences reported in this paper have been deposited in GenBank under the project accession number AAEL00000000.'. In addition, the received date should have been 14 March 2004. □