Published in final edited form as:

J Nutr Biochem. 2013 November ; 24(11): 1963–1969. doi:10.1016/j.jnutbio.2013.06.007.

Three promoters regulate the transcriptional activity of the human *holocarboxylase synthetase* gene^{1,2}

Mengna Xia¹, Sridhar A. Malkaram^{1,3}, and Janos Zemleni^{1,*}

¹Department of Nutrition and Health Sciences, University of Nebraska-Lincoln, Lincoln, NE, 68583-0806

Abstract

Holocarboxylase synthetase (HLCS) is the only protein biotin ligase in the human proteome. HLCS-dependent biotinylation of carboxylases plays crucial roles in macronutrient metabolism. HLCS appears to be an essential part of multiprotein complexes in the chromatin that cause gene repression and contribute toward genome stability. Consistent with these essential functions, HLCS knockdown causes strong phenotypes including shortened life span and low stress resistance in *Drosophila melanogaster*, and de-repression of long-terminal repeats in humans, other mammalian cell lines, and *Drosophila*. Despite previous observations that the expression of HLCS depends on biotin status in rats and in human cell lines, little is known about the regulation of HLCS expression. The goal of this study was to identify promoters that regulate the expression of the human *HLCS* gene. Initially, the human *HLCS* locus was interrogated *in silico* using predictors of promoters including sequences of *HLCS* mRNA and expressed sequence tags, CpG islands, histone marks denoting transcriptionally poised chromatin, transcription factor binding sites, and *DNaseI* hypersensitive regions. Our predictions revealed three putative *HLCS* promoters, denoted P1, P2, and P3. Promoters lacked a TATA box, which is typical for housekeeping genes. When the three promoters were cloned into a luciferase reporter plasmid, reporter gene activity was at least three times background noise in human breast, colon, and kidney cell lines; activities consistently followed the pattern P1 > P3 > P2. Promoter activity depended on the concentration of biotin in culture media, but the effect was moderate. We conclude that we have identified promoters in the human *HLCS* gene.

Keywords

Biotin; holocarboxylase synthetase; human; promoter

²Supported in part by funds provided through the Hatch Act. Additional support was provided by NIH grants DK063945, DK077816, and USDA CSREES grant 2006-35200-17138.

© 2013 Elsevier Inc. All rights reserved.

*To whom correspondence should be addressed. Phone: (402) 472 3270; fax: (402) 472 1587; jzemleni2@unl.edu. Address: Department of Nutrition and Health Sciences, University of Nebraska-Lincoln, 316C Ruth Leverton Hall, Lincoln, NE 68583-0806.

³Present address: Department of Biology, West Virginia University, Institute, WV 25112-1000.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Author disclosures: M. Xia, S.A. Malkaram and J. Zemleni, no conflicts of interest.

1. Introduction

Holocarboxylase Synthetase (HLCS) is the only enzyme in the human proteome capable of catalyzing the binding of biotin to proteins [1, 2]. HLCS is a monomeric protein encoded by a single gene localized on chromosome 21 (21q22.13, Genbank Ng_016193). Full-length HLCS comprises 726 amino acids with a predicted molecular weight of 81 kDa [3]. Three HLCS transcripts plus additional splicing variants originate in exons 1, 2, and 3 of the gene [3]; methionines (M) 1, 7, and 58 in exons 6 and 7 have been identified as possible translation start sites [4]. HLCS proteins, migrating with apparent sizes of 62, 64, 76, 82, and 86 kDa during gel electrophoresis, have been detected in human placenta and in bovine liver by using anti-HLCS. The 76-kDa, 82 kDa, and 86 kDa bands probably represent HLCS with translation start sites in M58, M7, and M1, respectively, whereas bands of lower molecular weight might have been caused by protein degradation [1, 4].

The "classical" role of HLCS is to catalyze the binding of biotin to five apocarboxylases. Acetyl-CoA carboxylases 1 and 2 are located in the cytoplasm and on the outer mitochondrial membrane, respectively; propionyl-CoA carboxylase, pyruvate carboxylase, and 3-methylcrotonyl-CoA carboxylase are located in the mitochondrial matrix [5]. These biotinylated (holo-)carboxylases catalyze key reactions in fatty acid synthesis, mitochondrial fatty acid import, degradation of odd-chain fatty acids, gluconeogenesis, and leucine catabolism, respectively. More recently, evidence has emerged suggesting that HLCS is not confined to the extranuclear space, but that HLCS also enters the nuclear compartment [6–9] and is part of mammalian chromatin [10–12]. HLCS isoforms with translational start sites in M1 and M58 are the preferred forms for nuclear import [8, 9]. Unambiguous evidence suggests that HLCS also catalyzes biotinylation of proteins other than carboxylases, including histone H3 [13] and the histone methyltransferase EHMT-1 [12]. Ongoing mass spectrometry studies in our laboratory revealed 108 novel proteins that contain covalently bound biotin (unpublished).

Consistent with the important roles of HLCS in intermediary metabolism, chromatin, and possible other functions, no living HLCS null individual has ever been reported, suggesting embryonic lethality. HLCS knockdown studies (~30% residual activity) produced phenotypes such as decreased life span and heat resistance in *Drosophila melanogaster* [10] and aberrant gene regulation in human cell lines [14–16]. Mutations in the human HLCS gene cause a substantial decrease in HLCS activity and metabolic abnormalities [17, 18]. Unless diagnosed and treated early, HLCS deficiency is uniformly fatal [19]. Single nucleotide polymorphisms in the *HLCS* gene decrease catalytic activity, but the biological importance of these polymorphisms is unknown [20].

Previous studies in rats and in human cell cultures suggest that the expression of HLCS depends on a sufficient supply of biotin [15, 21, 22]. However, these findings were not reproduced when investigating HLCS mRNA levels in lymphocytes from healthy adults where the inter-individual variation is large [23]. Little is known about the promoters in the human *HLCS* gene and their regulation, with the notable exception of a random screen of gene promoters on chromosome 21 by using reporter gene assays, which led to the tentative identification of a putative HLCS promoter [24]. The major objective of this study was to conduct a comprehensive analysis of HLCS promoters in the human genome. Supporting objectives included (a) the identification of transcriptional start sites, (b) the investigation of cell lineage specific activities of individual promoters, and (c) the identification of regulatory elements in these promoters.

2. Methods and Materials

2.1. Prediction of putative HLCS promoters and transcription factor binding sites

HLCS promoters were predicted using transcriptome, epigenome, and exon and promoter prediction algorithms. First, the tentative location of HLCS promoters was identified by using transcript information deposited in the ensembl genome browser [25], and the SwitchGear Genomics Transcription Start Sites (TSS) and Eponine Predicted TSS databases in the UCSC genome browser [26]. The predictions for transcriptional start sites were aligned with the following datasets and online tools to increase the confidence in predictions and to position the HLCS more accurately; all databases and tools were accessed through the UCSC genome browser (hg18, hg19) using the blat alignment tool, if applicable. (i) K4-trimethylated histone H3 (H3K4me3) and K27-acetylated H3 (H3K27ac) are epigenetic marks for transcriptionally active chromatin and are overrepresented in transcriptionally poised and active gene promoters and enhancers [27]; evidence suggests that H3K27ac distinguishes active enhancers from inactive/poised enhancer elements containing H3K4me1 alone [28, 29]. (ii) More than 50% of human genes initiate transcription from CpG dinucleotide-rich regions and these CpG islands are widely recognized markers of gene promoters [30]. CpG content was predicted using the normalized CpG fraction method spanning -1.5 kb through +1.5 kb of repeat-masked sequence from the putative TSS [31]. CpG fractions were calculated as described previously with repeats being masked; a value of 0.35 was used as a cut-off for distinguishing high and low CpG content promoters [32]; The expression of genes from high and low CpG content promoters depends on adjacent histone modifications [32]. (iii) The First-Exon and Promoter (FirstEF) Prediction tool recognizes structural and compositional features such as CpG islands, promoter regions and first splice-donor sites with 86% accuracy and was used for predicting HLCS promoters [33]. (iv) Expressed Sequence Tags (EST) are markers for gene transcription and were mapped in the HLCS locus using the EST database in the UCSC genome browser. In addition, GenBank was interrogated for entries regarding human HLCS mRNA. (v) RNA polymerase II (Pol II) binding sites are markers for transcriptionally poised promoters; pausing of Pol II pausing occurs at most promoters and is the major rate-limiting step in early transcription at these promoters [34]. Pol II binding sites were predicted using the UCSC ENCODE Pol II ChIP-seq tracks.

Transcription Factor binding sites were identified using the ENCODE Transcription Factor Binding Tracks in the UCSC genome browser [26, 35]: The data consist of 2229 ChIP-seq tracks generated by multiple participating laboratories using 196 distinct transcription factor specific antibodies. The transcription factor binding sites were computed as the presence of coverage peaks averaging over all ChIP-seq tracks specific to a transcription factor in 100 bp genomic intervals along the HLCS promoter region. Sequences were subjected to analysis by Consite or JASPAR to select for human transcription factors [36, 37]. A cut-off score of >70% was chosen, which is a conservative approach when predicting transcription factor binding sites on the plus and minus strand [38]. TSS were predicted using the Database of Transcriptional Start Sites [39]. *DNaseI* hypersensitive regions were mapped using the ENCODE Digital DNaseI Hypersensitivity Clusters Database; predictions are based on data obtained by Faire-seq, DNase-seq, and DNase Dgf protocols [40].

2.2. Creation of HLCS reporter gene constructs

Three HLCS promoters were predicted using the *in silico* approaches described above and were denoted P1, P2, and P3 (see *Results*). Genomic DNA was extracted from primary human IMR90 lung fibroblasts using DNeasy[®] spin columns (Qiagen). HLCS promoters were amplified by PCR using high fidelity Platinum[®] Pfx DNA Polymerase (Invitrogen) and the following primers (Online Supplementary Table 1): HLCS_P1_4399 and

HLCS_P1_5505, HLCS_P2_12877 and HLCS_P2_14901, and HLCS_P3_28055 and HLCS_P3_29315. PCR products were purified using 1% agarose gels. PCR products corresponding to P1, P2 and P3 were purified using the GenCatch™ Gel Extraction Kit (Epoch Life Science) and cloned into pSTBlue-1 to generate P1STBlue-1, P2STBlue-1, and P3STBlue-1, respectively. The sequences of all plasmids in this paper were confirmed by sequencing (Eurofins MWG Operon). P1STBlue-1 and P2STBlue-1 were digested using *MluI* and *HindIII*, and P3STBlue-1 was digested with *KpnI* and *HindIII* for subcloning into the pGL3 Basic Vector, thereby creating reporter gene plasmids HLCS-P1GL3₋₁₁₂₈, HLCS-P2GL3₋₂₀₅₄, and HCS-P3GL3₋₁₂₈₂.

Truncated reporter constructs were created using HLCS-P1GL3₋₁₁₂₈, HLCS-P2GL3₋₂₀₅₄, and HCS-P3GL3₋₁₂₈₂ as templates and by substituting various forward primers for HLCS_P1₋₁₁₂₈, HLCS_P2₋₂₀₅₄, and HLCS_P3₋₁₂₈₂ (Online Supplementary Table 1). PCR products were purified, cloned, and subcloned as described for the full-length constructs. Truncation constructs were denoted HLCS-P1GL3₋₈₅₁, HLCS-P1GL3₋₅₅₃, and HLCS-P1GL3₋₂₆₂ for P1, HLCS-P2GL3₋₁₇₂₄, HLCS-P2GL3₋₁₄₀₂, HLCS-P2GL3₋₁₁₆₇, HLCS-P2GL3₋₈₆₅, HLCS-P2GL3₋₅₂₅, and HLCS-P2GL3₋₂₄₉ for P2, and HCS-P3GL3₋₉₆₉, HCS-P3GL3₋₆₅₁, and HCS-P3GL3₋₃₁₃ for P3.

2.3. Cell transfection and reporter gene analysis

Human embryonic kidney carcinoma HEK293 cells, mammary carcinoma MCF7 cells, colon carcinoma HCT116 cells, and IMR90 fibroblasts were purchased from American Tissue Culture collection (Manassas, MD) and cultured following the repository's instructions. Primary IMR90 fibroblasts were obtained and processed at <20 passages to prevent effects of senescence. In some experiments, cells were cultured in biotin-defined media containing 25, 250, and 10,0pmol/l biotin, representing plasma levels observed in biotin-deficient, biotin-sufficient, and biotin-supplemented adults [41–43]. Biotin-defined media were prepared as described previously and cell viability was monitored routinely using the trypan blue exclusion test [43]. Cells were transfected with HLCS luciferase constructs and RSV- β -galactosidase (control for transfection efficiency) using TurboFect™ (Thermo Scientific, Inc.) as described previously and following the manufacturer's instructions [44]; baseline values were assessed by transfecting cells with the promoter-free pGL3 Basic vector and RSV- β -galactosidase. Luciferase and β -galactosidase activities were quantified as described previously [44]. Each data point represents the mean of three independent assays, each run in triplicate. In preliminary time-course studies, we identified the times post transfection when luciferase activities were maximal and used that time in all subsequent analyses. For MCF7 cells, activities peaked 72 h after transfection; for HEK293 and HCT116 cells, activities peaked 48 h after transfection. Reporter gene activities are expressed in units of luciferase activity in HLCS promoter-driven pGL3 compared with pGL3 Basic, and normalized by transfection efficiency.

2.4. Statistics

Outliers were identified using the Median Absolute Deviation procedure and excluded from further analysis. Data were assessed for heterogeneous variances using the Bartlett's test and, if applicable, log transformed prior to further analysis. One way ANOVA was used for data analysis; the Tukey-Kramer test was used for posthoc comparisons. JMP 9.0.0 (SAS Institute, Inc.) was used for statistical analysis. Differences were considered significant if $0.05 < p < 0.1$. Note that the more liberal cut off ($p < 0.1$) was applied only in the analysis of the data presented in Fig. 8 and is associated with an increased type I error rate (false positives). Data were expressed as means $\pm$ standard error.

3. Results

3.1. Bioinformatics prediction of HLCS promoters

Three putative promoters were predicted in the *HLCS* gene, when the human genome was analyzed using ensembl transcript data, the SwitchGear Genomics TSS database, the Eponine Predicted TSS database, and FirstEF (Fig. 1). Note that HLCS promoters were mapped based on GenBank ID NG_016193.1. All three promoters are located upstream of exon 4 in the *HLCS* gene and span sequences of 1128 bp (P1), 2054 bp (P2), and 1282 bp (P3). P1 is located in HLCS exon 1, P2 is located in exon 2, and P3 is located in intron 2/exon 3.

Next, confidence in predictions was strengthened by conducting a blat analysis of the human reference genome, mRNA, and EST the UCSC browser. Our queries positioned the three promoters in the genome (Fig. 2). Promoters P1, P2, and P3 had positions similar to those of multiple 5'-ends of mRNA and EST. These mRNA and EST have been detected in various tissues as per UCSC browser (Table 1).

The content of cytosines and guanines varies between 49% and 57% in P1, P2, and P3 (Table 2). P1 is a high CpG content promoter, whereas P2 and P3 are low CpG content promoters. Promoter methylation patterns are consistent in cell lines derived from various tissues, as assessed by the ENCODE HudsonAlpha Methyl-seq database (Fig. 3). The abundance and localization of CpG islands are consistent with the assumption that P1, P2, and P3 are gene promoters.

Additional evidence in support of our predictions was provided through bioinformatics analysis of histone modification marks. H3K4me3 is enriched at TSS [45–47]; H3K4me3 and H3K27ac are markers for transcriptionally active *low CpG content promoters* and *high CpG content promoters*, respectively [32, 48]. Distinct H3K4me3 peaks were visible for P1 and P3 from all five human cell lines that were analyzed using UCSC genome browser data, suggesting that these promoters are functional and transcriptionally active (Fig. 4). The abundance of H3K4me3 in promoters P1 and P3 was lower in GM12878 normal B-lymphocyte cells compared with other cell lines. The abundance of H3K4me3 was generally low in P2: H3K4me3 peaks were small, yet above background, in P2 in GM12878 normal B-lymphocyte cells, HEMC normal mammary epithelial cells, and HSMM normal skeletal muscle myoblasts xxx cells, and absent in HepG2 hepatocarcinoma cells and K562 leukemia cells. The H3K27ac mark was abundant in the P1 locus in all cell lines assessed; again the enrichment was considerably small in GM12878 cells compared with other cell lines. In P2, the H3K27ac mark was absent. Pexhibited a cell line specific pattern of H3K27ac enrichment: the mark was greatly enriched in K562 cells and produced a comparably small signal in the four other cell lines that were assessed.

Pol II binding sites in the HLCS genes are consistent with the pattern of histone modifications. Pol II was greatly enriched in P1 and P3, but produced only a weak signal in P2 (Fig. 5).

Transcription factor binding sites were overrepresented in the putative promoters compared with other regions in the *HLCS* gene, predicted using UCSC ENCODE Transcription factor Binding Chip-seq tracks (Online Supplementary Fig. 1). No TATA boxes were detected in any of the three promoters, which is indicative of housekeeping genes [49]. Databases for predicting transcription factor binding sites do not always produce consistent results. We increased the confidence of predictions by integrating data from three databases (UCSC ENCODE Transcription factor Binding Chip-seq tracks, Consite, and JASPAR) into an illustration that shows only transcription factor binding sites predicted in the three promoter

regions (Fig. 6). This procedure resulted in the identification of eight, one, and four transcription factor binding sites in P1, P2, and P3, respectively, between -191 and +449 bp from the TSS in P1, -160 bp to TSS in P2, -2 and -626 bp to TSS in P3. Note that Fig. 6 also marks the truncation sites for subsequent reporter gene assays. Transcription factor binding sites localized in *DNaseI* hypersensitive regions (Fig. 6 and Online Supplementary Figure 2).

3.2. Assessment of HLCS promoter activities by reporter gene constructs

All three putative promoters produced reporter gene activities produced distinct signals when MCT7 cells, HCT116 cells, and HEK293 cells were transfected with HLCS-P1GL3-₁₁₂₈, HLCS-P2GL3-₂₀₅₄, and HCS-P3GL3-₁₂₈₂ (Fig. 7, panels A–). The signals were at least three times the signal observed with empty vector pGL3 Basic. Promoter activity showed a consistent pattern across cell lines P1 >> P3 > P2. The activity of P1 exhibited a dose dependence regarding the concentration of biotin in culture media in HEK293 cells, but the effect was moderate and not statistically significant (Fig. 7, panel D).

Truncations of reporter constructs narrowed down regions in promoters that appear to be important for transcriptional activity. For example, when base pairs between HLCS_P1₄₉₇₄ and HLCS_P1₅₂₆₅ were deleted in P1, promoter activity decreased by about 50% (Fig. 8). That region contains two binding sites for AP2alpha, and one binding site each for YY1 and Max. When the E2F binding site was removed in P2, promoter activity decreased by 50%. However, the activity of P2 is low compared with P1 and P3; note that the x-axes are not drawn to scale in Fig. 8. The promoter activity decreased by 50% when binding sites for Max and HNF4A were deleted in P3.

4. Discussion

This paper characterizes three functional promoters in the human HLCS gene. P1 is identical to a putative HLCS promoter reported in a previous study of promoters on human chromosome 21 [24]. That study, however, was not designed to offer in-depth insights into any of the promoters located on chromosome 21, in contrast to the detailed bioinformatics and reporter gene analyses carried out in this paper. The discovery of functional promoters P2 and P3 further adds to the novelty of this paper. Confidence is high that P1 has greater activity than P3, and that the activity of P2 is fairly low across tissues. However, it is worthwhile noting that transcripts have been identified that can be linked with each of these promoters. Understanding the regulation of HLCS is a biologically meaningful advance in our knowledge of biotin metabolism, considering the strong phenotypes described for HLCS knockdown and the essential roles of biotinylated carboxylases and HLCS in macronutrient metabolism and epigenetic phenomena, respectively.

We caution that the regulation of HLCS expression also occurs at levels other than promoter activity including splicing and translational start sites [3, 4]. These processes affect the cellular distribution of HLCS. For example, evidence suggests that HLCS with translational start sites in M1 and M58 enter the nuclear compartment, whereas nuclear entry of HLCS with a start site in M7 is unlikely [8, 9]. In addition, it is entirely unclear whether posttranslational modifications of HLCS contribute toward the regulation of its catalytic activity. For example, *in silico* analysis of HLCS for possible modifications reveals the existence of several possible phosphorylation sites (not shown).

This report suggests that biotin has a fairly moderate effect on the activity of HLCS promoters, consistent with previous observations in various model systems. For example, the expression of HLCS depends on biotin availability in well-controlled and homogenous organisms such as rats and human cell cultures [21, 22], whereas no such effect can be

reproduced in a biotin-defined, yet heterogeneous population in an outpatient study in adults [50]. We propose that HLCS is a ubiquitously expressed housekeeping gene and that effects of biotin on promoter activity are small, consistent with the absence of a TATA box in any of the three HLCS promoters.

A few uncertainties remain for consideration in future studies, in addition to the posttranslational modifications described above. One such uncertainty is the effect of single nucleotide polymorphisms in HLCS promoters on transcriptional activity. We have conducted such studies in the coding region of the HLCS gene and identified a number of polymorphisms that impaired the catalytic activity of HLCS [20]. The majority of these polymorphisms lowered the affinity of HLCS for biotin and, hence, biotin supplementation is a fruitful approach to minimize the effects of those polymorphisms in afflicted individuals.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Chiba Y, Suzuki Y, Aoki Y, Ishida Y, Narisawa K. Purification and properties of bovine liver holocarboxylase synthetase. *Arch Biochem Biophys*. 1994; 313:8–14. [PubMed: 8053691]
2. Suzuki Y, Aoki Y, Ishida Y, Chiba Y, Iwamatsu A, Kishino T, et al. Isolation and characterization of mutations in the human holocarboxylase synthetase cDNA. *Nat Genet*. 1994; 8:122–128. [PubMed: 7842009]
3. Yang X, Aoki Y, Li X, Sakamoto O, Hiratsuka M, Kure S, et al. Structure of human holocarboxylase synthetase gene and mutation spectrum of holocarboxylase synthetase deficiency. *Hum Genet*. 2001; 109:526–534. [PubMed: 11735028]
4. Hiratsuka M, Sakamoto O, Li X, Suzuki Y, Aoki Y, Narisawa K. Identification of holocarboxylase synthetase (HCS) proteins in human placenta. *Biochim Biophys Acta*. 1998; 1385:165–171. [PubMed: 9630604]
5. Zemleni J, Wijeratne SS, Hassan YI. Biotin. *BioFactors* (Oxford, England). 2009; 35:36–46.
6. Narang MA, Dumas R, Ayer LM, Gravel RA. Reduced histone biotinylation in multiple carboxylase deficiency patients: a nuclear role for holocarboxylase synthetase. *Hum Mol Genet*. 2004; 13:15–23. [PubMed: 14613969]
7. Chew YC, Camporeale G, Kothapalli N, Sarath G, Zemleni J. Lysine residues in N- and C-terminal regions of human histone H2A are targets for biotinylation by biotinidase. *J Nutr Biochem*. 2006; 17:225–233. [PubMed: 16109483]
8. Bailey LM, Wallace JC, Polyak SW. Holocarboxylase synthetase: correlation of protein localisation with biological function. *Arch Biochem Biophys*. 2010; 496:45–52. [PubMed: 20153287]
9. Bao B, Wijeratne SS, Rodriguez-Melendez R, Zemleni J. Human holocarboxylase synthetase with a start site at methionine-58 is the predominant nuclear variant of this protein and has catalytic activity. *Biochem Biophys Res Commun*. 2011; 412:115–120. [PubMed: 21802411]
10. Camporeale G, Giordano E, Rendina R, Zemleni J, Eissenberg JC. Drosophila holocarboxylase synthetase is a chromosomal protein required for normal histone biotinylation, gene transcription patterns, lifespan and heat tolerance. *J Nutr*. 2006; 136:2735–2742. [PubMed: 17056793]
11. Singh D, Pannier AK, Zemleni J. Identification of holocarboxylase synthetase chromatin binding sites using the DamID technology. *Anal Biochem*. 2011; 413:55–59. [PubMed: 21303649]
12. Li Y, Hassan YI, Moriyama H, Zemleni J. Holocarboxylase synthetase interacts physically with euchromatic histone-lysine N-methyltransferase, linking biotinylation with histone methylation. *J Nutr Biochem*. 2013 (in press).
13. Bao B, Pestinger VIHY, Borgstahl GEO, Kolar C, Zemleni J. Holocarboxylase synthetase is a chromatin protein and interacts directly with histone H3 to mediate biotinylation of K9 and K18. *J Nutr Biochem*. 2011; 22:470–475. [PubMed: 20688500]

14. Chew YC, West JT, Kratzer SJ, Ilvarsonn AM, Eissenberg JC, Dave BJ, et al. Biotinylation of histones represses transposable elements in human and mouse cells and cell lines, and in *Drosophila melanogaster*. *J Nutr*. 2008; 138:2316–2322. [PubMed: 19022951]
15. Gralla M, Camporeale G, Zempleni J. Holocarboxylase synthetase regulates expression of biotin transporters by chromatin remodeling events at the SMVT locus. *J Nutr Biochem*. 2008; 19:400–408. [PubMed: 17904341]
16. Pestinger V, Wijeratne SSK, Rodriguez-Melendez R, Zempleni J. Novel histone biotinylation marks are enriched in repeat regions and participate in repression of transcriptionally competent genes. *J Nutr Biochem*. 2011; 22:328–333. [PubMed: 20691578]
17. Suzuki Y, Yang X, Aoki Y, Kure S, Matsubara Y. Mutations in the holocarboxylase synthetase gene HLCS. *Human Mutation*. 2005; 26:285–290. [PubMed: 16134170]
18. National Center for Biotechnology Information. [accessed 7/18/2008] Online Mendelian Inheritance in Man. <http://www.ncbi.nlm.nih.gov/sites/entrez?db=omim>
19. Thuy LP, Belmont J, Nyhan WL. Prenatal diagnosis and treatment of holocarboxylase synthetase deficiency. *Prenat Diagn*. 1999; 19:108–112. [PubMed: 10215065]
20. Esaki S, Malkaram SA, Zempleni J. Effects of single nucleotide polymorphisms in the human holocarboxylase synthetase gene on enzyme catalysis. *Eur J Hum Genet*. 2012; 20:428–433. [PubMed: 22027809]
21. Rodriguez-Melendez R, Cano S, Mendez ST, Velazquez A. Biotin regulates the genetic expression of holocarboxylase synthetase and mitochondrial carboxylases in rats. *J Nutr*. 2001; 131:1909–1913. [PubMed: 11435506]
22. Kaur Mall G, Chew YC, Zempleni J. Biotin requirements are lower in human Jurkat lymphoid cells but homeostatic mechanisms are similar to those of HepG2 liver cells. *J Nutr*. 2010; 140:1086–1092. [PubMed: 20357078]
23. Eng WK, Giraud D, Schlegel VL, Wang D, Lee BH, Zempleni J. Identification and assessment of markers of biotin status in healthy adults. *British Journal of Nutrition*. 2013 (in press).
24. Warnatz HJ, Querfurth R, Guerasimova A, Cheng X, Haas SA, Hufton AL, et al. Functional analysis and identification of cis-regulatory elements of human chromosome 21 gene promoters. *Nucleic Acids Res*. 2010; 38:6112–6123. [PubMed: 20494980]
25. EMBL-EBI. [accessed 2/20/2012] Ensembl Genomes. <http://www.ensemblgenomes.org/>
26. UCSC. [accessed 2/20/2012] UCSC Genome Browser. <http://www.genome.ucsc.edu>
27. Kouzarides, T.; Berger, SL. Chromatin modifications and their mechanism of action. In: Allis, CD.; Jenuwein, T.; Reinberg, D., editors. *Epigenetics*. Cold Spring Harbor, NY: Cold Spring Harbor Press; 2007. p. 191–209.
28. Heintzman ND, Hon GC, Hawkins RD, Kheradpour P, Stark A, Harp LF, et al. Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature*. 2009; 459:108–112. [PubMed: 19295514]
29. Creighton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci USA*. 2010; 107:21931–21936. [PubMed: 21106759]
30. Vavouri T, Lehner B. Human genes with CpG island promoters have a distinct transcription-associated chromatin organization. *Genome Biol*. 2012; 13:R110. [PubMed: 23186133]
31. Saxonov S, Berg P, Brutlag DL. A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc Natl Acad Sci USA*. 2006; 103:1412–1417. [PubMed: 16432200]
32. Karlic R, Chung HR, Lasserre J, Vlahovicek K, Vingron M. Histone modification levels are predictive for gene expression. *Proc Natl Acad Sci USA*. 2010; 107:2926–2931. [PubMed: 20133639]
33. Davuluri RV, Grosse I, Zhang MQ. Computational identification of promoters and first exons in the human genome. *Nat Genet*. 2001; 29:412–417. [PubMed: 11726928]
34. Core LJ, Waterfall JJ, Gilchrist DA, Fargo DC, Kwak H, Adelman K, et al. Defining the status of RNA polymerase at promoters. *Cell Rep*. 2012; 2:1025–1035. [PubMed: 23062713]
35. Encode Project Consortium. The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science*. 2004; 306:636–640. [PubMed: 15499007]

36. Sandelin A, Wasserman WW, Lenhard B. ConSite: web-based prediction of regulatory elements using cross-species comparison. *Nucleic Acids Res.* 2004; 32:W249–W252. [PubMed: 15215389]
37. Bryne JC, Valen E, Tang MH, Marstrand T, Winther O, da Piedade I, et al. JASPAR, the open access database of transcription factor-binding profiles: new content and tools in the 2008 update. *Nucleic Acids Res.* 2008; 36:D102–D106. [PubMed: 18006571]
38. Wasserman WW, Sandelin A. Applied bioinformatics for the identification of regulatory elements. *Nat Rev Genet.* 2004; 5:276–287. [PubMed: 15131651]
39. [data accessed: 6/1/2012] Database of Transcriptional Start Sites, release 8.0. 2011. <http://dbtss.hgc.jp/>
40. [accessed 6/1/2012] ENCODE Digital DNaseI Hypersensitivity Clusters Database. 2011. <http://genome.ucsc.edu/cgi-bin/hgTrackUi?hgsid=327733093&c=chr21&g=wgEncodeRegDnaseClustered>
41. Mock DM, Lankford GL, Mock NI. Biotin accounts for only half of the total avidin-binding substances in human serum. *J Nutr.* 1995; 125:941–946. [PubMed: 7722698]
42. Zemleni J, Helm RM, Mock DM. In vivo biotin supplementation at a pharmacologic dose decreases proliferation rates of human peripheral blood mononuclear cells and cytokine release. *J Nutr.* 2001; 131:1479–1484. [PubMed: 11340103]
43. Manthey KC, Griffin JB, Zemleni J. Biotin supply affects expression of biotin transporters, biotinylation of carboxylases, and metabolism of interleukin-2 in Jurkat cells. *J Nutr.* 2002; 132:887–892. [PubMed: 11983808]
44. Griffin JB, Rodriguez-Melendez R, Zemleni J. The nuclear abundance of transcription factors Sp1 and Sp3 depends on biotin in Jurkat cells. *J Nutr.* 2003; 133:3409–3415. [PubMed: 14608051]
45. Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, et al. High-resolution profiling of histone methylations in the human genome. *Cell.* 2007; 129:823–837. [PubMed: 17512414]
46. Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, et al. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet.* 2007; 39:311–318. [PubMed: 17277777]
47. Liu Y, Han D, Han Y, Yan Z, Xie B, Li J, et al. Ab initio identification of transcription start sites in the Rhesus macaque genome by histone modification and RNA-Seq. *Nucleic Acids Res.* 2011; 39:1408–1418. [PubMed: 20952408]
48. Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature.* 2007; 448:553–560. [PubMed: 17603471]
49. Yang C, Bolotin E, Jiang T, Sladek FM, Martinez E. Prevalence of the initiator over the TATA box in human and yeast genes and identification of DNA motifs enriched in human TATA-less core promoters. *Gene.* 2007; 389:52–65. [PubMed: 17123746]
50. Eng WK, Giraud D, Schlegel VL, Wang D, Lee BH, Zemleni J. Identification and assessment of markers of biotin status in healthy adults. *British Journal of Nutrition.* 2013:1–9.

Fig. 1.
Location of HLCS promoters in the human reference genome (NG_016193). Base pairs (bp) denote the exact location on human chromosome 21.

Fig. 2.
Alignment of the putative HLCS promoters with human mRNA and EST.

Fig. 3.

HLCS promoter methylation in human cell lines. Gray boxes denote methylated DNA and black boxes denote non-methylated DNA. Cell lines: 1, BG02ES embryonic stem cell; 2, BG02ES-EBD: embryonic stem cell; 3, GM12878 B-lymphocyte; 4, GM12891 B-lymphocyte; 5, GM12892 B-lymphocyte; 6, H1-Hesc embryonic stem cells; 7, H9ES-AFP embryonic stem cell; 8, H9ES-AFP+ embryonic stem cell; 9, H9ES-CM embryonic stem cell; 10, H9ES-E embryonic stem cell; 11, H9ES-EB embryonic stem cell; 12, H9ES-EBD embryonic stem cell; 13, HAL adult liver; 14, HCT-116 colorectal carcinoma; 15, HFL11W fetal liver; 16, HFL24W fetal liver; 17, Hela-S3 cervical carcinoma; 18, HepG2 hepatocellular carcinoma; 19, K562 leukemia; 20, H9ES embryonic stem cell.

Fig. 4. Abundance of H3K4me3 and H3K27ac histone modification marks in human GM12878 normal B-lymphocyte cells, HepG2 hepatocarcinoma cells, HMEC normal mammary epithelial cells, HSMN normal skeletal muscle myoblasts cells, and K562 leukemia cells.

Fig. 5. RNA polymerase II binding sites in HLCS promoters P1, P2, and P3, predicted by using ENCODE Pol II ChIP-seq tracks in the UCSC genome browser. HLCS promoters are marked by black boxes at the bottom of each panel.

Fig. 6. Transcription factor binding sites in the human *HLCS* gene, consolidated from predictions in databases UCSC ENCODE Transcription factor Binding Chip-seq tracks, Consite, and JASPAR. *DNase I* hypersensitive regions are underlined; transcription factor binding sites are boxed. TSS, transcription start sites. Truncation sites for generating reporter gene plasmids are marked using arrows and primer identifiers (compare Online Supplementary Table 1 and Fig. 7).

Fig. 7. Luciferase activities HLCS-P1GL3₁₁₂₈, HLCS-P2GL3₂₀₅₄, and HCS-P3GL3₁₂₈₂ in MCF7 cells (panel A), HCT116 cells (panel B), and HEK293 cells (panel C). Panel D depicts the activity of HLCS-P1GL3₁₁₂₈ in HEK293 cells cultured in biotin-supplemented, biotin-sufficient, and biotin-deficient media. Values are mean $\pm$ standard errors (^{a,b,c} Bars not sharing the same letter are significantly different; $p < 0.05$; $n = 3$).

Fig. 8. Luciferase activities in truncated reporter gene constructs in HEK293 cells. Values are mean $\pm$ standard errors (^{a,b,c,d} Bars not sharing the same letter are significantly different; $p < 0.1$; $n = 3$).

Table 1

HLCS promoters regulate the expression of distinct mRNA and EST in various human tissues

Promoter	mRNA/EST	Database ID	Tissue
P1	mRNA	D23672.2	Liver
		BC060787	Placenta, normal tissue
	EST	DA097639	Cerebellum
		DA334377	Hippocampus
		CB988654	Placenta
		DA858766	Placenta
		BX474953	n/a [/]
		DA714809	n/a
		DA715549	n/a
		DA755966	n/a
P2	mRNA	AJ001864	Fetal tissues, brain
	EST	AK307940	Tongue, tumor tissue
		DA102762	Cerebellum
		DA241621	Brain
		DA404459	Thalamus
		DB194634	Trachea
P3	mRNA	DN996277	Breast cancer
	EST	DB290499	Uterus
		DB469252	Hippocampus
		DC307332	Cerebellum
		DC346222	Tongue, tumor tissue
		HY124438	Brain
		AU143539	n/a
		DA755010	n/a
		DC363327	n/a

[/]
n/a, not available.

Table 2

HLCS promoter methylation.

Promoter	GC content	CpG
	% of bases	Score ²
P1	54	0.49
P2	49	0.28
P3	57	0.34

²Calculated as described in [31].