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 then 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[27]. Hydroxogossenes (5.5 mg protein) were extracted and resuspended in 150 μl 99% 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 program (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 HMMAIGN. 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 model of amino acid substitution. The first 5,000 generations were discarded as burn-in. Consensus trees satisfying the more than 50 majority rule were run for 100,000 generations, with sampling every 50 generations. The genome sequence of Thermus thermophilus was used as an outgroup.

Evolutionary relationships of NuoF and NuoE in a selection of organisms were explored with Treeview, and probabilities of branch partitions were calculated. The genome tree was rooted with an outgroup of the closest homologs of NuoF and NuoE.

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.I.J. and C.F.C.) and a National Aeronautics and Space Administration Astrobiology grant to UCLA. P.I.J. is a Burroughs Welcome 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.I.J. (johnsonp@ucla.edu).

The authors declare that they have no competing financial interests.

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

C. parvum also infects other mammals. Here we describe the eight-chromosome ~9.2-million-base genome of C. hominis. 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 may 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 and chromosomes of the C. parvum genome 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, respectively). The density of 25–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, TGGGGCCA and TGGATGGA, over-represented in other apicomplexans, 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 55 rRNA genes are present. As in P. falciparum, two tRNA elements 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 (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.

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 (acyt-CoA) and NADPH, which is then reduced to NADP+, releasing hydrogen, by a Narf-like [Fe]-hydrogenase, as in Giardia, 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, passing electrons from malate to an electron transport system composed of elements of Complexes I and III and an alternative oxidase system with O2 as electron acceptor and producing no additional ATP.

Enzymes for metabolism of glycogen, starch and amylpectin are present, which is consistent with suggestions that amylpectin represents an energy reserve for sporozoites. Loss of glucose-6-phosphate 1-dehydrogenase and other enzymes of the pentose phosphate pathway suggests that, unlike P. falciparum and other apicomplexans, 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—such as 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\textsuperscript{1} 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 \textit{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 \textit{C. hominis}, and, unlike \textit{P. falciparum}, \textit{C. hominis} has a large complement of amino acid transporters.

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

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{Figure1.png}
\caption{Schematic representation of the \textit{C. hominis} chromosomes. Tracks indicate \textit{C. hominis} contigs (blue), sequence gaps (white) and physical gaps (red) (a); HAPPY\textsuperscript{a} markers (b); positions of the octamers TGCGCCA (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 \textit{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 \textit{C. parvum} sequences (red triangles show sequence gaps), with the first base at the top.
}
\end{figure}
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. As previously suggested, 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*, 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.

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.

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. 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 adenylyl cyclase, cyclic-AMP phosphodiesterase and protein kinase A, indicating the presence of the cAMP-mediated signalling pathway. Trimeric G protein, often involved in the activation of cAMP-mediated signalling, was not found in *C. hominis*, indicating that, as in Kinoplastida 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,
enzymes associated with acidocalcisomes, and calmodulin imply that \( \text{Ca}^{2+} \) transport and sequestering are functional. Protein kinase C receptors indicate that \textit{C. hominis} has the ability to signal by activation of soluble cytoplasmic receptor-associated kinases.

No mitochondrial DNA sequences were found in \textit{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 cardioloipin and performs some related mitochondrial functions, and mitochondrial marker chaperonin 60 was localized to this structure\(^2\). Core enzymes of [Fe–S] cluster biosynthesis, namely CpFd1, LscU, LscS, mt-HSP70, mtFNR and frataxin, have been reported in 

\textit{Cryptosporidium}\(^5\,8\), and we were not surprised to observe proteins involved in electron transport. We used CDART\(^7\) 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, \textit{C. hominis}, like the microsporidian \textit{Encephalitozoon cuniculi}\(^2\), another obligate intracellular parasite, contains a minimal set of these proteins. These results imply significant mitochondrial function in \textit{C. hominis} and indicate that the previously reported organelle\(^9\) is an atypical mitochondrion.

\textit{Cryptosporidium} apparently lacks an apicoplast\(^1\,2\,3\), and searches of the \textit{C. hominis} genome identified no apicoplast-encoded genes (Supplementary Table 9). Some putative nuclear-encoded apicoplast genes, for example acetyl-CoA carboxylase 1 precursor and adenylyl cyclase\(^2\,2\), 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 

\textit{Cryptosporidium} lost an ancestral apicoplast. The presence of d-glucose-6-phosphate ketol-isomerase and 2-phospho-D-glycerate hydrolyase, 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 

\textit{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\(^1\,15\).

The \textit{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\(^2\). However, proteins implicated in the regulation of transport and enhancement of the release of dense granule proteins\(^2\) are present. As for \textit{Plasmodium}, a typical Golgi structure is not apparent in \textit{C. hominis}\(^7\). However, the presence of secretory organelles implies the existence of a functional endoplasmic reticulum and Golgi, and \textit{C. hominis} encodes proteins similar to many related components, including the NSF/SNAP/SNARE/Rab machinery, which participates in dense granule release\(^2\), and the rhoptry biogenesis mediator activator protein 1, involved in endoplasmic-reticulum–Golgi-organelle protein trafficking. The endoplasmic-reticulum–Golgi-organelle machinery of \textit{C. hominis} therefore seems similar to that of other apicomplexans.

As described above, \textit{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 \textit{C. hominis} and \textit{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 \textit{C. parvum} genes not found in \textit{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 \textit{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, \(\beta\)-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 \textit{Cryptosporidium} in all life-cycle stages.

As previously noted, our analysis shows that \textit{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. \textit{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\(^9\), that \textit{Cryptosporidium} progenitors exhibited the TCA cycle, \(\beta\)-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 \textit{Plasmodium} and, we assume, in the progenitor. However, these systems seem to have been deleted cleanly in \textit{Cryptosporidium}, leaving few identifiable residual genes or pseudogenes. Thus, the \textit{Cryptosporidium} genome is a mosaic resulting from multiple lateral gene transfers and selective gene deletion.

The tailored physiology of \textit{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 \textit{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 \textit{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 a ∼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 kb 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.

Acknowledgements
We thank S. Hendricks, M. R. C. Carvalho, S. Tula, G. Kazanina, J. Power, A. Holzgreve, N. Ebashi, E. Butt, B. Sutton, S. Millet, W. Vogel and B. Constance for their technical contributions to this project and J. Elhai, D. Mallone, L. M. Wen, T. Zwierszycki 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.ucdavis.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 tumor dormancy in hepatocellular cancer

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

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

Hepatocellular carcinoma is generally refractory to clinical treatment. Here, we report that inactivation of the MYC oncogene is sufficient to induce sustained regression of invasive liver cancers. MYC inactivation resulted in massive 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 tumorgenesis in the most clinically difficult cancers. Oncogene inactivation uncovers the pluripotent capacity of cancer cells 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. Even brief inactivation of a single oncogene can