

UC Davis

UC Davis Previously Published Works

Title

Investigation of high gamma-glutamyltransferase syndrome in California Thoroughbred racehorses

Permalink

<https://escholarship.org/uc/item/0r81q012>

Journal

Journal of Veterinary Internal Medicine, 36(6)

ISSN

0891-6640

Authors

Peng, Sichong
Magdesian, K Gary
Dowd, Joseph
et al.

Publication Date

2022-11-01

DOI

10.1111/jvim.16582

Peer reviewed

STANDARD ARTICLE

Investigation of high gamma-glutamyltransferase syndrome in California Thoroughbred racehorses

Sichong Peng¹ | K. Gary Magdesian²  | Joseph Dowd³ | Jeffrey Blea⁴ |
Ryan Carpenter⁵ | Wayne Ho⁶ | Carrie J. Finno¹ 

¹Department of Population Health and Reproduction, University of California, Davis, California, USA

²Department of Medicine and Epidemiology, School of Veterinary Medicine, University of California, Davis, California, USA

³Equine Medical and Surgical Group, Arcadia, California, USA

⁴Von Bleucher, Blea, Hunkin, DVM, Inc, Sierra Madre, California, USA

⁵DVM, Sierra Madre, California, USA

⁶Southern California Equine Foundation, Arcadia, California, USA

Correspondence

Carrie Finno, UC Davis School of Veterinary Medicine, Room 4206 Vet Med 3A, One Shields Ave, Davis, CA 95616, USA.
Email: cjfinno@ucdavis.edu

Funding information

National Center for Advancing Translational Sciences, Grant/Award Number: L40 TR001136; Southern California Equine Foundation, Grant/Award Number: NA; UC Davis Center for Equine Health, Grant/Award Number: NA

Abstract

Background: Increases in serum gamma-glutamyltransferase (GGT) activity have been reported in Thoroughbred (TB) racehorses and associated with maladaptation to training but the underlying etiology remains unknown.

Hypothesis/Objectives: Classify the etiology of high GGT syndrome in racing TBs by assessment of pancreatic enzymes, vitamin E concentrations, and both a candidate gene and whole genome association study. We hypothesized that a genetic variant resulting in antioxidant insufficiency or pancreatic dysfunction would be responsible for high GGT syndrome in TBs.

Animals: A total of 138 California racing TBs. Amylase: n = 31 affected (serum GGT activity ≥ 60 IU/L), n = 52 control (serum GGT activity < 40 IU/L). Lipase: n = 19 affected, n = 35 control. Serum α -tocopherol concentrations: n = 32 affected, n = 46 control. Genome-wide association study (GWAS): 36 affected, 58 control. Whole genome sequencing: n = 5 affected, n = 5 control.

Methods: Biochemical and vitamin analytes were compared among cohorts. A GWAS was performed and a subset of TBs underwent whole genome sequencing to interrogate candidate genes and positional genetic regions.

Results: Serum lipase and amylase activity and α -tocopherol concentrations did not differ between groups. No genetic variants were identified in 2 candidate genes (*UGT1A1* and *GGT1*) that associated with the phenotype. Four single nucleotide polymorphisms (SNPs) approached a suggestive association with the phenotype ($P = 2.15 \times 10^{-5}$), defining a 100 kb region on chromosome 5 surrounding *cluster of differentiation 1a* (*CD1A1*), a transmembrane gene related to the major histocompatibility complex.

Conclusions and Clinical Importance: An underlying genetic etiology may exist for high GGT syndrome in racing TBs, similar to genetic disorders in humans.

Abbreviations: GGT, gamma-glutamyltransferase; GWAS, genome-wide association study; MHC, major histocompatibility complex; TB, Thoroughbred; WGS, whole-genome sequencing; α -TOH, α -tocopherol.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. *Journal of Veterinary Internal Medicine* published by Wiley Periodicals LLC on behalf of American College of Veterinary Internal Medicine.

KEYWORDS

amylase, genetics, lipase, vitamin E

1 | INTRODUCTION

Gamma-glutamyltransferase (GGT) is a membrane-bound enzyme that catalyzes the transfer of gamma-glutamyl groups from gamma-glutamyl peptides. It is important in glutathione metabolism, amino acid absorption, and protection against oxidative injury. The enzyme is found in the biliary tract and is a marker of hepatobiliary disorders and cholestasis in horses. Increases in serum GGT activity have been reported in Thoroughbred (TB) racehorses and associated with cumulative exercise load and potential maladaptation to training.¹⁻⁴ Reasons for these increases in GGT are unknown. Affected horses often show little histologic evidence of liver damage.⁵ It has been postulated that increases in GGT activity in racing TBs are associated with infection with equine hepatitis virus, pegivirus or Theiler's disease-associated virus. However, recent studies indicated that infection with these viruses did not associate with high GGT syndrome in TBs.^{5,6}

In TBs from the east coast of the United States (NY, FL and KY), serum GGT activity above the reference interval (0-50 IU/L)⁵ was detected in 18% of horses tested, with similar prevalence estimates across 4 training farms.⁵ Although significantly higher in control horses, selenium concentrations were within the reference interval for all horses sampled.⁵ Metabolomic profiling identified increased abundance of pyroglutamic acid and taurine-conjugated bile acids, with decreased abundance of vitamin B6.⁵ It therefore was concluded that high GGT syndrome is a complex metabolic disorder. A more recent study using matched case-control pairs of 3 Thoroughbreds and 4 Standardbreds profiled before and after exercise identified that glutathione recycling was enhanced in high GGT horses. Additionally, mild cholestasis was evident in this population.⁶

We aimed to investigate 3 potential interlinked causes of high GGT syndrome in TB racehorses: (a) aberrant pancreatic function, (b) vitamin E deficiency, and (3) an underlying genetic etiology. Increases in GGT activity could result from pancreatic disorders in the horse,^{7,8} similar to findings in humans.⁹ Additionally, because high GGT activity is associated with oxidative stress,^{5,10} antioxidant deficiencies could play a role in the etiology. Although selenium concentrations have been examined previously in high GGT syndrome,⁵ vitamin E concentrations only recently were examined in a small cohort of TBs with high GGT syndrome.⁶

Because of the lack of evidence supporting an infectious or toxic etiology for high GGT syndrome in TB racehorses and comparative inherited diseases in humans, we hypothesized that a genetic variant may be responsible for this condition in TBs. In humans, genetic diseases exist that lead to increases in unconjugated bilirubin concentrations (Gilbert's syndrome) with concurrent increases in GGT activity.¹¹ Liver function remains normal, but increases in unconjugated bilirubin concentrations can occur during periods of stress, including strenuous exercise. Gilbert's syndrome is associated with genetic mutations in

UDP-glucuronosyltransferase 1 (UGT1A1).¹¹⁻¹³ In humans, genetic variants in the *gamma-glutamyltransferase 1* (GGT1) gene itself can result in chronic pancreatitis and increased risk of pancreatic cancer.¹⁴ Therefore, these 2 candidate genes were profiled. Lastly, a distinct syndrome has been described in humans with isolated high GGT activity, no other serum biochemical abnormalities or evidence of viral hepatitis, and no history of drug or alcohol abuse.¹⁵ No genetic variants have been associated with this novel inherited disease in humans. Thus, we undertook a genome-wide association study (GWAS) to perform a more broad genome-wide analysis of high GGT syndrome in TBs.

2 | MATERIALS AND METHODS

2.1 | Animals

All procedures were approved by the University of California Institutional Animal Care and Use Committee (IACUC #20751; approved October 3, 2018). A total of 138 actively racing TB horses at 2 racetracks in Southern California were screened for serum GGT activity and trainers were asked to complete a survey (supplementary file 1). Affected horses were defined as horses with serum GGT activity ≥ 60 IU/L and unaffected horses as horses with serum GGT activity < 40 IU/L. Any horses with serum GGT activity between these 2 values were excluded from the analysis. Samples from 2 unaffected horses were requested from the same trainer and barn as each affected horse to control for environmental factors. Performance histories were requested from the horses with increased GGT activity.

2.2 | Sample Collection

Blood (20 mL) was collected at a single resting time point into serum and EDTA tubes. Samples from affected and control horses were matched by trainer, barn and day. Serum samples were immediately processed for serum biochemistry, including lipase activity, and remaining samples were stored light-protected at -80°C for future α -tocopherol analysis as previously described.¹⁶ An aliquot of serum was processed for measurement of serum amylase activity by a commercial laboratory (IDEXX Laboratories, Westbrook ME). The EDTA samples were centrifuged at $2000 \times g$ for 15 min, and the buffy coat was isolated and shipped on ice to UC Davis for DNA isolation.

2.3 | Data Analysis

The effect of trainer was analyzed using Fisher's exact contingency testing with significance set at $P < .05$. Testing for normal distribution of age and

serum GGT activity was performed using a Shapiro-Wilk test, and the appropriate test employed (parametric unpaired *t* test or non-parametric Mann-Whitney test) to compare analyte concentrations between the 2 phenotypic categories. Simple linear regression was used to determine the effect of high GGT activity on serum amylase concentrations.

2.4 | Power Calculation for Genetic Studies

2.4.1 | Genome-wide association study (GWAS)

To the authors' knowledge, prevalence studies for high GGT syndrome are not currently available. Therefore, we used our screening of 138 actively racing horses, with 48 horses classified as "high GGT syndrome" by our case definition (i.e., serum GGT activity ≥ 60 IU/L) to detect disease prevalence at 34.7%. A detectable relative risk power analysis therefore was estimated for a binary (case-control) phenotype. Power calculations were performed using Power for Genetic Association analyses (PGA) for detection of relative risk alleles = 2 with r^2 set at 0.9 and alpha of 0.05, after accounting for multiple tests.¹⁷ Eighty percent power was achievable with a sample size of $n = 40$ cases and $n = 60$ controls.

2.4.2 | Whole genome association study

A power analysis was performed to determine the minimum number of TBs needed to detect a segregating variant using Fisher's exact test. Because the mode of inheritance is unknown, both recessive and dominant models were evaluated with full penetrance. If the variant resulting in high GGT syndrome was inherited in an autosomal recessive manner and no carriers were inadvertently included as unaffected horses, $n = 2$ horses per group could identify a segregating variant at $P < .05$. However, if inadvertent carriers were included as unaffected horses or if the disease is inherited in an autosomal dominant manner, at least $n = 5$ horses per group would be required to obtain a significant association at $P < .05$.

2.5 | DNA isolation and GWAS

Genomic DNA was isolated from whole blood samples according to the WIZARD Blood DNA Extraction Kit protocol (Promega, Madison, WI). Purified DNA samples were stored at -20°C . A total of 94 samples (36 affected and 58 control) were genotyped using the Affymetrix MNEc670k array.¹⁸ Only SNPs that passed quality control settings (minor allele frequency $> 1\%$, genotyping across individuals $> 90\%$) were selected. A genome-wide efficient mixed model association was performed using GEMMA software using the standardized relatedness matrix option ($-\text{gk } 2$).¹⁹ Population stratification was estimated by assessing the genomic control inflation factor (λ). A Bonferroni correction for 409 613 tests (the number of useable SNPs) from the GEMMA analysis was performed and significance set at $P = 1 \times 10^{-7}$. A second pruned analysis was performed to account

for the extent of linkage disequilibrium in the TB.²⁰ The SNPs were pruned at $r^2 < 0.6$, with 197 137 SNPs remaining after pruning. Remapping of SNP locations was performed to convert EquCab2.0 coordinates to EquCab3.0.²¹ Evaluation of candidate regions was performed using data from the UCSC Table Browser,²² Ensembl²³ and the Functional Annotation of Animal Genomes project.^{24,25}

2.6 | Whole genome sequencing

Genomic DNA from $n = 5$ affected and $n = 5$ unaffected unrelated TBs (Table S1) was sequenced on the Illumina HiSeq3000/4000 with a targeted average depth of $30\times$. Sequencing quality control was performed by FastQC,²⁶ MultiQC,²⁷ Cutadapt,²⁸ and Trim-Galore. Reads were subsequently aligned to EquCab3 using Burrows-Wheeler Aligner's MEM algorithm.²⁹ After alignment, reads were further filtered based on mapping quality ($Q > 30$) using SAMtools.³⁰ Aligned reads were used in subsequent variant calling via freebayes³¹ ($-\text{use-best-n-alleles } 4$ and $-\text{limit-coverage } 50$) and DELLY, using default parameters.³² A Fisher's exact test was applied at each identified variant to test for association of the genotype with GGT phenotype using SnpSift,³³ and functional effects of the variants were predicted using SnpEff.³⁴ Variants were filtered first by variant Phred score quality (threshold of $Q \geq 30$) followed by segregation using Fischer's exact test with SnpSift at $P_{\text{allelic}} < .05$.³³

To identify regions of association for putative regulatory regions, the equine Functional Annotation of Animal Genomes (FAANG) datasets were used.³⁵ Tissue-specific chromatin marks³⁶ and chromatin accessibility regions³⁷ were identified from liver samples as previously described. Hepatic CCCTC-binding factor (CTCF) sites were identified via CTCF ChIP-seq. Briefly, sequence reads were aligned to reference genome EquCab3 using bwa²⁹ and subsequently filtered to remove PCR duplicates and multi-mapping reads using samtools³⁰ and sambamba.³⁸ The CTCF binding regions then were determined using MACS2 peak calling algorithm for paired-end data.³⁹

To interrogate the candidate genes *GGT1* and *UGT1A1*, variants were filtered using SnpSift³³ using EquCab3.0 coordinates for each candidate gene and allowing for 1 kb up- and down-stream from the annotated transcription start and end sites, respectively, using the Ensembl annotation for EquCab3.0 (http://m.ensembl.org/Equus_caballus/Info/Annotation). In addition to assessing significantly associated variants called with freebayes,³¹ structural variants were identified using DELLY³² and bam files were visually inspected using Integrative Genome Viewer⁴⁰ in the positional and biologic candidate gene regions for any structural variants, including duplications, inversions and large deletions or insertions.

3 | RESULTS

3.1 | Overall population

Of the 138 actively racing TB horses screened at 2 racetracks in Southern California, $n = 69$ had GGT activity < 40 IU/L (range, 13-39; mean \pm SD, 28 ± 7.38 IU/L), $n = 21$ were between 40-60 IU/L (47.5

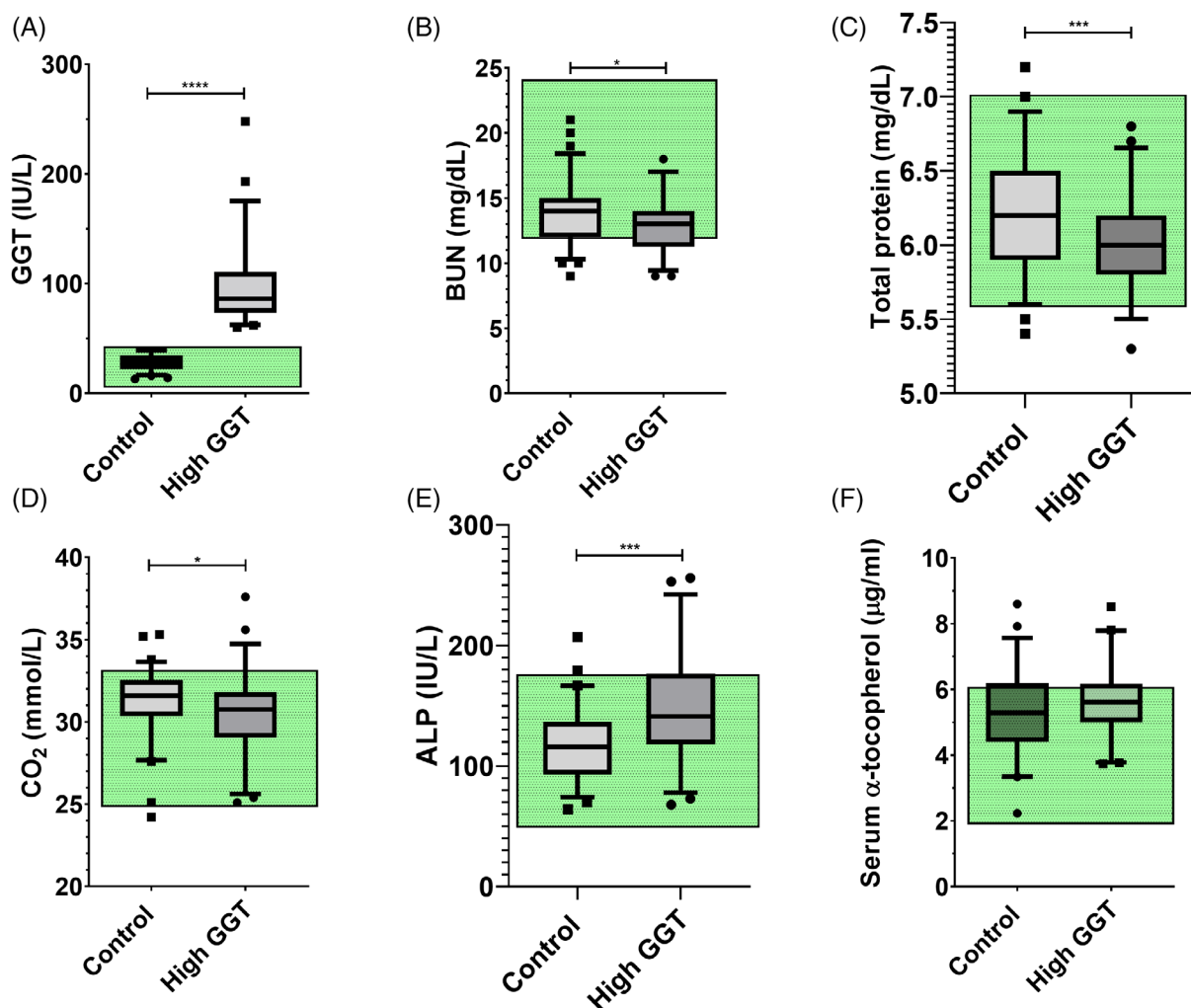


FIGURE 1 A, Serum gamma-glutamyltransferase (GGT), B, blood urea nitrogen (BUN), C, total protein, D, total carbon dioxide (CO_2), E, alkaline phosphatase (ALP) and F, α -tocopherol concentrations in Thoroughbreds with high GGT activity (≥ 60 IU/L) as compared to control Thoroughbreds (GGT < 40 IU/L). Green boxes denote reference interval limits for each analyte. Unpaired t test or Mann-Whitney test, * $P < .05$, ** $P < .01$, *** $P < .001$, **** $P < .0001$

± 5.61 IU/L) and $n = 48$ were ≥ 60 IU/L (range, 60-248; 96.1 ± 35.6 IU/L). High GGT horses were matched to at least 1 control in all cases. In accordance with our study design, serum GGT activities of the HIGH GGT population were significantly higher ($P < .0001$) than our control population in the final GWAS cohort (Figure 1A). When comparing the control population (< 40 IU/L GGT) to the HIGH GGT horses (≥ 60 IU/L), no significant differences of age (control, 1-15 yrs; median, 3 y; HIGH GGT, 1-14 y; median, 2 y; $P = .12$) or sex ($P = .69$) were identified.

Trainer information was not available for $n = 36$ horses. For the remaining $n = 102$ horses, 9 trainers were represented, with 2-35 horses per trainer (median, 6 horses/trainer). Six trainers had ≥ 6 horses in the sample set and were included in the analysis that identified no association ($P = .16$) between trainer and low (< 40 IU/L) vs high (≥ 60 IU/L) serum GGT activity in their cohort of trained horses. Within the $n = 48$ TBs with high GGT activity, $n = 34$ had a performance history reported by the veterinarian, with 17 (50%) reported as

poor performers and 17 (50%) performing as expected. Performance histories were not obtained from the normal GGT cohort.

3.2 | Biochemical analytes

In addition to serum GGT activity, complete serum biochemistry was available in 65/69 control horses and 48/48 HIGH GGT horses. No significant differences were found between HIGH GGT and control horses for the following analytes: glucose, creatinine, albumin, calcium, creatine kinase, phosphorus, total bilirubin, chloride, potassium, sodium, aspartate aminotransferase, direct bilirubin, lactate dehydrogenase and sorbitol dehydrogenase. Although most results remained within the reference intervals, BUN (reference interval, 12-24 mg/dL), total protein (reference interval, 5.2-7.0 g/dL) and total carbon dioxide (reference interval, 25-33 mmol/L) were significantly lower in HIGH GGT horses as compared to controls ($P = .03$, Figure 1B; $P = .007$,

FIGURE 2 A, Serum amylase activity in Thoroughbreds with high GGT activity (≥ 60 IU/L) as compared to control Thoroughbreds (GGT < 40 IU/L). Green boxes denote reference intervals specified by the laboratory. Although there was no significant difference in serum amylase activity between the 2 groups, simple linear regression B, identified that serum GGT activity could reliably predict a portion of the variance in serum amylase activity ($P = .002$). C, However, when this outlier (pink in B) was excluded, the association was no longer significant ($P = .16$)

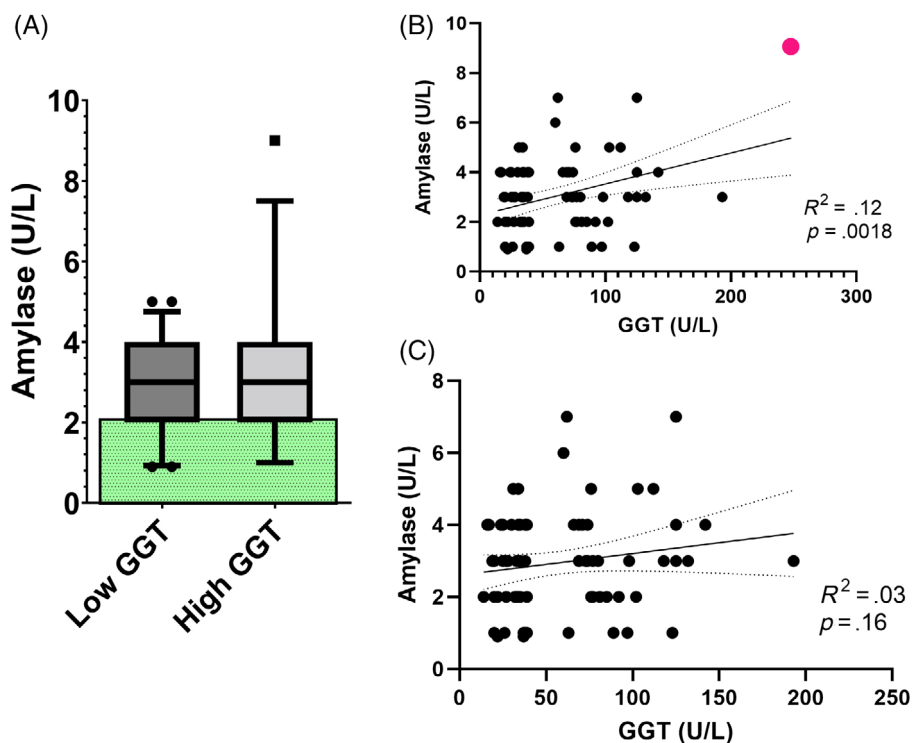


Figure 1C; and $P = .03$, Figure 1D, respectively), whereas alkaline phosphatase (ALP) activity (reference interval, 50-170 IU/L) was significantly higher ($P < .0001$; Figure 1E).

3.3 | Serum α -tocopherol concentrations

From the study population, serum α -tocopherol concentrations were available on 58/69 control and 45/48 HIGH GGT horses. No significant difference was found between groups ($P = .36$) and all results were within or above the reference interval (≥ 2 $\mu\text{g/mL}$; Figure 1F).¹⁶

3.4 | Lipase and amylase activity

Serum lipase activity was assayed in 45/69 control horses and 28/48 HIGH GGT horses and no significant differences were observed between groups ($P = .49$; reference interval, < 87 IU/L⁴¹). Serum amylase activity was assayed in 44/69 control horses and 33/48 HIGH GGT horses. Although on a linear scale no significant difference was found between groups ($P = .11$), many values exceeded the reference interval provided by IDEXX (≤ 2 U/L; Figure 2A). Simple linear regression identified that serum GGT activity could reliably predict a portion of the variance in serum amylase activity ($P = .002$; Figure 2B). The highest serum GGT activity measured was 248 IU/L, with the corresponding highest amylase activity of 9 IU/L. However, when this high value (pink, Figure 2B) was removed from the analysis, the prediction was no longer significant ($P = 0.16$; Figure 2C).

3.5 | Study population - GWAS

A study population subset of 94 TBs (Table S1) was selected based on the highest and lowest serum GGT activity and quality of DNA samples. Of these, 36 horses were selected with a serum GGT activity ≥ 60 IU/L (range, 60-248; 102 ± 45.5 IU/L). These HIGH GGT horses included 13 stallions and 7 geldings (20 males) and 15 mares, ranging in age from 2-15 y (median 2 y). These included all horses with a performance history reported (poor and normal performance). For the control population, 58 trainer- and barn-matched horses were selected with a serum GGT activity < 40 IU/L (range, 13-39; 28.1 ± 7.52 IU/L). This control population included 22 stallions and 12 geldings (34 males), 22 mares and 2 of unknown sex, ranging in age from 2-13 y (median, 3 y). No significant differences in age ($P = .17$) or sex ($P = .83$) were found between HIGH GGT and control populations for the GWAS.

3.6 | Genome-wide association study

After genotyping, 1 control gelding (serum GGT activity = 28 IU/L) had a genotyping rate < 0.95 and was excluded from subsequent analysis, leaving 57 control horses for the full analysis. Before association analysis, a principal component analysis was performed to investigate the sources of genetic variance within the genotyped population. Clustering was not based on GGT activity or sex but did identify 2 distinct clusters of TBs (Figure 3A). Further investigation identified common pedigrees in each cluster. Even with this apparent clustering, only a low level of population stratification existed ($\lambda = 1.001$).

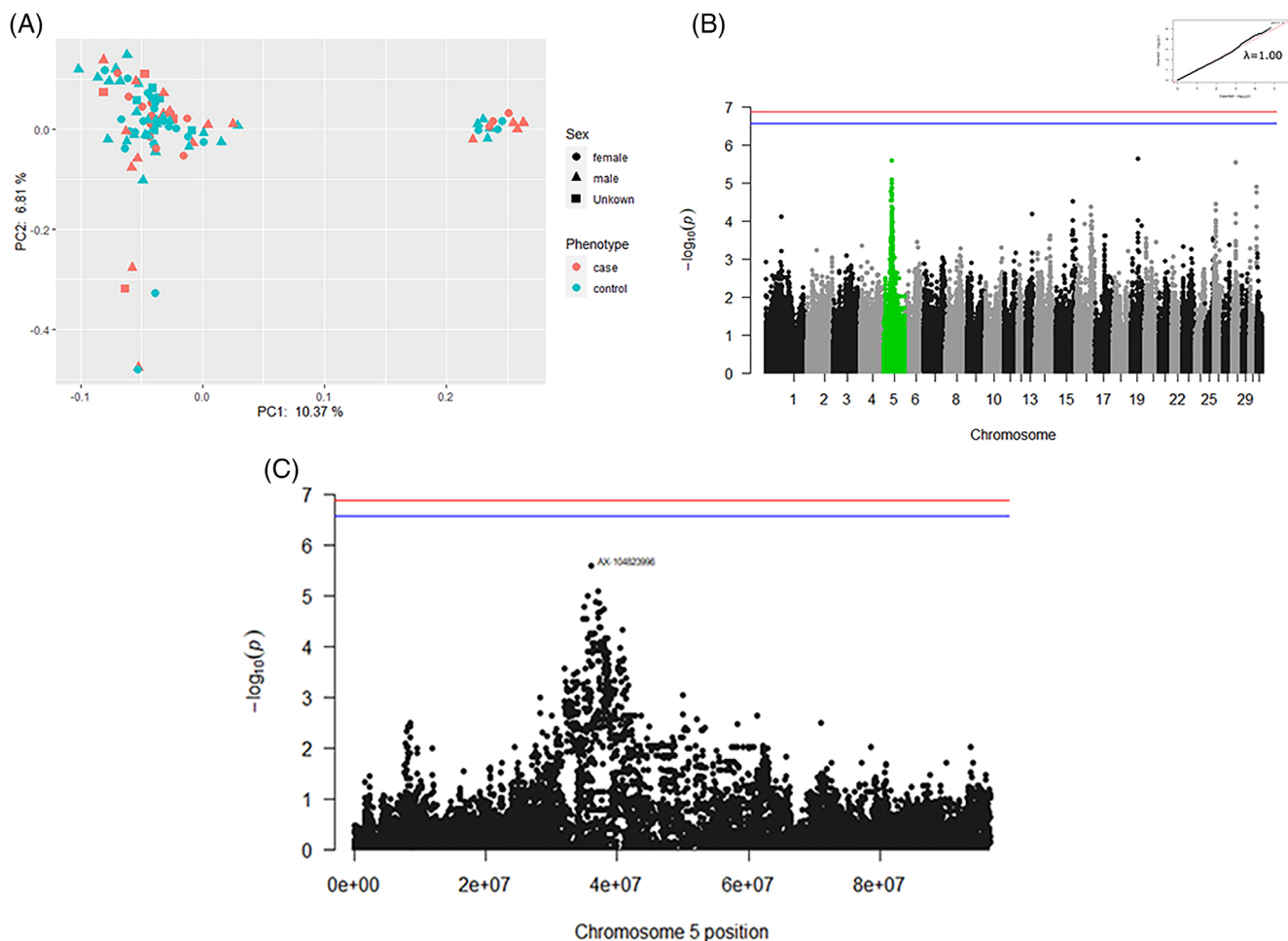


FIGURE 3 A, Principal component analysis identified clustering that was not based on GGT activity (color) or sex (shape) but did identify 2 distinct clusters of TBs. B, Manhattan plot of the mixed linear model association analysis for high GGT syndrome in Thoroughbreds. A genome-wide significance association after Bonferroni correction (red line; $P = 1 \times 10^{-7}$; $-\log_{10}P > 6.9$) was not achieved nor did any SNPs achieve suggestive significance at a false discovery rate of 10% (blue line; $P = 2.3 \times 10^{-7}$; $-\log_{10}P > 6.6$). Inset; quartile-quartile (Q-Q) plot demonstrating minimal population stratification ($\lambda = 1.001$). However, a region of interest ($\log_{10}P > 5.6$) was identified on chromosome 5 (green label in B and region depicted in C)

Using a mixed linear model to account for this small degree of population stratification, an association analysis did not identify any genome-wide significant SNPs after a stringent Bonferroni correction ($-\log_{10}P > 6.9$; Figure 3B, Table S2). However, a region of interest ($\log_{10}P > 5.6$) was identified on chromosome (chr) 5 (Figure 3C). The top 4 SNPs defined an approximately 100 kb region on chromosome 5 from 36 007 315-36 106 165 bp. After pruning of SNPs in linkage disequilibrium, the same region was identified but extended 1.3 Mb (chr5:36-37.4 Mb; Figure S2). Within this extended region of association, 28 Ensembl genes and gene predictions were identified (Table S3). Within the narrower region of association (chr5:36007315-36 106 165 bp), only 1 gene was annotated in Ensembl, *cluster of differentiation 1a* (*CD1A1*), a transmembrane gene related to the major histocompatibility complex (MHC).

Two other single SNPs (EquCab2.0 coordinates chr19:31502688 and chr28:17330962) achieved similar significance to the region on

chromosome 5 (Table S2). When converted to EquCab3.0 coordinates (chr19:34034769 and chr28:18365164, respectively) and evaluated, no annotated genes were found within 50 kb of either SNP.

3.7 | Positional candidate region on chromosome 5

Whole-genome sequencing from $n = 5$ HIGH GGT (68-125 IU/L; 3 stallions, 1 gelding, 1 mare; aged 3-5 y) and $n = 5$ control (17-34 IU/L; 1 stallion, 1 gelding, 3 mares; aged 3-5 y; Table S1) was performed to an average of 30 \times coverage. Within the 100 Kb positional candidate region on chromosome 5, 5 variants were significantly associated with the HIGH GGT phenotype, using a conservative $P_{\text{allelic}} < .01$. These 5 SNPs (Table 1) were Ref/Ref in all cases and either Alt/Ref or Alt/Alt in all controls. One of these SNPs was intronic in *CD1A1*, 3 were upstream of *CD1A1* and 1 was intergenic. When aligned to the available FAANG datasets, 3 of the WGS SNPs were located in a

regulatory region upstream of *CD1A1* that contained both H3K27me3 (repressive) mark and a mark for open chromatin (Figure 4).

3.8 | *GGT1* and *UGT1A1* candidate genes

Within the Ensembl annotation (<https://uswest.ensembl.org/index.html>), *GGT1* is not fully annotated in the horse. Based on comparative data and using RNA sequencing information from the Functional Annotation of Animal Genomes (FAANG) initiative,^{24,25} the corresponding Ensembl gene was identified as ENSECAG0000000097 (chr8: 3066855-3 077 075; transcript ENSECAT0000000268.2). *UGT1A1* is annotated in the horse (ENSECAG00000023519; chr6: 20358172-20 474 496), and coordinates for the longest transcript (ENSECAT00000025670.3) were selected to cover as many putative exons as possible. There were no SNPs approaching significance in either of these candidate genes from our GWAS (Figure 3B). However, we investigated these further using WGS. Within 1 kb up- and downstream of these 2 candidate transcripts, no genetic variants were identified that significantly associated with the HIGH GGT phenotype ($P_{\text{allelic}} < .01$). The RNA expressions of *GGT1* and *UGT1A1* were both high in equine liver (average gene-level transcript per million [TPM] of 27.4 and 207.0, respectively).

4 | DISCUSSION

In our efforts to investigate 3 potential interlinked causes of isolated increased GGT activity in TB racehorses, including aberrant pancreatic function, vitamin E deficiency, and an underlying genetic etiology, we identified a positional candidate gene, *CD1A1*, for this syndrome. Additionally, we determined that pancreatic enzyme activities are normal in HIGH GGT horses and confirmed that vitamin E concentrations are within normal limits in this larger cohort of horses.⁶ Additionally, we have excluded genetic variants in *GGT1* and *UGT1A1* (the gene associated with Gilbert's disease in humans) as potential candidates for high GGT syndrome in horses.

Within the serum biochemical profiles, our findings identified increased ALP activity in the HIGH GGT cohort, similar to a recent study that evaluated $n = 20$ high GGT (defined as ≥ 50 IU/L) and $n = 20$ matched control (defined as ≤ 36 IU/L) TBs.⁵ The authors postulated that this increase was caused by cholestasis, because they also had identified increased total bile acid concentrations in this cohort⁵ and in an additional case-control study.⁶ We did not profile bile acids in our horses, but we detected other differences in analytes that had not been examined previously, including lower BUN, total protein and CO₂ concentrations (albeit within reference intervals; Figure 1). Lower BUN and protein concentrations can occur with chronic liver disease in horses. Thus, our data support that cholestasis or biliary duct disease likely plays a direct or indirect role in high GGT syndrome, which is consistent with recent reports.⁶

In humans, serum GGT activity has been suggested as a marker of oxidative stress¹⁰ after a series of studies showed that both dietary

TABLE 1 Five SNPs within the 100 Kb positional candidate region on chromosome 5 for high GGT syndrome in Thoroughbreds that associated at the conservative $P_{\text{allelic}} < .01$.

Chr	Pos	Ref	Alt	Qual	P-val	Type	Nearest Gene	SNPEFF effect	High GGT horses	Controls					
chr5	36 024 204	C	A	16 741.8	0.005418	intron	CD1A1	Modifier	C/C	C/C	C/C	C/A	C/A	C/C	C/A
chr5	36 050 538	A	G	27 543.4	0.005418	upstream	CD1A1	Modifier	A/A	A/A	A/A	G/A	G/A	G/A	G/A
chr5	36 050 560	C	A	18 411.2	0.005418	upstream	CD1A1	Modifier	C/C	C/C	C/C	A/A	A/C	A/C	A/C
chr5	36 050 565	C	T	18 713.4	0.005418	upstream	CD1A1	Modifier	C/C	C/C	C/C	T/T	T/C	T/C	T/C
chr5	36 084 543	G	A	3197.82	0.005418	intergenic	CD1A1	Modifier	G/G	G/G	G/G	A/A	A/A	A/A	G/G

Abbreviations: Alt, alternate allele; Chr, chromosome; Pos, position; Qual, quality; Ref, reference allele; SNPEFF, effect predictions from SNPEff software.

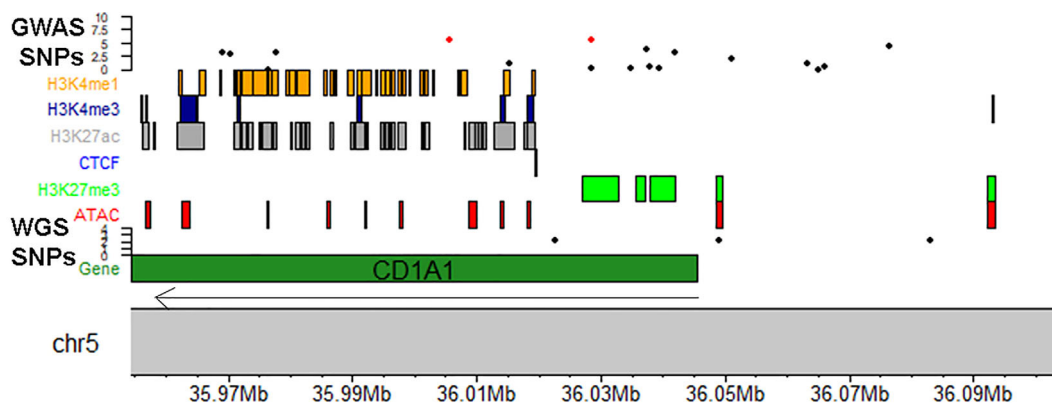


FIGURE 4 Within the narrow region of suggested association on chromosome 5 (36007315-36 106 165 bp; plot extends from chr5:35956000-36 106 165 to capture entire *CD1A1* gene), 5 variants identified by whole genome sequencing (WGS) of $n = 5$ high GGT and $n = 5$ control horses were significantly associated with the HIGH GGT phenotype, using a conservative $P_{\text{allelic}} < .01$. One of these SNPs was intronic in *CD1A1*, 3 were upstream of *CD1A1* (all located in the same region at 36.05 Mb) and 1 was intergenic (see WGS SNPs pane). GWAS SNPs, with their associated P values, are delineated in black in the top panel (GWAS SNPs), with most associated SNPs in red. When aligned to the available FAANG datasets (H3K4me1 orange, H3K4me3 purple, H3K27ac gray, CTCF blue, H3K27me3 green and ATAC red) for liver, 3 of the WGS SNPs (chr5: 36050538, chr5: 36050560 and chr5: 36050565) were located in an H3K27me3 mark and region of open chromatin

and serum antioxidants predicted future serum GGT activity.^{42,43} A subsequent study of >9000 adult participants found that serum carotenoids (α -carotene, β -carotene, β -cryptoxanthin, zeaxanthin/lutein, lycopene) and vitamin C were strongly and inversely correlated with serum GGT activity within its normal range, but serum vitamin E concentrations were not correlated.⁴⁴ In horses, a previous study identified a significant decrease in whole blood selenium concentrations in high GGT TBs, albeit all values remained within the reference interval.⁵ Serum vitamin E (α -tocopherol), another important component of the antioxidant system, was assessed in a small case-control cohort of racing TBs with high GGT activity and no deficiencies were identified.⁶ Our findings were similar in our larger cohort of profiled horses. All horses had α -tocopherol concentrations within the reference intervals with no significant difference between HIGH GGT horses and controls. Thus, although selenium concentrations may differ slightly with high GGT syndrome, marked antioxidant deficiencies do not appear to occur when using selenium and α -tocopherol as markers. This conclusion is further supported by recent work that did not identify any differences in the concentrations of advanced oxidation protein products or the redox balance, as assessed by calculating the oxidant status index, in a small cohort of high GGT TBs and Standardbreds.⁶

A genetic basis for high GGT syndrome has been suggested based on the comparison to Gilbert's syndrome in humans.⁴⁵ Gilbert's syndrome is characterized by mild chronic unconjugated hyperbilirubinemia in the absence of liver disease or overt hemolysis. Although the previous study⁵ and our study did not identify any differences in total bilirubin or direct bilirubin concentrations, serum bilirubin concentrations fluctuate in people with Gilbert's syndrome and often fall within normal limits.⁴⁶ Hepatic glucuronidation is essential for efficient biliary excretion of bilirubin and is approximately 30% of normal in patients with Gilbert's syndrome.^{47,48} The enzyme UGT1A1 is required for conjugation and subsequent excretion of bilirubin. Thus, this gene and

GGT1 were considered candidate genes and investigated in a subset of horses in our study. No genetic variants were found within either of these genes that significantly associated with the high GGT phenotype. Therefore, high GGT syndrome does not appear to have the same underlying molecular basis as Gilbert's syndrome in humans.

To pursue a more global genetic investigation, we performed a GWAS on a cohort of high GGT and matched control TBs. Although no SNPs achieved genome-wide significance, a region of interest was identified on chromosome 5. Of the 5 WGS variants that segregated with high GGT in this region on chromosome 5, the pairwise correlations were almost perfect at $r = 0.83-1$. Within this region, there is only 1 annotated gene, *CD1A1*. *CD1A1* encodes a protein primarily expressed in epidermal dendritic Langerhans cells. In humans, diseases associated with genetic variants in *CD1A1* include histiocytosis, characterized by an abnormal increase in monocytes, macrophages and dendritic cells. Clinically, this disorder manifests most often in children aged 5-10 years of age as dermal rashes and systemic clinical signs such as jaundice, vomiting, hormonal problems, headache and spontaneous fractures.⁴⁹ When compared to skin in the FAANG mRNA-seq data set, *CD1A1* was very lowly expressed in liver (gene-level TPM: 0.32). However, because GGT is an enzyme on the cellular membrane, active transcription may not be required for its function. Additionally, our FAANG dataset identified that 3 of our WGS SNPs overlapped with an H3K27me3 mark and region of open chromatin. Although H3K27me3 is a repressive mark, active chromatin regions are refractory to polycomb repressive complex 2 activity, thereby preventing H3K27me3-mediated repression of transcriptionally active genes.⁵⁰ Thus, H3K27me3 is classified as a promiscuous repressive chromatin mark. These regulatory variants therefore may warrant further investigation in high GGT syndrome.

Despite not obtaining a definitive genetic cause for high GGT syndrome, an underlying genetic basis still may exist. Isolated increases in serum GGT activity are not found in most other breeds of

horses, with the exception of Standardbreds.^{2,51} If the underlying etiology for increases in serum GGT activity were infectious or toxic, it would be expected that other breeds would be similarly affected, including racing Quarter Horses. However, overall amount and duration of exercise could be environmental factors that play a role in these isolated increases in serum GGT activity.²⁻⁴

Based on the metabolite profiling performed previously,⁵ altered lipid metabolism was suspected, with an increase in palmitic acid abundance and a decrease in propionyl-carnitine abundance in high GGT horses as compared to controls. High GGT horses also had lower vitamin B6 concentrations, which could contribute to carnitine deficiency that may be manifested with exercise.⁵ Hepatic disease, because of inborn errors of metabolism, is relatively common in humans and requires metabolite testing or genomic sequencing to arrive at the appropriate diagnosis.⁵² Thus, continued genetic investigation of high GGT syndrome is warranted.

Limitations of our study include that not all horses were able to be sampled for all analyses, thus altering the overall number for each analyte. Unfortunately, this limitation often occurs with prospective studies using client-owned animals. Additionally, although our power analysis suggested that $n = 40$ affected horses and $n = 60$ unaffected horses were required to achieve 80% power in our GWAS, the quality of DNA for some horses limited our genotyped cases to only $n = 36$ affected and $n = 58$ unaffected horses. It was evident from the Q-Q plot (Figure 3B) that our study may have been underpowered, which is why we chose to pursue the region on chromosome 5 although it did not achieve genome-wide significance.

Another limitation is that certain medications can lead to high GGT activity, and extensive screening for medications was not performed in our study. In order to minimize any effects of a potential drug, we endeavored to sample across training barns and match our cases to at least 2 controls from the same barn. Additionally, these horses were in active race training at racetracks in southern California when these samples were taken. As such, they would have undergone routine drug testing at the end of every race. We therefore investigated the medication violation tracking log through the California Horse Racing Board, and none of these horses were on the violation list from 2018-2020. Lastly, although our whole genome power analysis suggested that we would have the power to detect putative associations with whole-genome variants at an uncorrected P -value of .05, any potential genetic heterogeneity or mis-phenotyping could lead to false negative findings for our candidate genes.

In conclusion, we have excluded an association of genetic variants in *GGT1* and *UGT1A1* with high GGT syndrome in TBs and instead identified a putative positional candidate gene, *CD1A1*, that warrants further investigation with this phenotype.

ACKNOWLEDGMENT

This project was supported by the Southern California Equine Foundation, Inc and the Center for Equine Health with funds provided by the State of California pari-mutuel fund and

contributions by private donors. Support for this work was provided by the National Institutes of Health (NIH) to C.J.F. (L40 TR001136). The authors acknowledge the trainers and staff at Santa Anita and Del Mar for their assistance in sample collection. We also acknowledge Karen Klawitter for her assistance in coordinating this project.

CONFLICT OF INTEREST

Authors declare no conflict of interest.

OFF-LABEL ANTIMICROBIAL

Authors declare no off-label use of antimicrobials.

INSTITUTIONAL ANIMAL CARE AND USE COMMITTEE (IACUC) OR OTHER APPROVAL DECLARATION

Approved by the University of California IACUC #20751, October 3, 2018.

HUMAN ETHICS APPROVAL DECLARATION

Authors declare human ethics approval was not needed for this study.

ORCID

K. Gary Magdesian  <https://orcid.org/0000-0001-9988-3381>

Carrie J. Finno  <https://orcid.org/0000-0001-5924-0234>

REFERENCES

1. Tyler-McGowan CM, Golland LC, Evans DL, Hodgson DR, Rose RJ. Haematological and biochemical responses to training and overtraining. *Equine Vet J Suppl.* 1999;31:621-625.
2. Mack SJ, Kirkby K, Malalana F, McGowan CM. Elevations in serum muscle enzyme activities in racehorses due to unaccustomed exercise and training. *Vet Rec.* 2014;174:145.
3. Rivero JL, van Breda E, Rogers CW, Lindner A, Sloet van Oldruitenborgh-Oosterbaan MM. Unexplained underperformance syndrome in sport horses: classification, potential causes and recognition. *Equine Vet J.* 2008;40:611-618.
4. Snow DH, Gash SP, Rice D. Field observations on selenium status, whole blood glutathione peroxidase and plasma gamma-glutamyl transferase activities in thoroughbred racehorses. In: Gillespie JR, Robinson NE, eds. *Equine Exercise Physiology 2*. Davis, CA: ICEEP Publications; 1987:494.
5. Mann S, Ramsay JD, Wakshlag JJ, Stokol T, Reed S, Divers TJ. Investigating the pathogenesis of high-serum gamma-glutamyl transferase activity in thoroughbred racehorses: a series of case-control studies. *Equine Vet J.* 2021.
6. Mann S, Abuelo A, Stokol T, et al. A case-control exercise challenge study on the pathogenesis of high serum gamma-glutamyl transferase activity in racehorses. *Equine Vet J.* 2022;54:39-51.
7. Brobst DF. Pancreatic function. In: Kaneko JJ, Harvey JW, Bruss ML, eds. *Clinical Biochemistry of Domestic Animals*. 5th ed. San Diego, CA: Academic Press; 1997:353-366.
8. Braun JP, Benard P, Burgat V, Rico AG. Gamma Glutamyl transferase in domestic animals. *Vet Res Commun.* 1983;6:77-90.
9. Ruttensburg AM, Goldbarg JA, Pineda EP. Serum gamma-Glutamyl Transpeptidase activity in hepatobiliary pancreatic disease. *Gastroenterol.* 1963;45:43-48.
10. Lee DH, Blomhoff R, Jacobs DR Jr. Is serum gamma glutamyl-transferase a marker of oxidative stress? *Free Radic Res.* 2004;38: 535-539.

11. Bosma PJ, Chowdhury JR, Bakker C, et al. The genetic basis of the reduced expression of bilirubin UDP-glucuronosyltransferase 1 in Gilbert's syndrome. *N Engl J Med*. 1995;333:1171-1175.
12. Aono S, Adachi Y, Uyama E, et al. Analysis of genes for bilirubin UDP-glucuronosyltransferase in Gilbert's syndrome. *Lancet*. 1995;345:958-959.
13. Koiwai O, Nishizawa M, Hasada K, et al. Gilbert's syndrome is caused by a heterozygous missense mutation in the gene for bilirubin UDP-glucuronosyltransferase. *Hum Mol Genet*. 1995;4:1183-1186.
14. Brand H, Diergaard B, O'Connell MR, Whitcomb DC, Brand RE. Variation in the gamma-glutamyltransferase 1 gene and risk of chronic pancreatitis. *Pancreas*. 2013;42:836-840.
15. Bibas M, Zampa G, Procopio A, Guaitolini R. High serum gamma-glutamyltransferase concentrations in a family. *N Engl J Med*. 1994;330:1832-1833.
16. Finno CJ, Valberg SJ. A comparative review of vitamin E and associated equine disorders. *J Vet Intern Med*. 2012;26:1251-1266.
17. Menashe I, Rosenberg PS, Chen BE. PGA: power calculator for case-control genetic association analyses. *BMC Genet*. 2008;9:36.
18. Schaefer RJ, Schubert M, Bailey E, et al. Developing a 670k genotyping array to tag ~2M SNPs across 24 horse breeds. *BMC Genomics*. 2017;18:565.
19. Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. *Nat Genet*. 2012;44:821-824.
20. Wade CM, Giulotto E, Sigurdsson S, et al. Genome sequence, comparative analysis, and population genetics of the domestic horse. *Science*. 2009;326:865-867.
21. Beeson SK, Schaefer RJ, Mason VC, McCue ME. Robust remapping of equine SNP array coordinates to EquCab3. *Anim Genet*. 2019;50:114-115.
22. Karolchik D, Hinrichs AS, Furey TS, et al. The UCSC table browser data retrieval tool. *Nucleic Acids Res*. 2004;32:D493-D496.
23. Howe KL, Achuthan P, Allen J, et al. Ensembl 2021. *Nucleic Acids Res*. 2021;49:D884-D891.
24. Burns EN, Bordbari MH, Mienaltowski MJ, et al. Generation of an equine biobank to be used for functional annotation of animal genomes project. *Anim Genet*. 2018;49:564-570.
25. Donnelly CG, Bellone RR, Hales EN, et al. Generation of a biobank from two adult thoroughbred stallions for the functional annotation of animal genomes initiative. *Front Genet*. 2021;12:650305.
26. Andrews S. A Quality Control Tool for High Throughput Sequence Data. *FastQC*; 2010.
27. Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*. 2016;32:3047-3048.
28. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal*. 2011;17:10-12.
29. Li H, Durbin R. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics*. 2009;25:1754-1760.
30. Li H, Handsaker B, Wysoker A, et al. The sequence alignment/map format and SAMtools. *Bioinformatics*. 2009;25:2078-2079.
31. Garrison E, Marth G. Haplotype-based variant detection from short-read sequencing. *arXiv Preprint*. 2012;12073907.
32. Rausch T, Zichner T, Schlattl A, Stutz AM, Benes V, Korbel JO. DELLY: structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics*. 2012;28:i333-i339.
33. Cingolani P, Patel VM, Coon M, et al. Using *Drosophila melanogaster* as a model for genotoxic chemical mutational studies with a new program. *SnpSift Front Genet*. 2012;3:35.
34. Cingolani P, Platts A, Wang le L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly*. 2012;6:80-92.
35. Peng S, Petersen JL, Bellone RR, et al. Decoding the equine genome: lessons from ENCODE. *Genes (Basel)*. 2021;12(11):1707.
36. Kingsley NB, Kern C, Creppe C, et al. Functionally annotating regulatory elements in the equine genome using histone mark ChIP-Seq. *Genes (Basel)*. 2019;11(1):3.
37. Peng S, Bellone R, Petersen JL, Kalbfleisch TS, Finno CJ. Successful ATAC-Seq From Snap-Frozen Equine Tissues. *Front Genet*. 2021;12:641788.
38. Tarasov A, Vilella AJ, Cuppen E, Nijman IJ, Prins P. Sambamba: fast processing of NGS alignment formats. *Bioinformatics*. 2015;31:2032-2034.
39. Zhang Y, Liu T, Meyer CA, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol*. 2008;9:R137.
40. Robinson JT, Thorvaldsdottir H, Winckler W, et al. Integrative genomics viewer. *Nat Biotechnol*. 2011;29:24-26.
41. Lavoie J, Hinchcliff KW. *Blackwell's Five-Minute Veterinary Consult: Equine*. 2nd ed. Wiley-Blackwell: Ames, IA; 2008.
42. Lee DH, Gross MD, Jacobs DR Jr. Cardiovascular Risk Development in Young Adults Study. Association of serum carotenoids and tocopherols with gamma-glutamyltransferase: the cardiovascular risk development in young adults (CARDIA) study. *Clin Chem*. 2004;50:582-588.
43. Lee DH, Steffen LM, Jacobs DR Jr. Association between serum gamma-glutamyltransferase and dietary factors: the coronary artery risk development in young adults (CARDIA) study. *Am J Clin Nutr*. 2004;79:600-605.
44. Lim JS, Yang JH, Chun BY, Kam S, Jacobs DR Jr, Lee DH. Is serum gamma-glutamyltransferase inversely associated with serum antioxidants as a marker of oxidative stress? *Free Radic Biol Med*. 2004;37:1018-1023.
45. Smith GW. Congenital Hyperbilirubinemia. In: Smith BP, van Metre DC, Pusterla N, eds. *Large Animal Internal Medicine*. 6th ed. St. Louis, MO: Elsevier; 2020:948.
46. Owens D, Evans J. Population studies on Gilbert's syndrome. *J Med Genet*. 1975;12:152-156.
47. Black M, Billing BH. Hepatic bilirubin udp-glucuronyl transferase activity in liver disease and gilbert's syndrome. *N Engl J Med*. 1969;280:1266-1271.
48. Arias IM, London IM. Bilirubin glucuronide formation in vitro; demonstration of a defect in Gilbert's disease. *Science*. 1957;126:563-564.
49. Fraitag S, Emile JF. Cutaneous histiocytoses in children. *Histopathology*. 2022;80:196-215.
50. Wiles ET, Selker EU. H3K27 methylation: a promiscuous repressive chromatin mark. *Curr Opin Genet Dev*. 2017;43:31-37.
51. Leleu C, Haentjens F. Morphological, haemato-biochemical and endocrine changes in young Standardbreds with 'maladaptation' to early training. *Equine Vet J Suppl*. 2010;42:171-178.
52. Chanprasert S, Scaglia F. Adult liver disorders caused by inborn errors of metabolism: review and update. *Mol Genet Metab*. 2015;114:1-10.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Peng S, Magdesian KG, Dowd J, et al. Investigation of high gamma-glutamyltransferase syndrome in California Thoroughbred racehorses. *J Vet Intern Med*. 2022; 36(6):2203-2212. doi:10.1111/jvim.16582