Multifaceted roles of SAMHD1 in cancer

Katie-May McLaughlin[1], Jindrich Cinatl jr.[2]*, Mark N. Wass[1]*, Martin Michaelis[1]*

[1] School of Biosciences, University of Kent, Canterbury, UK

[2] Institute of Medical Virology, Goethe-University, Frankfurt am Main, Germany

* Correspondence to: Jindrich Cinatl jr. (Cinatl@em.uni-frankfurt.de), Mark N. Wass (M.N.Wass@kent.ac.uk), Martin Michaelis (M.Michaelis@kent.ac.uk)

Abstract

SAMHD1 is discussed as a tumour suppressor protein, but its potential role in cancer has only been investigated in very few cancer types. Here, we performed a systematic analysis of the TCGA (adult cancer) and TARGET (paediatric cancer) databases, the results of which did not suggest that SAMHD1 should be regarded as a *bona fide* tumour suppressor. SAMHD1 mutations that interfere with SAMHD1 function were not associated with poor outcome, which would be expected for a tumour suppressor. High SAMHD1 tumour levels were associated with increased survival in some cancer entities and reduced survival in others. Moreover, the data suggested differences in the role of SAMHD1 between males and females and between different races. Often, there was no significant relationship between SAMHD1 levels and cancer outcome. Taken together, our results indicate that SAMHD1 may exert pro- or anti-tumourigenic effects and that SAMHD1 is involved in the oncogenic process in a minority of cancer cases. These findings seem to be in disaccord with a perception and narrative forming in the field suggesting that SAMHD1 is a tumour suppressor. A systematic literature review confirmed that most of the available scientific articles focus on a potential role of SAMHD1 as a tumour suppressor. The reasons for this remain unclear but may include confirmation bias and publication bias. Our findings emphasise that hypotheses, perceptions, and assumptions need to be continuously challenged by using all available data and evidence.

Introduction

Sterile α motif and histidine-aspartic domain containing protein 1 (SAMHD1) was initially discovered in dendritic cells and named dendritic cell-derived IFN-γ induced protein (DCIP) [Li et al., 2000; Coggins et al., 2020]. SAMHD1 is indeed involved in the regulation of interferon signalling, and SAMHD1 mutations are associated with Aicardi-Goutieres syndrome, an autoimmune inflammatory disorder characterised by a dysregulated interferon response [Rice et al., 2009; Mauney & Hollis, 2018; Coggins et al., 2020].

In the meantime, SAMHD1 has been shown to exert a range of additional functions [Mauney & Hollis, 2018; Coggins et al., 2020; Chen et al., 2021]. As a deoxynucleoside triphosphate hydrolase (dNTPase), that cleaves deoxynucleoside triphosphates (dNTPs) into deoxynucleosides and triphosphate, SAMHD1 plays, together with enzymes that catalyse dNTP biosynthesis, an important role in the maintenance of balanced cellular dNTP pools [Mauney & Hollis, 2018; Coggins et al., 2020; Chen et al., 2021]. Since imbalances in cellular dNTP pools affect cell cycle regulation and DNA stability, SAMHD1 is also involved in the regulation of these processes [Chen et al., 2021].

In addition to controlling cellular dNTP levels, SAMHD1 has been shown to maintain genome integrity by a range of further mechanisms, including maintenance of telomere integrity, inhibition of LINE-1 retrotransposons, facilitation of homologous recombination-mediated double-strand break repair and DNA end joining, and prevention of R-loop formation at transcription-replication conflict regions [Herold et al., 2017a; Akimova et al., 2021; Chen et al., 2021; Park et al., 2021]. Additionally, low SAMHD1 levels have been detected in chronic lymphocytic leukaemia (CLL), lung cancer, cutaneous T-cell lymphoma, AML, colorectal cancer, and Hodgkin lymphoma. Moreover, loss-of-function SAMHD1 mutations have been described in cancer types including CLL and colorectal cancer [Herold et al., 2017a; Mauney & Hollis, 2018; Coggins et al., 2020; Chen et al., 2021]. Due to these observations, SAMHD1 is being considered as a tumour suppressor protein.

However, the potential role of SAMHD1 in cancer diseases is more complex. It also recognises and cleaves the triphosphorylated, active forms of a range of anti-cancer nucleoside analogues. In this context, SAMHD1 has been described as a clinically relevant resistance factor in acute myeloid leukaemia (AML) and acute lymphoblastic leukaemia (ALL) against nucleoside analogues including cytarabine, decitabine, and nelarabine [Herold et al., 2017b; Schneider et al., 2017; Knecht et al., 2018; Oellerich et al., 2019; Rothenburger et al., 2020].

So far, the potential tumour suppressor activity of SAMHD1 has only been investigated in a few cancer types. To establish a broader understanding of the role of SAMHD1 in cancer, we here performed a systematic analysis of mutation data, gene expression data, and cancer patient survival data provided by The Cancer Genome Atlas (TCGA) [Cancer Genome Atlas Research Network, 2008] and the Therapeutically Applicable Research To Generate Effective Treatments (TARGET) (https://ocg.cancer.gov/programs/target) databases. The TCGA provided data from 9,703 patients with 33 different types of adult cancer and the TARGET database from 1,091 patients with seven different paediatric cancer types.

Results

High *SAMHD1* expression is not consistently associated with increased survival

Although SAMHD1 has recently been considered as a tumour suppressor protein [Herold et al., 2017a; Mauney & Hollis, 2018; Coggins et al., 2020], high *SAMHD1* expression in tumour tissues was not associated with favourable outcomes across all patients in the TCGA (Figure 1A). In contrast, high *SAMHD1* expression was associated with favourable outcome across all patients in the paediatric cancer database TARGET (Figure 1A).

Considering the individual cancer categories in the TCGA database, high *SAMHD1* expression was significantly associated with increased survival in nine out 33 of cancer categories (Figure 1B, Supplementary Table 1) and with poor outcome in five cancer categories (Figure 1B, Supplementary Table 1). This indicates that the role of SAMHD1 differs between cancer types and that it does not always function as a tumour suppressor. Similarly, high SAMHD1 expression was significantly correlated with longer survival in only one cancer type (osteosarcoma) in the TARGET database but with reduced survival in two others (acute lymphoblastic leukaemia, Wilm's tumour) (Figure 1B, Supplementary Table 2).

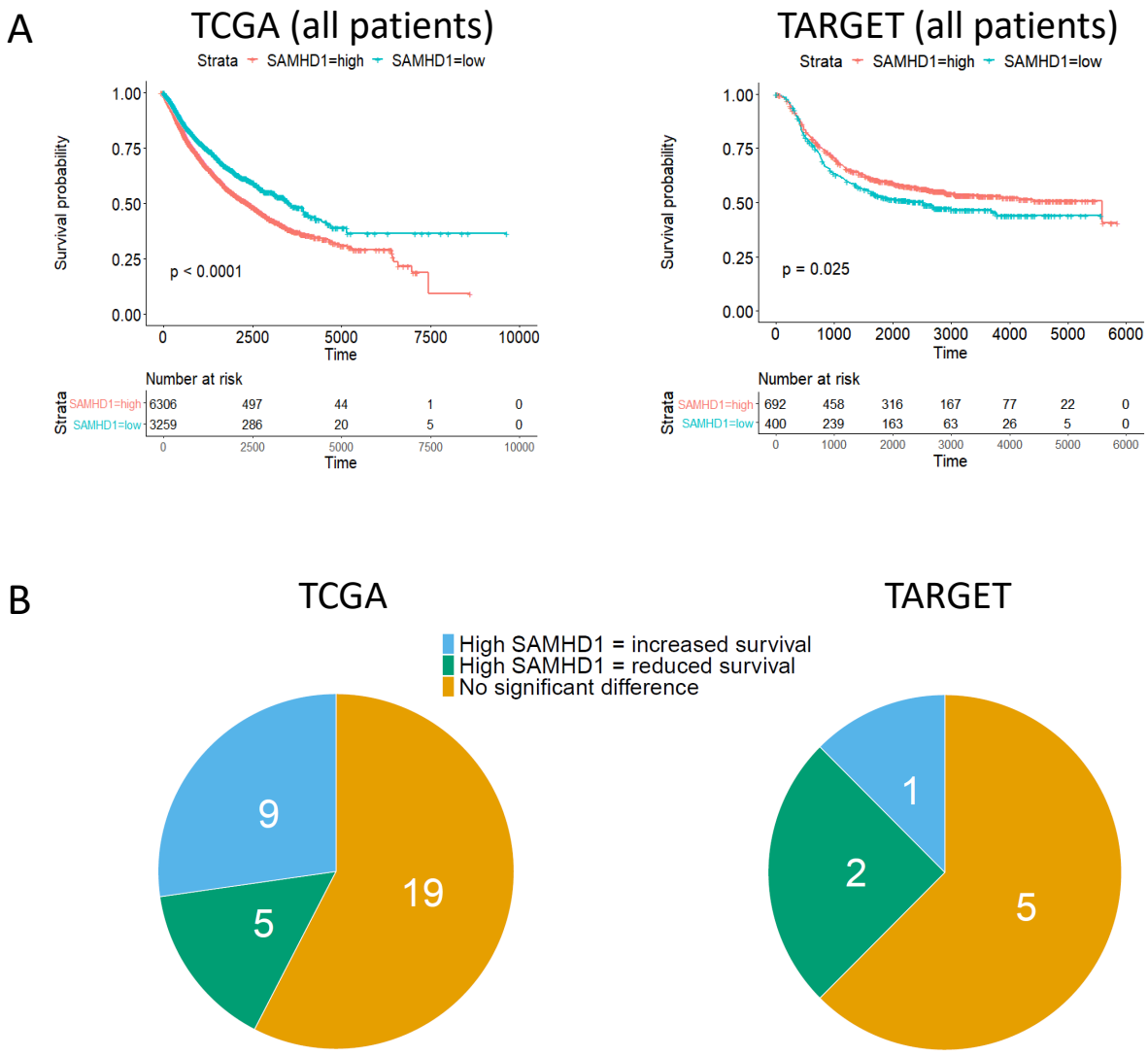


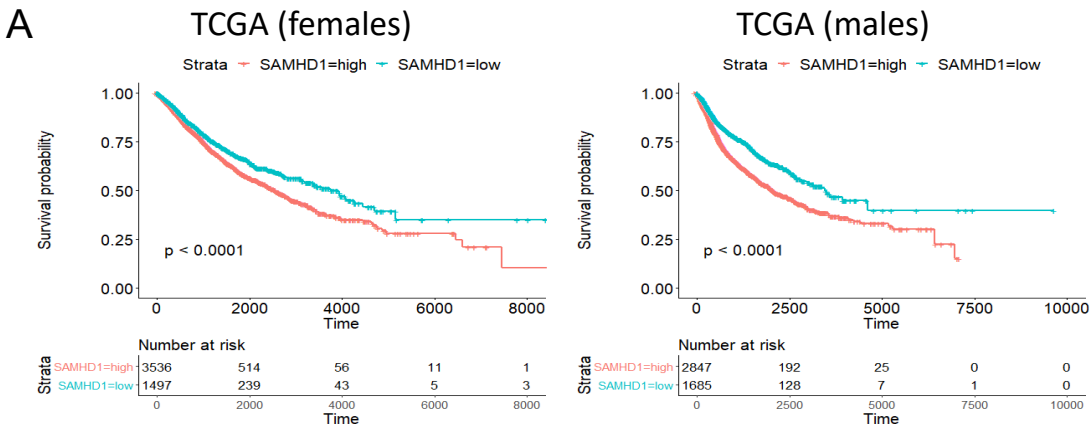
Figure 1. Effect of *SAMHD1* expression in cancer patients. A) Kaplan Meier plots indicating survival in cancer patients with tumours characterised by high or low *SAMHD1* expression (as determined by best separation) across all patients in the TCGA and TARGET databases. P-values were determined by log-rank test. B) Pie charts indicating the number of cancer types for which high SAMHD1 expression was associated with increased survival, reduced survival, or not significantly associated with survival based on data from the TCGA and TARGET databases. Data are presented in Supplementary Table 1 and Supplementary Table 2.

Role of *SAMHD1* expression in the context of sex

Next, we analysed the role of SAMHD1 in males and females in TCGA and TARGET. In TCGA, there were no sex-specific differences with regard to the association of SAMHD1 with survival time across all cancer types (Figure 2A). However, some discrepancies became visible upon the comparison of the role of SAMHD1 in the 27 cancer entities that occur in both females and males (Figure 2B, Supplementary Table 3). When we did not consider statistical significance levels, high SAMHD1 levels were associated with higher 5-year survival rates in 13 cancer entities across all patients, in 17 cancer entities in female patients, and in 14 cancer entities in male patients (Figure 2B, Supplementary Table 3).

When we only considered comparisons in which the 5-year survival rates were significantly different ($p<0.05$) between high and low SAMHD1-expressing tumours for at least one comparison (across all patients, in females, and/ or males), differences reached significance for only one sex in ten cancer types (Figure 2C, Supplementary Table 3). Consistent findings were obtained for three cancer types (LAML, LGG, SARC, all abbreviations for cancer entities are provided in Supplementary Table 1 and the legend of Figure 2).

Although differences did not reach our cut-off value for statistical significance ($p < 0.05$), trends were detected indicating opposite effects of SAMHD1 on disease outcome between the sexes in four cancer entities, (Figure 2D, Supplementary Table 3). In kidney renal clear cell carcinoma (KIRC), low SAMHD1 levels were associated with a higher 5-year survival rate (68%) than high SAMHD1 levels (50%) in males ($p=0.009$). In contrast, high SAMHD1 levels were related to higher survival in females (72% vs. 55%), with the p-value being close to significance ($p=0.056$). In three other cancer types (KIRP, PAAD, STAD), low SAMHD1 levels were associated with higher 5-year survival in females and with lower 5-year survival in males (Figure 2D, Supplementary Table 3).



B

High SAMHD1
Low SAMHD1

	All patients	Females	Males
ACC	High SAMHD1	High SAMHD1	High SAMHD1
BLCA	Low SAMHD1	Low SAMHD1	Low SAMHD1
BRCA	High SAMHD1	High SAMHD1	High SAMHD1
CHOL	High SAMHD1	High SAMHD1	High SAMHD1
COAD	High SAMHD1	Low SAMHD1	High SAMHD1
DLBC	High SAMHD1	High SAMHD1	High SAMHD1
ESCA	Low SAMHD1	High SAMHD1	Low SAMHD1
GBM	High SAMHD1	High SAMHD1	Low SAMHD1
HNSC	High SAMHD1	High SAMHD1	High SAMHD1
KICH	Low SAMHD1	High SAMHD1	Low SAMHD1
KIRC	Low SAMHD1	High SAMHD1	Low SAMHD1
KIRP	Low SAMHD1	Low SAMHD1	High SAMHD1
LAML	Low SAMHD1	Low SAMHD1	Low SAMHD1
LGG	Low SAMHD1	Low SAMHD1	Low SAMHD1
LIHC	Low SAMHD1	High SAMHD1	Low SAMHD1
LUAD	High SAMHD1	High SAMHD1	High SAMHD1
LUSC	Low SAMHD1	Low SAMHD1	Low SAMHD1
MESO	High SAMHD1	High SAMHD1	High SAMHD1
PAAD	Low SAMHD1	Low SAMHD1	High SAMHD1
PCPG	High SAMHD1	High SAMHD1	Low SAMHD1
READ	High SAMHD1	High SAMHD1	High SAMHD1
SARC	High SAMHD1	High SAMHD1	High SAMHD1
SKCM	Low SAMHD1	High SAMHD1	Low SAMHD1
STAD	Low SAMHD1	Low SAMHD1	High SAMHD1
THCA	High SAMHD1	High SAMHD1	High SAMHD1
THYM	Low SAMHD1	Low SAMHD1	Low SAMHD1
UVM	Low SAMHD1	Low SAMHD1	Low SAMHD1

C

High SAMHD1
Low SAMHD1
Not significant

	All patients	Females	Males
ACC	High SAMHD1		High SAMHD1
BLCA	Low SAMHD1		Low SAMHD1
HNSC	High SAMHD1		High SAMHD1
KIRC	Low SAMHD1		Low SAMHD1
KIRP		Low SAMHD1	
LAML	Low SAMHD1	Low SAMHD1	Low SAMHD1
LGG	Low SAMHD1	Low SAMHD1	Low SAMHD1
LUAD	High SAMHD1		
MESO	High SAMHD1		High SAMHD1
PAAD		Low SAMHD1	
READ	High SAMHD1	High SAMHD1	
SARC	High SAMHD1	High SAMHD1	High SAMHD1
STAD		Low SAMHD1	
THCA	High SAMHD1	High SAMHD1	
UVM	Low SAMHD1		

D

	5-year survival rate			**5-year survival rate**		
	female			**male**		
	SAMHD1	**SAMHD1**		**SAMHD1**	**SAMHD1**	
	high	**low**	**p-value**	**high**	**low**	**p-value**
KIRC	72%	55%	0.056	50%	68%	0.009
KIRP	26%	81%	0.0001	81%	62%	0.073
PAAD	7%	51%	0.031	35%	34%	0.11
STAD	37%	50%	0.0036	41%	28%	0.13

Figure 2. *SAMHD1* expression levels and 5-year survival rates in dependence of sex based on TCGA data. A) Kaplan Meier plots indicating sex-specific survival in cancer patients with tumours characterised by high or low *SAMHD1* expression (as determined by best separation). P-values were determined by log-rank test. B) Heatmap indicating the association of *SAMHD1* expression and 5-year survival rates (blue: high SAMHD1 associated with higher survival rates, yellow: low SAMHD1 associated with higher survival rates). C) Heatmap indicating cancer entities in which high *SAMHD1* expression (blue) or low *SAMHD1* expression (yellow) is significantly (p<0.05) associated with higher 5-year survival rates. D) Cancer entities in which SAMHD1 displays a trend towards differing roles by sex. Blue indicates higher survival rates in patients with tumours with high SAMHD1 levels, yellow in patients with low SAMHD1 levels. Abbreviations: ACC, Adrenocortical carcinoma; BLCA, Bladder Urothelial Carcinoma; BRCA, Breast invasive carcinoma; CESC, Cervical squamous cell carcinoma and endocervical adenocarcinoma; CHOL, Cholangiocarcinoma; COAD, Colon adenocarcinoma; DLBC,

Lymphoid Neoplasm Diffuse Large B-cell Lymphoma; ESCA, Oesophageal carcinoma; GBM, Glioblastoma multiforme; HNSC, Head and Neck squamous cell carcinoma; KICH, Kidney Chromophobe; KIRC, Kidney renal clear cell carcinoma; KIRP, Kidney renal papillary cell carcinoma; LAML, Acute Myeloid Leukaemia; LGG, Low Grade Glioma; LIHC, Liver hepatocellular carcinoma; LUAD, Lung adenocarcinoma; LUSC, Lung squamous cell carcinoma; MESO, Mesothelioma; OV, Ovarian serous cystadenocarcinoma; PAAD, Pancreatic adenocarcinoma; PCPG, Pheochromocytoma and Paraganglioma; PRAD, Prostate adenocarcinoma; READ, Rectum adenocarcinoma; SARC, Sarcoma; SKCM, Skin Cutaneous Melanoma; STAD, Stomach adenocarcinoma; TGCT, Testicular Germ Cell Tumours; THCA, Thyroid carcinoma; THYM, Thymoma; UCEC, Uterine Corpus Endometrial Carcinoma; UCS, Uterine Carcinosarcoma; UVM, Uveal Melanoma.

In TARGET, high SAMHD1 levels were significantly associated with increased survival across all cancer types in females (Figure 3A) but not in males (Figure 3B). For seven paediatric cancer types, data were available for both sexes. When we did not consider statistical significance levels, higher 5-year survival rates were recorded for SAMHD1 patients with SAMHD1 high tumours across all patients, in one entity in female patients, and in four entities in male patients (Figure 3B, Supplementary Table 4.

When we only considered cancer types in which the 5-year survival rates were significantly different ($p<0.05$) between high and low SAMHD1-expressing tumours for at least one comparison (across all patients, in females, and/ or males), differences reached significance only for females in two cancer types (Figure 3C, Supplementary Table 4). Taken together, these data suggest that the role of SAMHD1 in cancer may differ between the sexes in some cancer types.

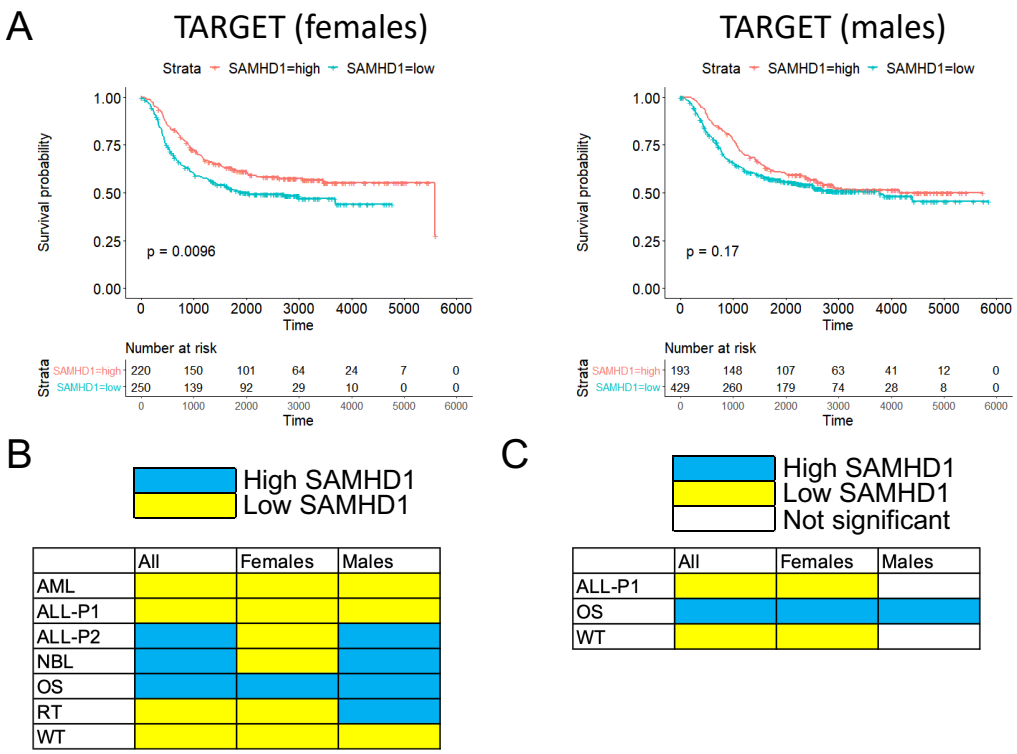


Figure 3. *SAMHD1* expression levels and 5-year survival rates in dependence of sex based on TARGET data. A) Kaplan Meier plots indicating sex-specific survival in cancer patients with tumours

151 characterised by high or low *SAMHD1* expression (as determined by best separation). P-values were
152 determined by log-rank test. B) Heatmap indicating the association of *SAMHD1* expression and 5-year
153 survival rates (blue: high SAMHD1 associated with higher survival rates, yellow: low SAMHD1
154 associated with higher survival rates). C) Heatmap indicating cancer entities in which high *SAMHD1*
155 expression (blue) or low *SAMHD1* expression (yellow) is significantly ($p<0.05$) associated with higher
156 5-year survival rates.

157

158 Notably, the potentially different roles between SAMHD1 in female and male cancer patients do not
159 appear to be the consequence of sex-specific discrepancies in *SAMHD1* expression. In TCGA, there
160 were no significant sex-specific differences in *SAMHD1* expression between tumour samples and
161 matched normal tissue samples from females and males, when we excluded sex-specific cancer types
162 (CESC, OV, PRAD, TGCT, UCEC, UCS) and BRCA (only 12 out of 1,089 tumour tissue samples from male
163 patients, only one out of 113 matched normal tissue samples from a male) (Supplementary Figure 1).
164 SAMHD1 expression was also not significantly different in males and females in the TARGET database
165 (Supplementary Figure 2).

166

167 **Role of *SAMHD1* expression in the context of race**

168 The vast majority of data in TCGA and TARGET are derived from white individuals, which reduces the
169 significance of the race-related data. In TCGA, high SAMHD1 expression was associated with reduced
170 overall survival in white patients (Figure 4). This reflects the findings obtained across all patients
171 (Figure 1A) and probably that 6,834 (82%) out of 8,319 patients, for whom race data are available, are
172 reported to be white. Apart from this, a significant difference in outcome in dependence of tumour
173 SAMHD1 levels was only detected in Native Hawaiian or other Pacific islander patients, in whom high
174 SAMHD1 was associated with improved survival (Figure 4). However, only 13 individuals fell into this
175 category. Cancer-type specific comparisons did not reveal significant differences in SAMHD1-related
176 outcomes between racial groups (Supplementary Figure 3, Supplementary Table 5), which may be due
177 to the low numbers of patients in most of the categories (Figure 4). SAMHD1 levels were generally
178 similar between the different race groups (Supplementary Figure 4). Only Native Hawaiian or other
179 Pacific islander patients displayed increased levels (Supplementary Figure 4).

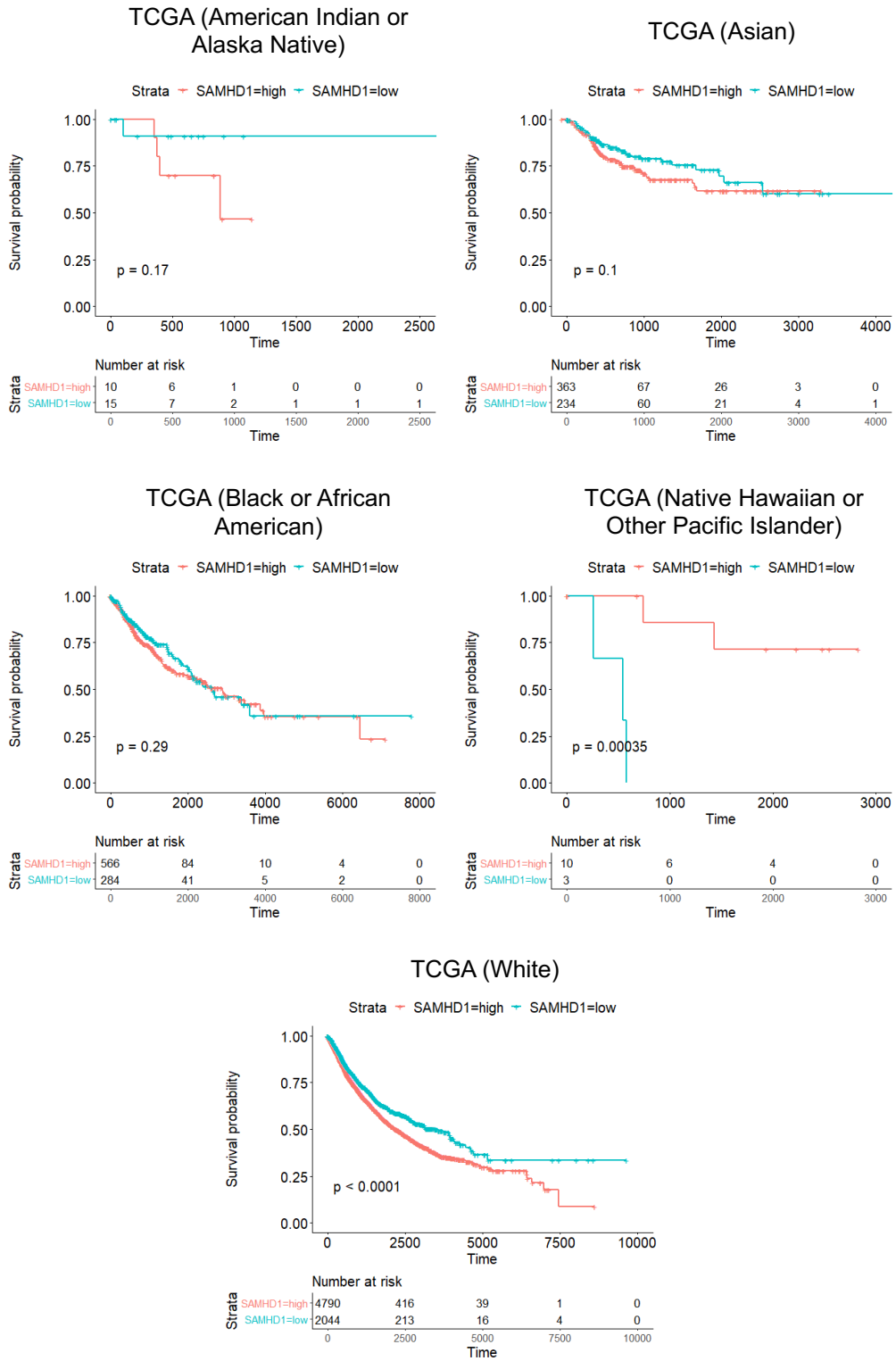
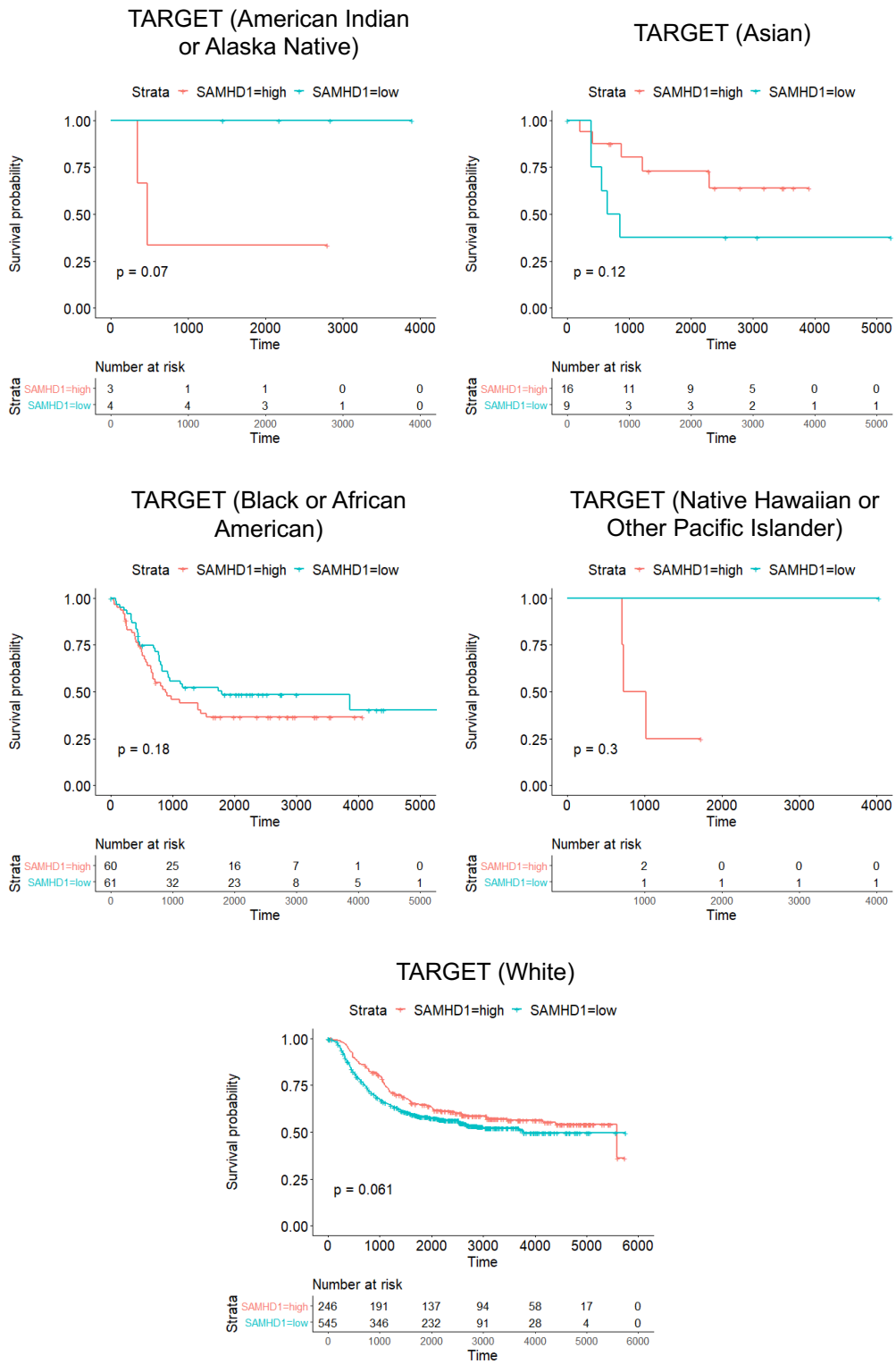


Figure 4. Kaplan-Meier plots indicating survival in cancer patients of different race with tumours characterised by high or low *SAMHD1* expression (as determined by best separation) based on TCGA data. P-values were determined by log-rank test.

185 Stratifying of patients in the TARGET database according to race provided some trends, which may
186 point towards differences, but the numbers are too low to draw firm conclusions (Figure 5,
187 Supplementary Figure 5, Supplementary Table 6). No significant differences were detected between
188 the SAMHD1 levels in the different race groups (Supplementary Figure 6).

189



190

191 **Figure 5**. Kaplan-Meier plots indicating survival in cancer patients of different race with tumours
192 characterised by high or low *SAMHD1* expression (as determined by best separation) based on TARGET
193 data. P-values were determined by log-rank test.

194

195 ***SAMHD1* expression in tumour vs matched normal samples**

196 To further investigate the potential role of SAMHD1 in cancer, we next compared *SAMHD1* expression
197 data in tumour tissue and matched normal samples, which were available for 695 patients and 21
198 cancer types in TCGA. Across all patients, there was no significant difference between the *SAMHD1*
199 FPKM (fragments per kilobase of transcript per million mapped reads) values of tumour samples and
200 matched normal samples (Wilcoxon signed-rank test p-value = 0.14). However, when stratifying by
201 cancer type, SAMHD1 levels significantly differed ($p<0.05$) between tumour samples and matched
202 control samples in seven cancer types (Supplementary Table 7). SAMHD1 was higher in matched
203 control samples suggesting tumour suppressor activity in three cancer types (BLCA, LUAD, LUSC) and
204 higher in tumour samples from four cancer types (KIRC, KIRP, LIHC, STAD) suggesting oncogenic action
205 (Supplementary Table 7).

206 Next, we compared the results on potential tumour suppressor or oncogenic functions of SAMHD1
207 from tumour and matched normal tissues (Supplementary Table 7) to those obtained from analysing
208 5-year survival in cancer patients with SAMHD1 low or high tumours (Figure 2; Supplementary Table
209 1). When we did not consider statistical significance levels, SAMHD1 levels were higher in control
210 tissues suggesting tumour suppressor activity in ten cancer entities (Figure 6A, Supplementary Table
211 7). In twelve of the 21 cancer types, both SAMHD1 levels in tumour and matched normal tissues and
212 the relationship of 5-year survival and tumour SAMHD1 levels indicated a similar role of SAMHD1, i.e.
213 tumour suppressor (higher *SAMHD1* expression in matched normal tissue, higher 5-year survival in
214 patients with SAMHD1 high tumours) or oncogenic (higher *SAMHD1* expression in tumour tissues,
215 higher 5-year survival in patients with SAMHD1 low tumours) activity (Figure 6A, Supplementary Table
216 7).

217 In the next step, only cancer entities were considered for which at least one of the comparisons had
218 resulted in a statistically significant ($p<0.05$) difference, leaving 13 cancer types (Figure 6B,
219 Supplementary Table 7). In seven of these 13 cancer types, the anticipated role of SAMHD1 (tumour
220 suppressor or oncogenic) coincided between both comparisons (Figure 6B, Supplementary Table 7).

221 In only three cancer entities (BLCA, KIRC, LUAD), the differences reached statistical significance for
222 both comparisons (Figure 6C, Supplementary Table 7). SAMHD1 consistently displayed oncogenic
223 activity in KIRC (Kidney renal clear cell carcinoma) and tumour suppressor activity in LUAD (Lung
224 adenocarcinoma). In BLCA (Bladder Urothelial Carcinoma), higher SAMD1 levels in matched normal
225 tissue samples suggested tumour suppressor activity, whereas higher 5-year survival in patients with
226 SAMHD1 low tumours suggested oncogenic effects (Figure 6C, Supplementary Table 7). Hence,
227 SAMHD1 may exert oncogenic activity in KIRC and tumour suppressor activity in LUAD, but clear
228 evidence is lacking for other cancer entities.

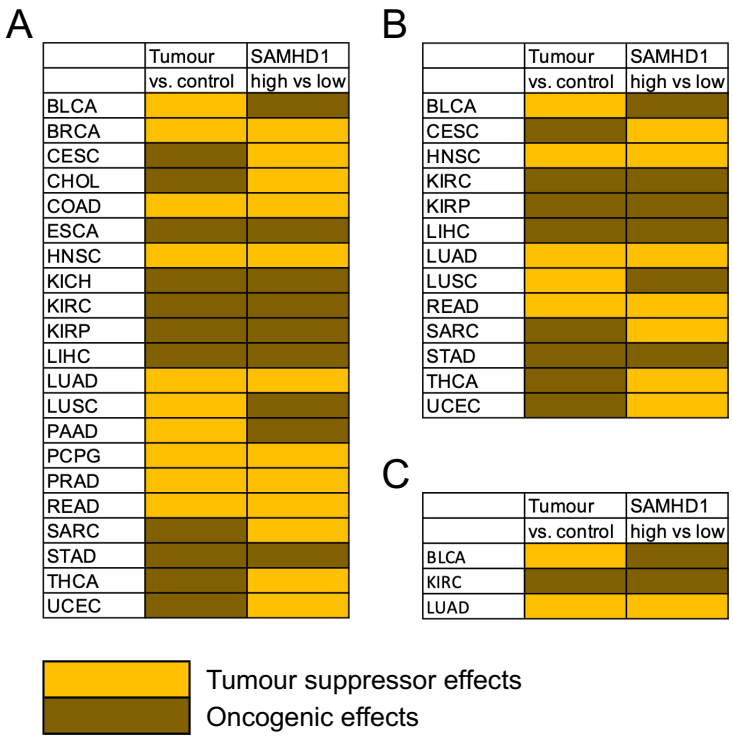


Figure 6. Tumour suppressor and oncogenic effects of SAMHD1 in different cancer types, as suggested by SAMHD1 levels in tumour tissues vs. matched normal tissues (Tumour vs. control) or the comparison of 5-year survival in patients with SAMHD1 high or low tumours (SAMHD1 high vs. low). Higher SAMHD1 levels in matched normal tissues were interpreted as tumour suppressor activity, while higher SAMHD1 levels in tumour tissues as indication of oncogenic effects. Higher 5-year survival in patients with SAMHD1 high tumours was construed as sign of tumour suppressor activity, higher 5-year survival in patients with SAMHD1 low tumours indication of oncogenic effects. A) Data for all available comparisons. B) Data for entities, in which at least the difference for one comparison reached statistical significance. C) Data for entities, in which the difference for both comparisons reached statistical significance.

SAMHD1 regulation by methylation and miRNAs

Promotor methylation and miRNAs have been described to be involved in SAMHD1 regulation [de Silva et al., 2013; Kohnken et al., 2017; Chen et al., 2021]. Tumour and normal sample SAMHD1 expression and promoter methylation beta values were available for 18 cancer types in TCGA. SAMHD1 promotor methylation significantly inversely correlated with SAMHD1 expression levels across all patients, but the correlation coefficient was moderate and the relationship appears weak (Figure 7A).

When we looked at the individual cancer types, an inverse correlation between SAMHD1 expression and promotor methylation was detected in 18 cancer entities (Supplementary Table 8). In eleven of these cancer types, the inverse correlations displayed p-values < 0.05 (Supplementary Table 8). The cancer types with the strongest inverse correlations between SAMHD1 expression and promotor methylation were TGCT (Testicular Germ Cell Tumours), THYM (Thymoma), and CHOL

253 (Cholangiocarcinoma) (Figure 7A). There were also 15 cancer entities with a direct correlation
254 between SAMHD1 expression and promotor methylation, only four of which were associated with a
255 p-value < 0.05. These data suggest that promotor methylation is one SAMHD1 regulation mechanism
256 among others and that its role differs between cancer types.

257 Eight miRNAs (mir-30a, mir-155, and six subtypes of mir-181) have been described to be involved in
258 SAMHD1 regulation [Jin et al., 2014; Pilakka-Kanthikeel et al., 2015; Kohnken et al., 2017; Riess et al.,
259 2017]. Moreover, 21 miRNAs were indicated to interact with SAMHD1 in the DIANA-TarBase v8
260 [http://www.microrna.gr/tarbase] [Karagkouni et al., 2018], an online resource that lists
261 experimentally validated miRNA/mRNA interactions. After the removal of overlaps, this resulted in a
262 list of 28 miRNAs with a documented effect on SAMHD1 (Supplementary Table 9).

263 Of these 28 miRNAs, 26 miRNAs were found to be inversely correlated with SAMHD1 expression in
264 one (mir-21) to 18 (mir-183) cancer entities (Supplementary Table 10). Six miRNAs (mir-23b, mir-30a,
265 mir-192, mir-181d, mir-218-1, mir-218-2) were significantly ($p<0.05$) inversely correlated with
266 SAMHD1 across all patients, with mir-23b showing the strongest inverse correlation ($R= -0.14$,
267 $p=2.2x10^{-16}$) (Figure 7B, Supplementary Table 10). The strongest inverse correlation was detected
268 between mir-30c-1 and SAMHD1 in THYM ($R= -0.85$, $p=0.02$) (Figure 7B, Supplementary Table 10).

269 Each of these 28 miRNAs were found to be inversely correlated with SAMHD1 expression in between
270 two (mir-155) and all 28 (mir-23b and mir-183) cancer entities (Supplementary Table 10). Six miRNAs
271 (mir-23b, mir-30a, mir-192, mir-181d, mir-218-1, mir-218-2) were significantly ($p<0.05$) inversely
272 correlated with SAMHD1 across all patients, with mir-23b showing the strongest inverse correlation
273 (Figure 7B, Supplementary Table 10). The strongest inverse correlation in a cancer type was detected
274 between mir-23b and SAMHD1 in TGCT ($R= -0.54$, $p=7.37x10^{-13}$) (Figure 7B, Supplementary Table 10).

275 Taken together, SAMHD1 levels are determined by complex regulation mechanisms that include
276 promotor methylation and miRNAs, together with post-translational modifications such as
277 phosphorylation and acetylation that have also been described [Coggins et al., 2020; Chen et al.,
278 2021].

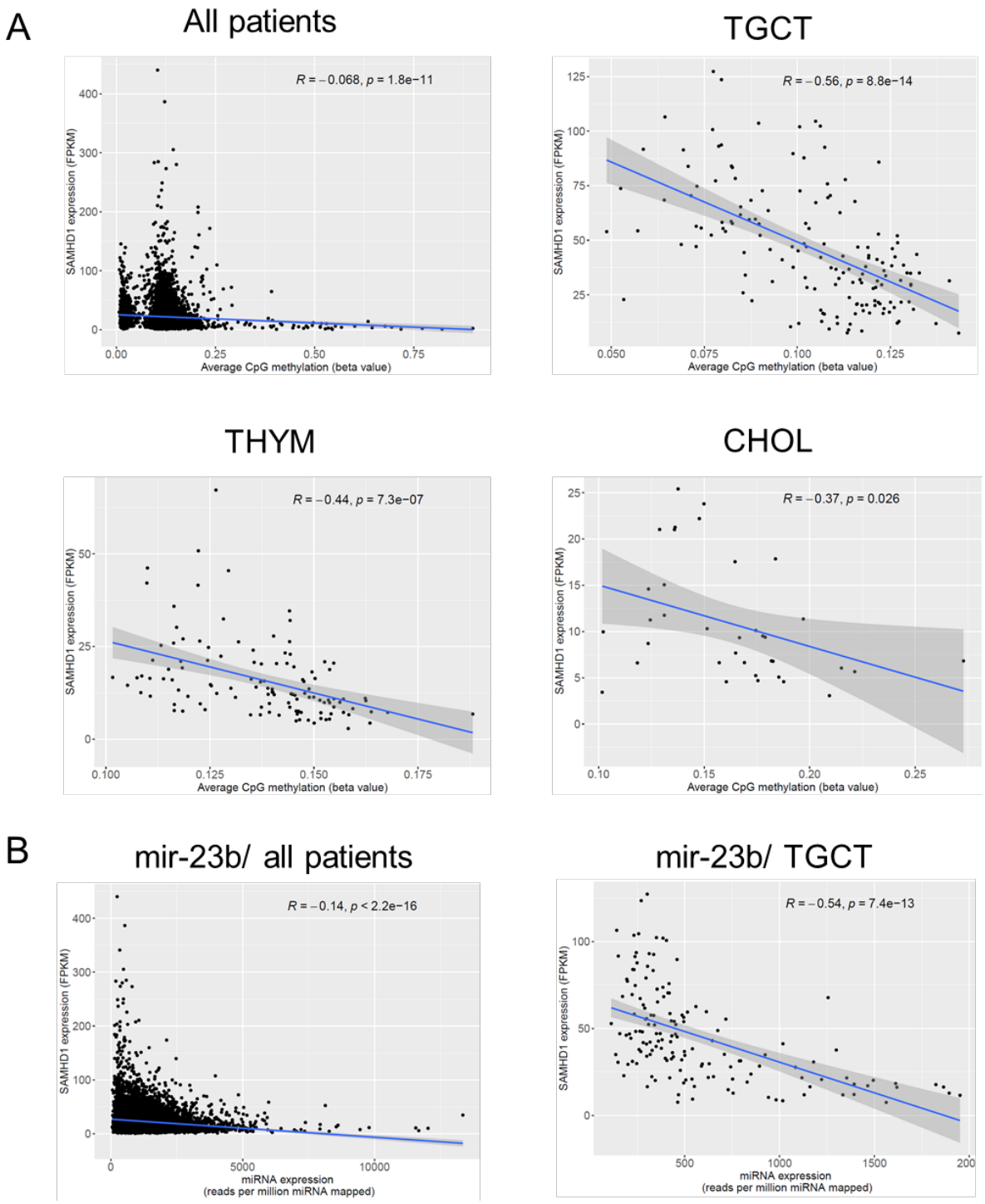


Figure 7. Inverse correlation between SAMHD1 promotor methylation levels or miRNA levels and SAMHD1 expression based on TCGA data. A) Correlation between SAMHD1 promotor methylation levels and SAMHD1 expression across all patients and in THYM patients, which displayed the strongest inverse correlation across all cancer types. Data for all cancer types are presented in Supplementary Table 8. B) Correlation of mir-23b with SAMHD1 expression across all patients and of mir-30c-1 with SAMHD1 in THYM. mir-23b was the miRNA that displayed the strongest inverse correlation with SAMHD1 across all patients. The inverse correlation between mir-30c-1 and SAMHD1 was the

287 strongest among all miRNAs in all cancer types. Data for all significant inverse correlations of miRNAs
288 and SAMHD1 across all cancer types are provided in Supplementary Table 10.

289

290 ***SAMHD1* mutations and patient survival**

291 *SAMHD1* mutations have been described in cancers including chronic lymphocytic leukaemia, T-cell
292 prolymphocytic leukaemia, mantle cell lymphoma, cutaneous T-cell lymphoma and colon cancer
293 [Clifford et al., 2014; Guièze et al., 2015; Merati et al., 2015; Amin et al., 2016; Rentoft et al., 2016;
294 Burns et al., 2018; Johansson et al., 2018; Nadeu et al., 2020; Bühler et al., 2021; Roider et al., 2021].

295 Mutation data was available for 10,149 patients in the TCGA. 15,351 out of 21,156 genes harboured
296 at least one non-synonymous mutation in one patient (Supplementary Table 11). The three most
297 commonly mutated genes were *TTN*, *MUC16*, and *TP53* (Supplementary Table 11). *TTN* and *MUC16*
298 encode the two longest human proteins (36,800 and 14,500 amino acids, respectively) that are
299 frequently found mutated. Mutations in these genes are commonly regarded not to be of functional
300 relevance and removed as artefacts or used as indicators of the mutational burden of tumours, while
301 *TP53* is known to be the most commonly mutated tumour suppressor gene [Lawrence et al., 2013;
302 Kim et al., 2017; Levine, 2020; Oh et al., 2020; Wang et al., 2020; Yang et al., 2020].

303 In total, *SAMHD1* was mutated 201 times, including 175 non-synonymous mutations in 159 patients
304 (1.57% of patients for whom mutation data was available) (Supplementary Table 11). This places
305 *SAMHD1* within the top 15.3% of most commonly mutated genes (Figure 8A, Supplementary Table
306 11). Among the 135 patients with *SAMHD1* mutant tumours for whom survival data were available,
307 *SAMHD1* mutations were associated with superior outcome (Figure 8B, Supplementary Table 12). In
308 18 of the 25 cancer types, in which *SAMHD1* mutations were detected, 5-year survival was higher in
309 patients with *SAMHD1* mutant tumours (Supplementary Table 12). However, the significance of these
310 data is limited due to the low number of SAMHD1 mutations. Notably, the p-value (0.07) was close to
311 significance in UCEC, the cancer type with the most SAMHD1 mutations (35/ 6.6% out of 527), in which
312 93.2% of patients with SAMHD1 mutant cancers survived for five years, in contrast to 76.1% of the
313 492 UCEC patients with SAMHD1 wild-type cancers (Figure 8C, Supplementary Table 12).

314 Although it is not possible to draw firm conclusions from these data, they do not support a general
315 tumour suppressive role of SAMHD1, as mutations in tumour suppressor genes would rather be
316 expected to result in shorter survival. For example, mutations in *TP53*, the most commonly mutated
317 tumour suppressor gene [Levine, 2020], were associated with reduced survival (Figure 8D).

318

319 ***SAMHD1* mutations are likely to be deleterious**

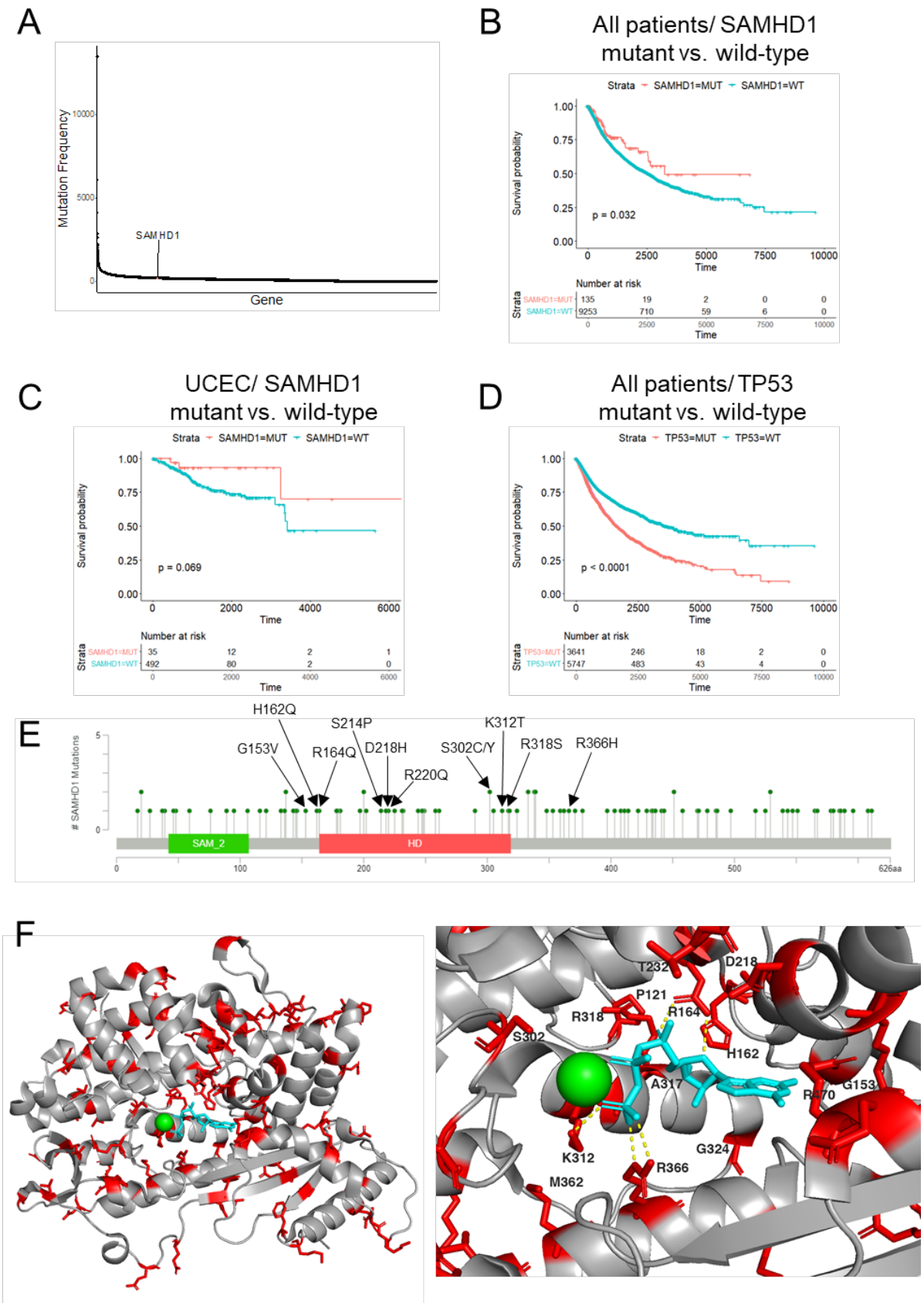
320 Twenty-nine of the mutations are likely to result in a loss of function, including 11 stop-gain, 11
321 frameshift, six splice site and one stop loss mutation. While 21 mutations were located in untranslated
322 regions (six 5' UTR, 15 3' UTR), four were in introns, three in-frame, 25 were synonymous, with the
323 remaining 104 resulting in nonsynonymous mutations.

324 50 mutations had already been described in cancer cells or were present in positions that had been
325 found mutated in cancer cells (Supplementary Table 11). Three of the *SAMHD1* mutations identified
326 in the TCGA (R143C/ UCEC patient, R145Q/ COAD patient, R290H/ STAD patient) were loss-of-function
327 mutations associated with Aicardi-Goutières syndrome [Rice et al., 2009; Mauney & Hollis, 2018;
328 Coggins et al., 2020; UniProt Consortium, 2021]. 18 nonsynonymous mutations occurred at positions

329 demonstrated to be important for SAMHD1 function by mutagenesis studies according to UniProt
330 (Figure 8E, Figure 8F, Supplementary Table 11). This was supported by structural analysis which
331 showed that ten non-synonymous mutations were located around the SAMHD1 active ligand binding
332 sites (Figure 8F).

333 Next, the SAMHD1 non-synonymous variants were analysed using SIFT [Sim et al., 2012], PolyPhen-2
334 [Ng & Henikoff, 2001; Sim et al., 2012; Vaser et al., 2016], Condel [González-Pérez & López-Bigas,
335 2011], and CADD [Kircher et al., 2014] to predict if they are likely to have an effect on protein function
336 (Supplementary Table 11). Approximately half of the amino acid changes were predicted to have a
337 significant impact on SAMHD1 function (SIFT: 63/104 (60.6%), Polyphen-2: 50/104 (48.1%) and
338 Condel: 54/104 (51.9%)). 72 of these variants also had a scaled CADD score of >20, which rates a
339 variant among the top 1% of the most deleterious changes. Five variants displayed CADD scores >30
340 (Figure 8E). 39 variants had a SIFT rating of 'tolerated', a PolyPhen-2 rating of 'benign' and a Condel
341 rating of 'neutral', of which 13 had a scaled CADD score of <10. Predictions for the remaining 17 amino
342 acid changes were inconsistent (i.e. contrasting SIFT, PolyPhen-2 and Condel predictions)
343 (Supplementary Table 11).

344 Taken together, many of the mutations appear to affect SAMHD1 function. However, loss-of-function
345 should typically be associated with reduced survival in tumour suppressor genes.



346

347 **Figure 8**. *SAMHD1* mutations in cancer tissues. A) *SAMHD1* was mutated 201 times, including 175 non-
348 synonymous mutations, which puts *SAMHD1* within the 15.3% of most commonly mutated genes. B)

349 Survival in patients with and without *SAMHD1* mutant tumours. C) Survival in UCEC (cancer type with
350 the most *SAMHD1* mutations) patients with and without *SAMHD1* mutant tumours. D) Survival in
351 patients with tumours with or without mutations in *TP53*, the most commonly mutated tumour
352 suppressor genes. E) Lollipop plot indicating locations of missense mutations in *SAMHD1*. Residues
353 predicted to be involved in ligand binding are labelled in bold. F) nonsynonymous mutations mapped
354 (coloured red) onto the SAMHD1 protein structure (Protein Databank identifier 6DWD [Knecht et al.,
355 2018] with bound clofarabine hydrochloride (indicated in cyan) and magnesium ion (green). The image
356 on the left shows the full structure and on the right the active site is displayed. Yellow dashed lines
357 indicate hydrogen bonds between mutated residues and ligand.

358

359 **Literature review SAMHD1 and cancer**

360 Our analysis of TCGA and TARGET data do not suggest that SAMHD1 generally functions as a tumour
361 suppressor protein. Neither *SAMHD1* mutations nor low SAMHD1 levels were consistently associated
362 with reduced survival. However, SAMHD1 is discussed as a potential tumour suppressor protein in the
363 literature [Herold et al., 2017a; Chen et al., 2021]. Next, we performed a systematic review to compare
364 our findings to those from the literature and to gain further insights into the narrative underlying the
365 perceived role of SAMHD1 in cancer.

366 The literature search was performed in PubMed (https://pubmed.ncbi.nlm.nih.gov) on 17th June 2021
367 using the search term "(((Cancer) OR (tumor) OR (tumour))) AND (SAMHD1)". It resulted in 150 hits,
368 including 35 articles with relevant original data and 15 relevant secondary literature articles (reviews,
369 editorials, comments) (Supplementary Figure 7, Supplementary Table 13).

370 The first articles reported on a potential role of SAMHD1 in cancer in 2013 [Clifford et al., 2014; de
371 Silva et al., 2014; Shi et al., 2014]. The first one reported on low SAMHD1 levels in patients with Sézary
372 syndrome, an aggressive subtype of cutaneous T-cell lymphoma, due to SAMHD1 methylation [de Silva
373 et al., 2014], while the second paper described SAMHD1 variants associated with hepatitis B virus- and
374 hepatitis C virus-induced hepatocellular carcinoma [Shi et al., 2014]. The third paper found SAMHD1
375 mutations in chronic lymphocytic leukaemia and proposed that these mutations promote leukaemia
376 development by affecting SAMHD1-mediated DNA repair [Clifford et al., 2014].

377 Among the 35 articles that reported on an association between SAMHD1 and cancer, four did not
378 (entirely) support the 'SAMHD1 is a tumour suppressor' narrative [Yang et al., 2017; Kodigepalli et al.,
379 2018; Shang et al., 2018; Xagoraris et al., 2021]. One study correlated high *SAMHD1* levels with
380 metastasis formation in colorectal cancer [Yang et al., 2017]. Kodigepalli et al., [2018] reported that
381 SAMHD1 knock-out increases acute myeloid leukaemia cell proliferation via PI3K signalling but inhibits
382 tumourigenesis potentially due to a lack of SAMHD1-mediated TNFalpha suppression. Notably, the
383 title of this study exclusively focused on the inhibitory effects of SAMHD1 on leukaemia cell
384 proliferation and did not mention its role in tumourigenesis [Kodigepalli et al., 2018]. One study
385 detected a SAMHD1 increase upon lung cancer progression [Shang et al., 2018], and the most recent
386 study, correlated the presence of SAMHD1 in Hodgkin lymphoma cells with unfavourable outcome
387 [Xagoraris et al., 2021]. Notably, this study [Xagoraris et al., 2021] even referred to SAMHD1 as "novel
388 tumour suppressor" in the title, although the study rather indicated an oncogenic role of SAMHD1.

389 The remaining articles largely focused on SAMHD1 mutations and reduced SAMHD1 levels in different
390 cancer types as well as on SAMHD1's potential role as a tumour suppressor involved in DNA repair
391 (Supplementary Table 13). *SAMHD1* mutations were detected in patients with hepatocellular
392 carcinoma [Shi et al., 2014], chronic lymphocytic leukaemia [Clifford et al., 2014; Guièze et al., 2015;

393 Amin et al., 2016; Kim et al., 2016; Burns et al., 2018], cutaneous T-cell lymphoma [Merati et al., 2015],
394 colorectal cancer [Rentoft et al., 2016], T-cell prolymphocytic leukaemia [Johansson et al., 2018], acute
395 myeloid leukaemia [Zhu et al., 2018], and mantle cell lymphoma [Nadeu et al., 2020; Bühler et al.,
396 2021].

397 In our TCGA analysis performed above (Figure 8, Supplementary Table 12), there is rather a trend
398 towards higher 5-year survival rates among hepatocellular carcinoma (LIHC) patients with *SAMHD1*
399 mutant tumours, although the significance of the data is limited due to low numbers (Supplementary
400 Table 12). All four patients with *SAMHD1* mutant hepatocellular carcinoma survived for five years,
401 while only 51.4% of 359 patients with *SAMHD1* wild-type hepatocellular carcinomas survived for five
402 years. In colorectal adenocarcinoma (COAD), there was no noticeable difference between the survival
403 of patients with *SAMHD1* mutant and *SAMHD1* wild-type tumours (Supplementary Table 12). Only in
404 rectal adenocarcinoma (READ), a trend suggested that patients with *SAMHD1* mutant tumours may
405 have a worse outcome. None out of five patients with *SAMHD1* mutant tumours survived for five
406 years, while 54.9% of 130 patients with *SAMHD1* wild-type tumours did (Supplementary Table 12).

407 Acute myeloid leukaemia (LAML) was the only cancer type in which a significant difference was
408 detected between patients with *SAMHD1* mutant and wild-type cancer cells. Patients with *SAMHD1*
409 mutant leukaemia cells had a higher 5-year survival rate (Supplementary Table 8). However, this is
410 most probably not due to general oncogenic activity, but because lack of SAMHD1 function results in
411 a higher activity of nucleoside analogues including cytarabine and decitabine that are SAMHD1
412 substrates and commonly used for LAML treatment [Schneider et al., 2017; Oellerich et al., 2019].

413 The study on colorectal cancer [Rentoft et al., 2016] was the only one that had used TCGA data.
414 However, it only used TCGA data to identify mutations, but did not compare survival in patients with
415 and without *SAMHD1* mutations.

416 *SAMHD1* expression levels have been suggested to impact on cutaneous T-cell lymphoma [de Silva et
417 al., 2014], lung cancer [Wang et al., 2014; Shang et al., 2018], colorectal cancer [Yang et al., 2017], and
418 acute myeloid leukaemia [Jiang et al., 2020]. Low SAMHD1 levels were described in cutaneous T-cell
419 lymphoma and acute myeloid leukaemia cells [de Silva et al., 2014; Jiang et al., 2020], supporting a
420 potential role as a tumour suppressor. TCGA did not contain data on *SAMHD1* expression in cutaneous
421 T-cell lymphoma or acute myeloid leukaemia cells relative to control cells.

422 In lung cancer, conflicting results were reported. One study found that SAMHD1 is down regulated in
423 lung cancer by methylation and inhibits tumour cell proliferation [Wang et al., 2014]. The other study
424 reported that SAMHD1 levels increase in the serum of lung cancer patients upon progression [Shang
425 et al., 2018]. Our analysis of SAMHD1 data found significantly higher 5-year survival rates in patients
426 with tumours displaying high *SAMHD1* expression levels (Supplementary Table 1) supporting the first
427 study. Notably, elevated SAMHD1 in the serum of lung cancer patients may not have been derived
428 from cancer tissue.

429 In colorectal cancer, low SAMHD1 levels were detected in tumour tissues relative to adjacent control
430 tissues [Yang et al., 2017], which agrees with a tumour suppressor function. However, higher SAMHD1
431 levels were associated with metastasis formation [Yang et al., 2017], rather supporting an oncogenic
432 role. The study included the analysis of TCGA data on colorectal cancer [Yang et al., 2017], but no
433 systematic analysis of SAMHD1 across different cancer entities.

434 The 15 relevant secondary literature articles all had narratives focussing on the potential role of
435 SAMHD1 as a tumour suppressor (Supplementary Table 13).

436

437 **Discussion**

438 SAMHD1 has been suggested to exert tumour suppressor functions due its role in maintaining genome
439 integrity and as an inhibitor of uncontrolled proliferation [Herold et al., 2017; Chen et al., 2021].
440 However, our analysis of TCGA and TARGET data does not suggest that SAMHD1 should be regarded
441 as a *bona fide* tumour suppressor. Notably, SAMHD1 mutations that interfere with SAMHD1 function
442 were not associated with poor outcome, which is something that would be expected from a tumour
443 suppressor. In agreement, no increased cancer formation has been described in SAMHD1-deficient
444 animal models [Kohnken et al., 2015].

445 Our results rather indicated that changes in SAMHD1 are involved in the oncogenic process in a
446 minority of cases and that it may exert pro- or antitumourigenic effects in different cancer types (and
447 perhaps individual tumours). Moreover, the role of SAMHD1 may differ between the sexes and
448 different races. These findings also show that our understanding of the processes underlying cancer
449 needs to improve further, before a broad paradigm shift towards tumour-agnostic approaches [Danesi
450 et al., 2021] can become a reality.

451 Notably, the interpretation of our findings may be affected by SAMHD1 being a triphosphohydrolase
452 that cleaves and inactivates the triphosphorylated forms of a number of nucleoside analogues
453 including cytarabine, decitabine, and nelarabine [Schneider et al., 2017; Oellerich et al., 2019;
454 Rothenburger et al., 2020]. However, most cancer diseases are not treated with SAMHD1 substrates.
455 Notably, KIRC (kidney renal clear cell carcinoma), the only cancer in which 5-year survival is
456 significantly lower in SAMHD1 high tumours and SAMHD1 levels are significantly higher in tumour
457 than in control tissues, suggesting an oncogenic role of SAMHD1, is not treated with nucleoside
458 analogues [Geynisman et al., 2021]. Hence, the absence of tumour suppressor activity and/ or
459 oncogenic effects cannot simply be explained by SAMHD1-mediated inactivation of nucleoside
460 analogue substrates.

461 Our findings demonstrating that SAMHD1 plays multifaceted (and often, if any, minor) roles in cancer
462 seem to be in disaccord with a perception and narrative forming in the field suggesting that SAMHD1
463 is a tumour suppressor [Herold et al., 2017; Chen et al., 2021]. A systematic review confirmed that
464 most of the available literature focuses on a potential role of SAMHD1 as a tumour suppressor. Among
465 35 original articles on the role of SAMHD1 in cancer, 31 discussed a potential tumour suppressor
466 function and three potential oncogenic effects. One article reported both potential tumour suppressor
467 and oncogenic activity, but only mentioned the anticipated tumour suppressor effects in the title
468 [Kodigepalli et al., 2018]. All 15 secondary literature articles (reviews, editorials, comments) had a
469 narrative built around SAMHD1 being a candidate tumour suppressor.

470 The narrative that SAMHD1 is a tumour suppressor has formed since 2013 around findings in a limited
471 number of cancer entities (Supplementary Table 13). Three reasons may contribute to the
472 perpetuation of such a narrative without much scrutiny. Firstly, SAMHD1 has been described to
473 maintain genome integrity by a range of different mechanisms [Herold et al., 2017; Akimova et al.,
474 2021; Chen et al., 2021; Park et al., 2021]. Hence, a potential tumour suppressor role is plausible and
475 convincing. Further research will have to show why SAMHD1's multifaceted roles in DNA repair do not
476 translate into a consistent and general tumour suppressor function.

477 The second potential reason is confirmation bias. Scientists (like everybody else) tend to accept
478 findings that support their own experiences, assumptions, and perceptions and to disregard evidence
479 that challenges them [Letrud & Hernes, 2019; Yanai & Lercher, 2021]. Thus, researchers are more
480 likely to look for data that support their hypothesis and not for those that contradict it. Notably, one

study referred to SAMHD1 as "novel tumour suppressor" in the title, although *SAMHD1* expression was described as an adverse prognostic factor in Hodgkin lymphoma [Xagoraris et al., 2021].

The final potential reason is publication bias, i.e. a focus on 'positive' findings that are easier to publish in more prestigious journals than 'negative' findings [Begley & Ioannidis, 2015; Nissen et al., 2016; Wass et al., 2019; Marks-Anglin & Chen, 2020]. In the case of studies investigating a potential role of SAMHD1 in cancer, this means that some studies that did not find a relationship between SAMHD1 and cancer may simply not have been published and that the publicly available data may not reflect all available data on the subject.

In conclusion, SAMHD1 can play multifaceted roles in cancer that may differ between different cancer types, the sexes, and races. In contradiction to the predominant narrative, SAMHD1 may exert oncogenic as well as tumour suppressor activity and may often be a minor (if any) player in carcinogenesis. Our findings emphasise that hypotheses, perceptions, and assumptions need to be continuously challenged by using all available data and evidence. In this context, it is important that all data are actually published and made available, even if they are not deemed particularly exciting by researchers. Finally, the increasing number of available data and databases should be effectively used to inform and challenge our research and research findings.

499 **Methods**

500 **Gene Expression and Clinical Data**

501 Gene expression data (FPKM values) from patient tumours were derived from The Cancer Genome
502 Atlas (TCGA) [Cancer Genome Atlas Research Network, 2008] via the GDC Data Portal
503 (https://portal.gdc.cancer.gov). The Bioconductor R package TCGAbiolinks was used to obtain
504 corresponding clinical data. Primary tumour gene expression data and clinical response data were
505 available for 9,572 patients (5,037 female, 4,535 male) with 33 different cancer types. Ages at
506 diagnosis ranged from 14 to 90 (median age at diagnosis = 61, no data for 113 patients). Data were
507 also downloaded for 694 matched normal tissue samples.

508 Gene expression (RPKM) values and clinical data were extracted for patients in the TARGET database
509 from the National Cancer Institute Office of Cancer Genomics TARGET data matrix
510 (https://ocg.cancer.gov/programs/target/data-matrix). Primary tumour sample data was available for
511 a total of 1,091 patients in TARGET (470 females and 593 males) with seven cancer types. Ages at
512 diagnosis ranged from six days to 32.41 years (median age at diagnosis was 5.4 years (1976 days)).

513 Tumour vs normal sample gene expression was compared using the wilcox.test function in R, which
514 performs the Mann Whitney U test for independent groups. Pairwise comparisons were made using
515 the Wilcoxon Signed Rank test. Pie charts were generated using ggplot2.

516 **Methylation and miRNA data**

517 TCGA methylation beta values and miRNA expression values (reads per million miRNA mapped) were
518 downloaded from the GDC Data Portal (https://portal.gdc.cancer.gov). Mean methylation beta values
519 for each CpG site in the SAMHD1 promoter for which data were available (cg02078758, cg00642209,
520 cg16430572, cg09128050, cg12099051, cg18861300, cg11094122, cg22769031, cg23888977,
521 cg09717261, cg24951864, cg06097592, cg22583967, cg10804363 and cg12517061) were calculated
522 per individual. Expression data for miRNAs which were listed in DIANA-TarBase v8 [Karagkouni et al.,
523 2018] as being experimentally validated to positively interact with SAMHD1 (n=21) along with eight
524 miRNAs shown in previous experiments to target SAMHD1 [Jin et al., 2014; Pilakka-Kanthikeel et al.,
525 2015; Kohnken et al., 2017; Riess et al., 2017] were extracted for analysis. Scatter plots and associated
526 Pearson correlations for methylation and miRNA expression with SAMHD1 expression were calculated
527 using the ggplot2 package in R.

528 **Survival analyses**

529 Cox proportional hazards regression was used to calculate the hazard ratio for cohorts expressing high
530 levels of SAMHD1. Overall survival (OS) was defined as days to last follow-up or death, as previously
531 described [Ng et al., 2016]. Calculations were performed using the R survminer and survival packages.
532 The 'surv_cutpoint' function was used to identify the optimal expression cut-off point to give the
533 lowest p-value for high vs low expression. We permitted the cut-off to be only between the 20th and
534 80th percentiles of gene expression values, as described by previously [Uhlen et al., 2017].

535 Kaplan-Meier survival curves were generated using R package ggsurvplot. P-values in each case were
536 the result of a log rank (Mantel-Cox) test, which assesses whether there is a significant difference
537 between the survival of two independent groups. Hazard ratios quoted refer to values for 'low' (below
538 the calculated optimal cut-off) expression for each gene in the model, with values >1 indicating
539 increased hazard (i.e. reduced OS) and values <1 indicating decreased hazard (i.e. increased OS).

540 **Mutation Data and Variant Effect Prediction**

Mutation data for 10,149 TCGA patients were downloaded from the GDC Data Portal (https://portal.gdc.cancer.gov). A dot plot displaying mutation frequencies of 21,156 genes was generated using ggplot2.

In order to assess the potential impact of mutations in SAMHD1, we used the online tool Variant Effect Predictor (VEP) [McLaren et al., 2016] to obtain reports from SIFT [Sim et al., 2012], PolyPhen-2 [Adzhubei et al., 2010], Condel [González-Pérez & López-Bigas, 2011] and CADD [Kircher et al., 2014].

A lollipop plot of SAMHD1 mutations was generated using the cBioPortal MutationMapper tool (https://www.cbioportal.org/mutation_mapper) [Cerami et al., 2012].

Literature review

Relevant articles were identified on 17th June 2021 by using the search term "(((Cancer) OR (tumor) OR (tumour))) AND (SAMHD1)" in PubMed (https://pubmed.ncbi.nlm.nih.gov) on the basis of the principles outlined in the PRISMA guidelines (http://prisma-statement.org). Articles in English were included into the analysis, when they contained original data on the role of SAMHD1 in cancer. Moreover, reviews that discussed the potential impact of SAMHD1 on cancer were used to analyse conceptions and the predominant narrative in the field. Two reviewers independently analysed articles for relevant information and then agreed a list of relevant articles.

Conflict of interest

There is no conflict of interest.

560 **References**

561 Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR.
562 A method and server for predicting damaging missense mutations. Nat Methods. 2010 Apr;7(4):248-
563 9. doi: 10.1038/nmeth0410-248.

564 Akimova E, Gassner FJ, Schubert M, Rebhandl S, Arzt C, Rauscher S, Tober V, Zaborsky N, Greil R,
565 Geisberger R. SAMHD1 restrains aberrant nucleotide insertions at repair junctions generated by DNA
566 end joining. Nucleic Acids Res. 2021 Mar 18;49(5):2598-2608. doi: 10.1093/nar/gkab051.

567 Amin NA, Seymour E, Saiya-Cork K, Parkin B, Shedden K, Malek SN. A Quantitative Analysis of Subclonal
568 and Clonal Gene Mutations before and after Therapy in Chronic Lymphocytic Leukemia. Clin Cancer
569 Res. 2016 Sep 1;22(17):4525-35. doi: 10.1158/1078-0432.CCR-15-3103.

570 Begley CG, Ioannidis JP. Reproducibility in science: improving the standard for basic and preclinical
571 research. Circ Res. 2015 Jan 2;116(1):116-26. doi: 10.1161/CIRCRESAHA.114.303819.

572 Bühler MM, Lu J, Scheinost S, Nadeu F, Roos-Weil D, Hensel M, Thavayogarajah T, Moch H, Manz MG,
573 Haralambieva E, Marques Maggio E, Beà S, Giné E, Campo E, Bernard OA, Huber W, Zenz T. SAMHD1
574 mutations in mantle cell lymphoma are recurrent and confer in vitro resistance to nucleoside
575 analogues. Leuk Res. 2021 May 3;107:106608. doi: 10.1016/j.leukres.2021.106608.

576 Burns A, Alsolami R, Becq J, Stamatopoulos B, Timbs A, Bruce D, Robbe P, Vavoulis D, Clifford R, Cabes
577 M, Dreau H, Taylor J, Knight SJL, Mansson R, Bentley D, Beekman R, Martín-Subero JI, Campo E,
578 Houlston RS, Ridout KE, Schuh A. Whole-genome sequencing of chronic lymphocytic leukaemia reveals
579 distinct differences in the mutational landscape between IgHV(mut) and IgHV(unmut) subgroups.
580 Leukemia. 2018 Feb;32(2):332-342. doi: 10.1038/leu.2017.177.

581 Cancer Genome Atlas Research Network. Comprehensive genomic characterization defines human
582 glioblastoma genes and core pathways. Nature. 2008 Oct 23;455(7216):1061-8. doi:
583 10.1038/nature07385.

584 Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson
585 E, Antipin Y, Reva B, Goldberg AP, Sander C, Schultz N. The cBio cancer genomics portal: an open
586 platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012 May;2(5):401-4.
587 doi: 10.1158/2159-8290.CD-12-0095.

588 Chen Z, Hu J, Ying S, Xu A. Dual roles of SAMHD1 in tumor development and chemoresistance to
589 anticancer drugs. Oncol Lett. 2021 Jun;21(6):451. doi: 10.3892/ol.2021.12712.

590 Clifford R, Louis T, Robbe P, Ackroyd S, Burns A, Timbs AT, Wright Colopy G, Dreau H, Sigaux F, Judde
591 JG, Rotger M, Telenti A, Lin YL, Pasero P, Maelfait J, Titsias M, Cohen DR, Henderson SJ, Ross MT,
592 Bentley D, Hillmen P, Pettitt A, Rehwinkel J, Knight SJ, Taylor JC, Crow YJ, Benkirane M, Schuh A.
593 SAMHD1 is mutated recurrently in chronic lymphocytic leukemia and is involved in response to DNA
594 damage. Blood. 2014 Feb 13;123(7):1021-31. doi: 10.1182/blood-2013-04-490847.

595 Coggins SA, Mahboubi B, Schinazi RF, Kim B. SAMHD1 Functions and Human Diseases. Viruses. 2020
596 Mar 31;12(4):382. doi: 10.3390/v12040382.

597 Danesi R, Fogli S, Indraccolo S, Del Re M, Dei Tos AP, Leoncini L, Antonuzzo L, Bonanno L, Guarneri V,
598 Pierini A, Amunni G, Conte P. Druggable targets meet oncogenic drivers: opportunities and limitations
599 of target-based classification of tumors and the role of Molecular Tumor Boards. ESMO Open. 2021
600 Apr;6(2):100040. doi: 10.1016/j.esmoop.2020.100040.

601 de Silva S, Wang F, Hake TS, Porcu P, Wong HK, Wu L. Downregulation of SAMHD1 expression
602 correlates with promoter DNA methylation in Sézary syndrome patients. J Invest Dermatol. 2014
603 Feb;134(2):562-565. doi: 10.1038/jid.2013.311.

604 Geynisman DM, Maranchie JK, Ball MW, Bratslavsky G, Singer EA. A 25 year perspective on the
605 evolution and advances in an understanding of the biology, evaluation and treatment of kidney cancer.
606 Urol Oncol. 2021 Jun 4:S1078-1439(21)00194-0. doi: 10.1016/j.urolonc.2021.04.038.

607 González-Pérez A, López-Bigas N. Improving the Assessment of the Outcome of Nonsynonymous SNVs
608 with a Consensus Deleteriousness Score, Condel. Am J Hum Genet. 2011 Apr 8;88(4):440-9. doi:
609 10.1016/j.ajhg.2011.03.004.

610 Guièze R, Robbe P, Clifford R, de Guibert S, Pereira B, Timbs A, Dilhuydy MS, Cabes M, Ysebaert L,
611 Burns A, Nguyen-Khac F, Davi F, Véronèse L, Combes P, Le Garff-Tavernier M, Leblond V, Merle-Béral
612 H, Alsolami R, Hamblin A, Mason J, Pettitt A, Hillmen P, Taylor J, Knight SJ, Tournilhac O, Schuh A.
613 Presence of multiple recurrent mutations confers poor trial outcome of relapsed/refractory CLL.
614 Blood. 2015 Oct 29;126(18):2110-7. doi: 10.1182/blood-2015-05-647578.

615 Herold N, Rudd SG, Sanjiv K, Kutzner J, Myrberg IH, Paulin CBJ, Olsen TK, Helleday T, Henter JI, Schaller
616 T. With me or against me: Tumor suppressor and drug resistance activities of SAMHD1. Exp Hematol.
617 2017a Aug;52:32-39. doi: 10.1016/j.exphem.2017.05.001.

618 Herold N, Rudd SG, Sanjiv K, Kutzner J, Bladh J, Paulin CBJ, Helleday T, Henter JI, Schaller T. SAMHD1
619 protects cancer cells from various nucleoside-based antimetabolites. Cell Cycle. 2017b Jun
620 3;16(11):1029-1038. doi: 10.1080/15384101.2017.1314407.

621 Jin C, Peng X, Liu F, Cheng L, Lu X, Yao H, Wu H, Wu N. MicroRNA-181 expression regulates specific
622 post-transcriptional level of SAMHD1 expression in vitro. Biochem Biophys Res Commun. 2014 Sep
623 26;452(3):760-7. doi: 10.1016/j.bbrc.2014.08.151.

624 Jiang H, Li C, Liu Z, Shengjing Hospital; Hu. Expression and Relationship of SAMHD1 with Other
625 Apoptotic and Autophagic Genes in Acute Myeloid Leukemia Patients. Acta Haematol. 2020;143(1):51-
626 59. doi: 10.1159/000500822.

627 Johansson P, Klein-Hitpass L, Choidas A, Habenberger P, Mahboubi B, Kim B, Bergmann A, Scholtysik
628 R, Brauser M, Lollies A, Siebert R, Zenz T, Dührsen U, Küppers R, Dürig J. SAMHD1 is recurrently
629 mutated in T-cell prolymphocytic leukemia. Blood Cancer J. 2018 Jan 19;8(1):11. doi: 10.1038/s41408-
630 017-0036-5.

631 Karagkouni D, Paraskevopoulou MD, Chatzopoulos S, Vlachos IS, Tastsoglou S, Kanellos I,
632 Papadimitriou D, Kavakiotis I, Maniou S, Skoufos G, Vergoulis T, Dalamagas T, Hatzigeorgiou AG.
633 DIANA-TarBase v8: a decade-long collection of experimentally supported miRNA-gene interactions.
634 ucleic Acids Res. 2018 Jan 4;46(D1):D239-D245. doi: 10.1093/nar/gkx1141.

635 Kim YA, Madan S, Przytycka TM. WeSME: uncovering mutual exclusivity of cancer drivers and beyond.
636 Bioinformatics. 2017 Mar 15;33(6):814-821. doi: 10.1093/bioinformatics/btw242.

637 Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating
638 the relative pathogenicity of human genetic variants. Nat Genet. 2014 Mar;46(3):310-5. doi:
639 10.1038/ng.2892.

640 Knecht KM, Buzovetsky O, Schneider C, Thomas D, Srikanth V, Kaderali L, Tofoleanu F, Reiss K, Ferreirós
641 N, Geisslinger G, Batista VS, Ji X, Cinatl J Jr, Keppler OT, Xiong Y. The structural basis for cancer drug

642 interactions with the catalytic and allosteric sites of SAMHD1. Proc Natl Acad Sci U S A. 2018 Oct
643 23;115(43):E10022-E10031. doi: 10.1073/pnas.1805593115.

644 Kodigepalli KM, Bonifati S, Tirumuru N, Wu L. SAMHD1 modulates in vitro proliferation of acute
645 myeloid leukemia-derived THP-1 cells through the PI3K-Akt-p27 axis. Cell Cycle. 2018;17(9):1124-
646 1137. doi: 10.1080/15384101.2018.1480218.

647 Kohnken R, Kodigepalli KM, Wu L. Regulation of deoxynucleotide metabolism in cancer: novel
648 mechanisms and therapeutic implications. Mol Cancer. 2015 Sep 29;14:176. doi: 10.1186/s12943-015-
649 0446-6.

650 Kohnken R, Kodigepalli KM, Mishra A, Porcu P, Wu L. MicroRNA-181 contributes to downregulation of
651 SAMHD1 expression in CD4+ T-cells derived from Sèzary syndrome patients. Leuk Res. 2017 Jan;52:58-
652 66. doi: 10.1016/j.leukres.2016.11.010.

653 Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, Sivachenko A, Carter SL, Stewart C, Mermel
654 CH, Roberts SA, Kiezun A, Hammerman PS, McKenna A, Drier Y, Zou L, Ramos AH, Pugh TJ, Stransky N,
655 Helman E, Kim J, Sougnez C, Ambrogio L, Nickerson E, Shefler E, Cortés ML, Auclair D, Saksena G, Voet
656 D, Noble M, DiCara D, Lin P, Lichtenstein L, Heiman DI, Fennell T, Imielinski M, Hernandez B, Hodis E,
657 Baca S, Dulak AM, Lohr J, Landau DA, Wu CJ, Melendez-Zajgla J, Hidalgo-Miranda A, Koren A, McCarroll
658 SA, Mora J, Crompton B, Onofrio R, Parkin M, Winckler W, Ardlie K, Gabriel SB, Roberts CWM, Biegel
659 JA, Stegmaier K, Bass AJ, Garraway LA, Meyerson M, Golub TR, Gordenin DA, Sunyaev S, Lander ES,
660 Getz G. Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature.
661 2013 Jul 11;499(7457):214-218. doi: 10.1038/nature12213.

662 Letrud K, Hernes S. Affirmative citation bias in scientific myth debunking: A three-in-one case study.
663 PLoS One. 2019 Sep 9;14(9):e0222213. doi: 10.1371/journal.pone.0222213.

664 Levine AJ. p53: 800 million years of evolution and 40 years of discovery. Nat Rev Cancer. 2020
665 Aug;20(8):471-480. doi: 10.1038/s41568-020-0262-1.

666 Li N, Zhang W, Cao X. Identification of human homologue of mouse IFN-gamma induced protein from
667 human dendritic cells. Immunol Lett. 2000 Nov 1;74(3):221-4. doi: 10.1016/s0165-2478(00)00276-5.

668 Marks-Anglin A, Chen Y. A historical review of publication bias. Res Synth Methods. 2020
669 Nov;11(6):725-742. doi: 10.1002/jrsm.1452.

670 Mauney CH, Hollis T. SAMHD1: Recurring roles in cell cycle, viral restriction, cancer, and innate
671 immunity. Autoimmunity. 2018 May;51(3):96-110. doi: 10.1080/08916934.2018.1454912.

672 McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A, Flicek P, Cunningham F. The Ensembl
673 Variant Effect Predictor. Genome Biol. 2016 Jun 6;17(1):122. doi: 10.1186/s13059-016-0974-4.

674 Merati M, Buethe DJ, Cooper KD, Honda KS, Wang H, Gerstenblith MR. Aggressive CD8(+)
675 epidermotropic cutaneous T-cell lymphoma associated with homozygous mutation in SAMHD1. JAAD
676 Case Rep. 2015 Jun 25;1(4):227-9. doi: 10.1016/j.jdcr.2015.05.003.

677 Nadeu F, Martin-Garcia D, Clot G, Díaz-Navarro A, Duran-Ferrer M, Navarro A, Vilarrasa-Blasi R, Kulis
678 M, Royo R, Gutiérrez-Abril J, Valdés-Mas R, López C, Chapaprieta V, Puiggros M, Castellano G, Costa
679 D, Aymerich M, Jares P, Espinet B, Muntañola A, Ribera-Cortada I, Siebert R, Colomer D, Torrents D,
680 Gine E, López-Guillermo A, Küppers R, Martin-Subero JI, Puente XS, Beà S, Campo E. Genomic and
681 epigenomic insights into the origin, pathogenesis, and clinical behavior of mantle cell lymphoma
682 subtypes. Blood. 2020 Sep 17;136(12):1419-1432. doi: 10.1182/blood.2020005289.

683 Ng PC, Henikoff S. Predicting deleterious amino acid substitutions. Genome Res. 2001 May;11(5):863-
684 74. doi: 10.1101/gr.176601.

685 Nissen SB, Magidson T, Gross K, Bergstrom CT. Publication bias and the canonization of false facts.
686 Elife. 2016 Dec 20;5:e21451. doi: 10.7554/eLife.21451.

687 Oellerich T, Schneider C, Thomas D, Knecht KM, Buzovetsky O, Kaderali L, Schliemann C, Bohnenberger
688 H, Angenendt L, Hartmann W, Wardelmann E, Rothenburger T, Mohr S, Scheich S, Comoglio F, Wilke
689 A, Ströbel P, Serve H, Michaelis M, Ferreirós N, Geisslinger G, Xiong Y, Keppler OT, Cinatl J Jr. Selective
690 inactivation of hypomethylating agents by SAMHD1 provides a rationale for therapeutic stratification
691 in AML. Nat Commun. 2019 Aug 2;10(1):3475. doi: 10.1038/s41467-019-11413-4.

692 Oh JH, Jang SJ, Kim J, Sohn I, Lee JY, Cho EJ, Chun SM, Sung CO. Spontaneous mutations in the single
693 TTN gene represent high tumor mutation burden. NPJ Genom Med. 2020 Jan 14;5:33. doi:
694 10.1038/s41525-019-0107-6.

695 Park K, Ryoo J, Jeong H, Kim M, Lee S, Hwang SY, Ahn J, Kim D, Moon HC, Baek D, Kim K, Park HY, Ahn
696 K. Aicardi-Goutières syndrome-associated gene SAMHD1 preserves genome integrity by preventing R-
697 loop formation at transcription-replication conflict regions. PLoS Genet. 2021 Apr 15;17(4):e1009523.
698 doi: 10.1371/journal.pgen.1009523.

699 Pilakka-Kanthikeel S, Raymond A, Atluri VS, Sagar V, Saxena SK, Diaz P, Chevelon S, Concepcion M, Nair
700 M. Sterile alpha motif and histidine/aspartic acid domain-containing protein 1 (SAMHD1)-facilitated
701 HIV restriction in astrocytes is regulated by miRNA-181a. J Neuroinflammation. 2015 Apr 8;12:66. doi:
702 10.1186/s12974-015-0285-9.

703 Rentoft M, Lindell K, Tran P, Chabes AL, Buckland RJ, Watt DL, Marjavaara L, Nilsson AK, Melin B, Trygg
704 J, Johansson E, Chabes A. Heterozygous colon cancer-associated mutations of SAMHD1 have
705 functional significance. Proc Natl Acad Sci U S A. 2016 Apr 26;113(17):4723-8. doi:
706 10.1073/pnas.1519128113.

707 Rice GI, Bond J, Asipu A, Brunette RL, Manfield IW, Carr IM, Fuller JC, Jackson RM, Lamb T, Briggs TA,
708 Ali M, Gornall H, Couthard LR, Aeby A, Attard-Montalto SP, Bertini E, Bodemer C, Brockmann K,
709 Brueton LA, Corry PC, Desguerre I, Fazzi E, Cazorla AG, Gener B, Hamel BC, Heiberg A, Hunter M, van
710 der Knaap MS, Kumar R, Lagae L, Landrieu PG, Lourenco CM, Marom D, McDermott MF, van der
711 Merwe W, Orcesi S, Prendiville JS, Rasmussen M, Shalev SA, Soler DM, Shinawi M, Spiegel R, Tan TY,
712 Vanderver A, Wakeling EL, Wassmer E, Whittaker E, Lebon P, Stetson DB, Bonthron DT, Crow YJ.
713 Mutations involved in Aicardi-Goutieres syndrome implicate SAMHD1 as regulator of the innate
714 immune response. Nat Genet. 2009 Jul;41(7):829-32. doi: 10.1038/ng.373.

715 Riess M, Fuchs NV, Idica A, Hamdorf M, Flory E, Pedersen IM, König R. Interferons Induce Expression
716 of SAMHD1 in Monocytes through Down-regulation of miR-181a and miR-30a. J Biol Chem. 2017 Jan
717 6;292(1):264-277. doi: 10.1074/jbc.M116.752584.

718 Roider T, Wang X, Hüttl K, Müller-Tidow C, Klapper W, Rosenwald A, Stewart JP, de Castro DG, Dreger
719 P, Hermine O, Kluin-Nelemans HC, Grabe N, Dreyling M, Pott C, Ott G, Hoster E, Dietrich S. The impact
720 of SAMHD1 expression and mutation status in mantle cell lymphoma: An analysis of the MCL Younger
721 and Elderly trial. Int J Cancer. 2021 Jan 1;148(1):150-160. doi: 10.1002/ijc.33202.

722 Rothenburger T, McLaughlin KM, Herold T, Schneider C, Oellerich T, Rothweiler F, Feber A, Fenton TR,
723 Wass MN, Keppler OT, Michaelis M, Cinatl J Jr. SAMHD1 is a key regulator of the lineage-specific

response of acute lymphoblastic leukaemias to nelarabine. Commun Biol. 2020 Jun 24;3(1):324. doi: 10.1038/s42003-020-1052-8.

Schneider C, Oellerich T, Baldauf HM, Schwarz SM, Thomas D, Flick R, Bohnenberger H, Kaderali L, Stegmann L, Cremer A, Martin M, Lohmeyer J, Michaelis M, Hornung V, Schliemann C, Berdel WE, Hartmann W, Wardelmann E, Comoglio F, Hansmann ML, Yakunin AF, Geisslinger G, Ströbel P, Ferreirós N, Serve H, Keppler OT, Cinatl J Jr. SAMHD1 is a biomarker for cytarabine response and a therapeutic target in acute myeloid leukemia. Nat Med. 2017 Feb;23(2):250-255. doi: 10.1038/nm.4255.

Shang Z, Qian L, Liu S, Niu X, Qiao Z, Sun Y, Zhang Y, Fan LY, Guan X, Cao CX, Xiao H. Graphene Oxide-Facilitated Comprehensive Analysis of Cellular Nucleic Acid Binding Proteins for Lung Cancer. ACS Appl Mater Interfaces. 2018 May 30;10(21):17756-17770. doi: 10.1021/acsami.8b05428.

Shi Y, Lv G, Chu Z, Piao L, Liu X, Wang T, Jiang Y, Zhang P. Identification of natural splice variants of SAMHD1 in virus-infected HCC. Oncol Rep. 2014 Feb;31(2):687-92. doi: 10.3892/or.2013.2895.

Sim NL, Kumar P, Hu J, Henikoff S, Schneider G, Ng PC. SIFT web server: predicting effects of amino acid substitutions on proteins. Nucleic Acids Res. 2012 Jul;40(Web Server issue):W452-7. doi: 10.1093/nar/gks539.

Uhlen M, Zhang C, Lee S, Sjöstedt E, Fagerberg L, Bidkhori G, Benfeitas R, Arif M, Liu Z, Edfors F, Sanli K, von Feilitzen K, Oksvold P, Lundberg E, Hober S, Nilsson P, Mattsson J, Schwenk JM, Brunnström H, Glimelius B, Sjöblom T, Edqvist PH, Djureinovic D, Micke P, Lindskog C, Mardinoglu A, Ponten F. A pathology atlas of the human cancer transcriptome. Science. 2017 Aug 18;357(6352):eaan2507. doi: 10.1126/science.aan2507.

UniProt Consortium. UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Res. 2021 Jan 8;49(D1):D480-D489. doi: 10.1093/nar/gkaa1100.

Vaser R, Adusumalli S, Leng SN, Sikic M, Ng PC. SIFT missense predictions for genomes. Nat Protoc. 2016 Jan;11(1):1-9. doi: 10.1038/nprot.2015.123.

Wang X, Duanmu J, Fu X, Li T, Jiang Q. Analyzing and validating the prognostic value and mechanism of colon cancer immune microenvironment. J Transl Med. 2020 Aug 28;18(1):324. doi: 10.1186/s12967-020-02491-w.

Wass MN, Ray L, Michaelis M. Understanding of researcher behavior is required to improve data reliability. Gigascience. 2019 May 1;8(5):giz017. doi: 10.1093/gigascience/giz017.

Xagoraris I, Vassilakopoulos TP, Drakos E, Angelopoulou MK, Panitsas F, Herold N, Medeiros LJ, Giakoumis X, Pangalis GA, Rassidakis GZ. Expression of the novel tumour suppressor sterile alpha motif and HD domain-containing protein 1 is an independent adverse prognostic factor in classical Hodgkin lymphoma. Br J Haematol. 2021 May;193(3):488-496. doi: 10.1111/bjh.17352.

Yanai I, Lercher M. Novel predictions arise from contradictions. Genome Biol. 2021 May 11;22(1):153. doi: 10.1186/s13059-021-02371-6.

Yang CA, Huang HY, Chang YS, Lin CL, Lai IL, Chang JG. DNA-Sensing and Nuclease Gene Expressions as Markers for Colorectal Cancer Progression. Oncology. 2017;92(2):115-124. doi: 10.1159/000452281.

Yang Y, Zhang J, Chen Y, Xu R, Zhao Q, Guo W. MUC4, MUC16, and TTN genes mutation correlated with prognosis, and predicted tumor mutation burden and immunotherapy efficacy in gastric cancer and pan-cancer. Clin Transl Med. 2020 Aug;10(4):e155. doi: 10.1002/ctm2.155.

765 Zhu KW, Chen P, Zhang DY, Yan H, Liu H, Cen LN, Liu YL, Cao S, Zhou G, Zeng H, Chen SP, Zhao XL, Chen
766 XP. Association of genetic polymorphisms in genes involved in Ara-C and dNTP metabolism pathway
767 with chemosensitivity and prognosis of adult acute myeloid leukemia (AML). J Transl Med. 2018 Apr
768 10;16(1):90. doi: 10.1186/s12967-018-1463-1.

769

770