


RESEARCH

Open Access



# Ganciclovir-induced mutations are present in a diverse spectrum of post-transplant malignancies

Hu Fang<sup>1,2†</sup>, Helen H. N. Yan<sup>3,4†</sup>, Rebecca A. Bilardi<sup>5,6†</sup>, Christoffer Flensburg<sup>5,6</sup>, Haocheng Yang<sup>1</sup>, Jayne A. Barbour<sup>1</sup>, Hoi Cheong Siu<sup>3</sup>, Michelle Turski<sup>7</sup>, Edward Chew<sup>5</sup>, Zhen Xu<sup>5,6</sup>, Siu T. Lam<sup>8</sup>, Rakesh Sharma<sup>8</sup>, Mengya Xu<sup>1</sup>, Junshi Li<sup>1</sup>, Ho W. Ip<sup>9</sup>, Carol Y. M. Cheung<sup>10</sup>, Michael S. Y. Huen<sup>1</sup>, E. Alejandro Sweet-Cordero<sup>7</sup>, Ian J. Majewski<sup>5,6\*</sup>, Suet Y. Leung<sup>3,4,8,11\*</sup> and Jason W. H. Wong<sup>1,4,8,11\*</sup> 

## Abstract

**Background:** Ganciclovir (GCV) is widely used in solid organ and haematopoietic stem cell transplant patients for prophylaxis and treatment of cytomegalovirus. It has long been considered a mutagen and carcinogen. However, the contribution of GCV to cancer incidence and other factors that influence its mutagenicity remains unknown.

**Methods:** This retrospective cohort study analysed genomics data for 121,771 patients who had undergone targeted sequencing compiled by the Genomics Evidence Neoplasia Information Exchange (GENIE) or Foundation Medicine (FM). A statistical approach was developed to identify patients with GCV-associated mutational signature (GCV<sup>sig</sup>) from targeted sequenced data of tumour samples. Cell line exposure models were further used to quantify mutation burden and DNA damage caused by GCV and other antiviral and immunosuppressive drugs.

**Results:** Mutational profiles from 22 of 121,771 patient samples in the GENIE and FM cohorts showed evidence of GCV<sup>sig</sup>. A diverse range of cancers was represented. All patients with detailed clinical history available had previously undergone solid organ transplantation and received GCV and mycophenolate treatment. RAS hotspot mutations associated with GCV<sup>sig</sup> were present in 9 of the 22 samples, with all samples harbouring multiple GCV-associated protein-altering mutations in cancer driver genes. In vitro testing in cell lines showed that elevated DNA damage response and GCV<sup>sig</sup> are uniquely associated with GCV but not acyclovir, a structurally similar antiviral. Combination treatment of GCV with the immunosuppressant, mycophenolate mofetil (MMF), increased the misincorporation of GCV in genomic DNA and mutations attributed to GCV<sup>sig</sup> in cell lines and organoids.

**Conclusions:** In summary, GCV can cause a diverse range of cancers. Its mutagenicity may be potentiated by other therapies, such as mycophenolate, commonly co-prescribed with GCV for post-transplant patients. Further

<sup>†</sup>Hu Fang, Helen H. N. Yan and Rebecca A. Bilardi are co-first authors.

\*Correspondence: majewski@wehi.edu.au; suetyi@hku.hk; jwhwong@hku.hk

<sup>1</sup> School of Biomedical Sciences, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China

<sup>3</sup> Department of Pathology, School of Clinical Medicine, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China

<sup>5</sup> The Walter and Eliza Hall Institute of Medical Research, 1G Royal Parade, Parkville, VIC 3052, Australia

Full list of author information is available at the end of the article



investigation of the optimal use of these drugs could help reduce GCV-associated mutagenesis in post-transplant patients.

**Keywords:** Ganciclovir, Mutagenesis, Mutational signature, DNA damage, Organ transplant

## Background

Recipients of solid organ or allogeneic haematopoietic stem cell transplants (HSCT) have a higher risk of developing cancer. It has been assumed that long-term immunosuppression and viral infection account for the elevated risk of post-transplant malignancies [1, 2]. Indeed, immunosuppressive drugs such as azathioprine are known to damage DNA and are associated with an increased risk of skin cancer [3]. Furthermore, transplant recipients experience higher rates of infection with oncogenic viruses, such as Epstein-Barr virus and herpesvirus 8, which are major risk factors for post-transplant lymphoproliferative disorders [4] and Kaposi sarcoma [5], respectively. Recently, it was reported that the antiviral ganciclovir (GCV) induces a distinctive mutational signature dominated by CA>AA substitutions [6]. This mutational signature was found in blood progenitor cells of patients after HSCT and in two patients from a survey of 3668 solid whole cancer genomes [6]. Furthermore, we had previously observed two individuals (H015 and WEHI-2) with a shared history of acute myeloid leukaemia, HSCT, and colorectal cancer (CRC) carrying the same CA>AA enriched mutational signature [7, 8]. Both patients had received GCV as part of their treatment following HSCT (Additional file 1: Table S1).

GCV and the prodrug valganciclovir are widely used in solid organ transplant and HSCT patients for prophylaxis and treatment of cytomegalovirus (CMV). This study sought to establish the broader prevalence of GCV-associated cancers by interrogating somatic mutations from 121,771 target panel sequenced cancer patients. Furthermore, since transplant patients are often simultaneously administered a range of other drugs, we used cell line and organoid exposure models to investigate treatments that may influence the penetrance and variability in mutational burden observed in GCV-associated malignancies.

## Methods

### Target panel sequencing cancer patient cohort

A set of 121,771 cancer patient samples with somatic mutation and limited clinical information were obtained from AACR Project Genomics Evidence Neoplasia Information Exchange (GENIE) release 11-public [9] ( $n=104,264$ ) and FoundationMedicine (FM) [10] ( $n=17,507$ ). AACR Project GENIE is a publicly accessible database of real-world cancer genomics panel sequencing data assembled from cancer centres worldwide and

includes the most common cancer types. The GENIE cohort consisted of a mixture of data from 93 different panels ranging from 6 to 1422 genes (average 544), spanning 6 bases to 9.95 megabases (average 1.06 megabases), with the number of mutations ranging from 0 to 9364 (average 7.938). The FM cohort consists of adult solid tumour samples that underwent genomic profiling on a single uniform platform with 287 genes, spanning 0.83 megabases as part of standard clinical care. The number of mutations in the cohort ranged from 1 to 366 (average 6.787) (see Additional file 1: Table S2 for detailed cohort characteristics). For patients with multiple samples in the database, only the most recent sample, indicated by participant age, was included. This combined dataset is used to identify the presence of GCV-associated signature across different cancer types.

### Accessibility analysis for COSMIC census genes

Somatic mutation data was obtained for a set of recurrently mutated genes from the COSMIC census v77 database [11]. Genes with candidate hotspot mutations were identified by selecting those genes with high variance in mutation counts across the gene. We restricted our analysis to the 50 genes with the highest variance and focused on mutations with at least 100 occurrences in COSMIC. TP53 was removed because the volume of hotspot sites dominated the analysis. We used the mutational signature for GCV (GCV<sup>sig</sup>, derived from H015 [8]) to assess the accessibility of hotspot mutations based on their trimer context. We also assessed accessibility at the level of each driver gene by summing the values across all hotspots in that gene.

### Cell line and organoid treatment and whole-genome sequencing

CRC cell lines H414 and HCT116, murine myeloid cells (*HoxA9-Meis1*-transformed primary cells), and a normal human colon organoid line were used as models for GCV treatment. Clonal H414 and HCT116 cells were seeded in a human plasma-like medium (Gibco, A4899101), for whole-genome sequencing experiments, or Dulbecco's modified Eagles' medium, for all other experiments, supplemented with 10% foetal bovine serum 24 h prior to drug treatment. To test cell viability, these cells were treated with GCV (Abcam) or ACV (Abcam) ranging from 0.1 to 1000  $\mu$ M. After 48-h incubation, the cell viability was assessed by trypan blue exclusion assay. For

mutation analysis, cells were treated with 100  $\mu\text{M}$  GCV, 100  $\mu\text{M}$  ACV, and/or 1  $\mu\text{M}$  MMF (Roche) for 48 h. A colony (2–4 cells) was then isolated and expanded, as single cell clones were found to be not viable after drug treatment. Following expansion, DNA was extracted for WGS with 150-bp PE sequencing at 30 $\times$  using an Illumina Novaseq sequencer.

Myeloid cell lines were generated by infecting E13.5 C57/BL6 liver cells with pMSCV-HoxA9-IRES-Meis1 (a gift from Guy Sauvageau). These cells were infected with pMIG (GFP+) and single cell cloned by plating in Dulbecco's modified Eagles' medium with 20% FBS, 0.3% BactoAgar, and 10 ng/mL mIL-3 (WEHI). Individual colonies were expanded and maintained in IMDM with 10% FBS and 10 ng/mL mIL-3. For growth inhibition experiments, myeloid cell lines were seeded and treated with ACV (Hospira AU) or GCV (Pharmaco AU) ranging from 0.00316 to 100  $\mu\text{M}$ . To assess mutagenicity, clonal cultures were treated with GCV and ACV continuously for 13–15 days, then single cell cloned and expanded in liquid culture. Control samples were either untreated or cultured in vehicle (DMSO). DNA was extracted for WGS with 150-bp PE sequencing at 250–600 M reads per sample using an Illumina Novaseq sequencer.

The normal colon organoid was previously generated from pooled sigmoid organoid derived from a 40-year-old female CRC patient [12]. Briefly, bulk sigmoid organoid was trypsinized into single cells, passed through a cell strainer, and then loaded to a BD Influx<sup>TM</sup> cell sorter (Biosciences). Single cells were selected based on their physical size and molecular granularity. The sorted cells were serially diluted and seeded in Matrigel for expansion. The clonal organoids were manually picked at 1–2 weeks post-sorting. Fresh medium (advanced DMEM/F12, 1 $\times$  GlutaMax, 1 $\times$  HEPES, 1 $\times$  P/S, 50% Wnt3a, 10% RSPO-1, 10% Noggin, 1 $\times$  B27, 50 ng/mL EGF, 200 ng/mL FGF10, 1 mM N-acetylcysteine, 1 nM gastrin, 2  $\mu\text{M}$  A83-01) was changed every 2–3 days and organoids were passaged every week. Organoids were treated with 1  $\mu\text{M}$  MMF (Roche), 20  $\mu\text{M}$  GCV (Hainan Poly Pharm Co Ltd), or in combinations (1  $\mu\text{M}$  MMF + 20  $\mu\text{M}$  GCV or 1  $\mu\text{M}$  MMF + 40  $\mu\text{M}$  GCV) continuously for 4–6 weeks. The fresh medium was changed every 2–3 days and organoids were passaged every week. DNA was extracted from the bulk culture as the organoids were found to be not viable as single clones after drug treatment. WGS was performed with 150-bp PE sequencing at 30 $\times$  using an Illumina Novaseq sequencer.

The concentration for each cell line treatment is provided in Additional file 1: Table S3. A detailed description of the methods associated with cell viability, DNA damage response, and mass spectrometry assays are provided in Additional file 2: Supplementary Methods.

### Mutation calling from whole-genome sequencing data

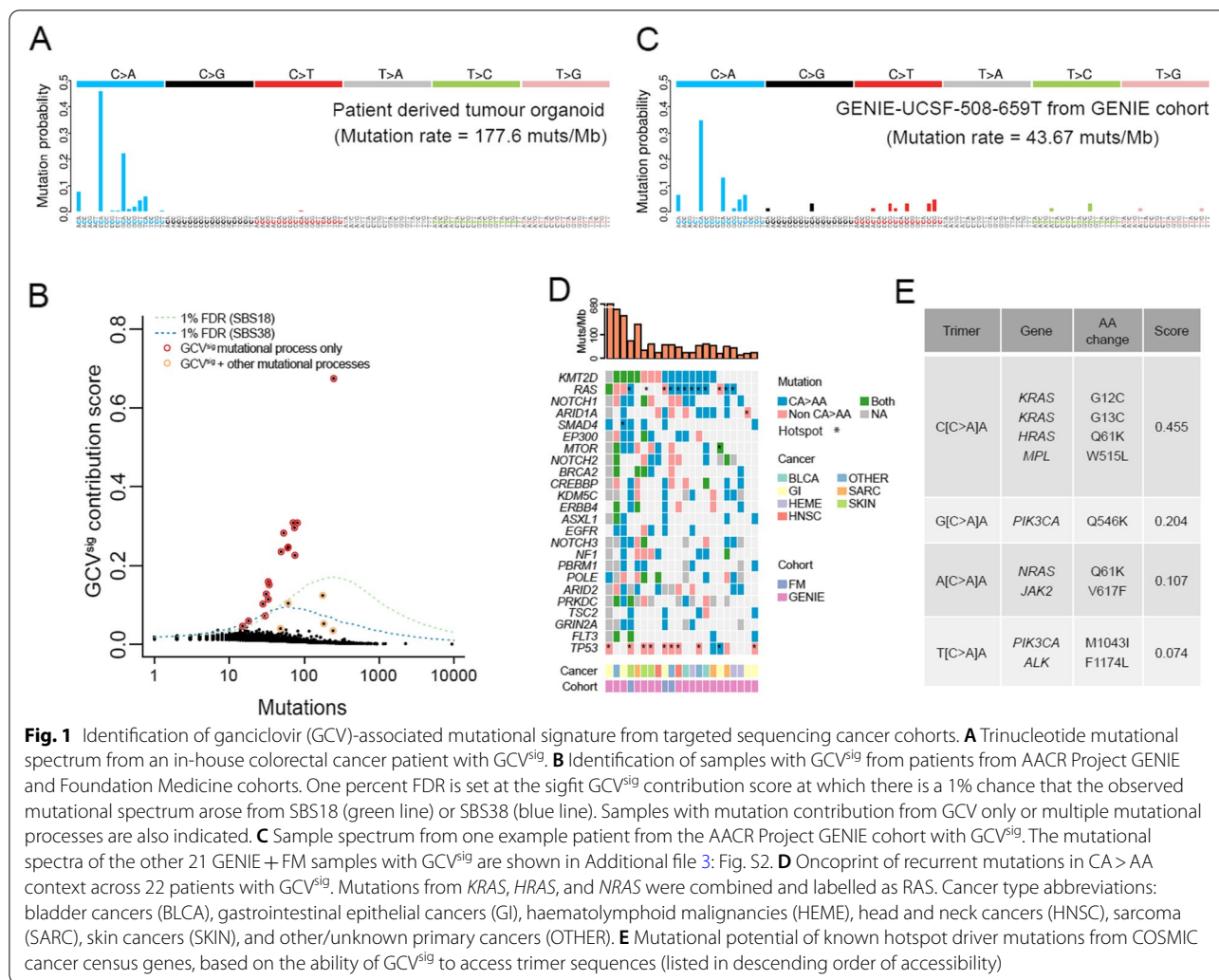
For H414 and HCT116 and normal human organoids, raw sequencing reads were aligned to the human (hg38) reference genome using bwa (v0.7.17-r1188) [13]. Mutations were called using MuTect2 (v4.2.5.0) [14] and Strelka2 (v2.9.7) [15] were run in paired normal-tumour mode, where the respective parental clone of each cell line was used as the “normal”, while the drug-treated or control cells were used as “tumour”. For MuTect2, mutation calls required support from 3 or more reads with at least one on each strand. The final set of mutations used for the downstream analysis required that a mutation was annotated as PASS by both MuTect2 and Strelka2 and that the mutation was not shared with any other sample.

For data from the murine cell line, reads were aligned to mm10 with bwa (v0.7.17-r1188) [13], and samples from the same clone were grouped for analysis with superFreq (v1.4.3) [16]. For signature analysis, we used somatic variants with somaticP > 0.5, read depth  $\geq 15$ , VAF  $\geq 0.25$ , and required VAF  $\leq 0.05$  in all other samples. Bases with at least 15 read depth were classified as callable. The median callable region for the clones was 2.0 Gbp (range 1.7 to 2.3 Gbp). Regions classified as Simple\_repeat or Low\_complexity by repeatMasker were excluded from the analysis.

### Statistical analysis

To identify patients with evidence of GCV-associated mutagenesis, single nucleotide substitutions from each sample were represented based on trinucleotide substitution frequency [17]. The GCV-associated signature (GCV<sup>sig</sup>) was obtained from patient H015 [8] (Fig. 1A). The R package, sigfit [18], was used to identify samples that exhibited GCV<sup>sig</sup>. Briefly, sigfit applies Bayesian inference to fit known mutational signatures to an observed mutational spectrum. Sigfit was used to compute the contribution of mutational signatures (referred to as contribution score) from COSMICv3 with the addition of GCV<sup>sig</sup> for each patient.

To estimate the false discovery rate (FDR), we simulated trinucleotide mutational spectra based on COSMIC mutational signature SBS18, which shared the highest similarity with GCV<sup>sig</sup> among common COSMIC mutational signatures (Additional file 1: Table S4). A minor contribution from the age-associated SBS5 was also included in the simulation as a background process. To this end, mutation spectra consisting of 0.95 SBS18 and 0.05 SBS5 were simulated using the R `r multinom` function. Sigfit was used to provide an estimate of GCV<sup>sig</sup> contribution for each simulated mutation spectrum. The process was repeated 1000 times for each mutation count ranging from 1 to 10,000. The 99th percentile GCV<sup>sig</sup> contribution score from the simulated SBS18 spectra



was set as the 1% FDR cutoff for identifying samples with GCV-associated mutagenesis. A GCV<sup>sig</sup> cutoff was also calculated for SBS38 spectra, but this was only applied to skin malignancies as it has been specifically associated with UV damage [19].

As sigfit assigns uniformly low contribution scores across all signatures when it cannot confidently identify any distinctive mutational processes, samples with less than 10 mutations were excluded as our analysis lacks the ability to confidently detect GCV<sup>sig</sup> below this mutation count based on our 1% FDR threshold (Additional file 3: Fig. S1A).

While this analysis identified cancers where GCV<sup>sig</sup> was dominant, we noted that the FDR cutoff based on 0.95 SBS18 contribution would be too stringent for cancers with large contributions from multiple mutational processes. To this end, we further simulated SBS18 and SBS38 spectra at a range of SBS contributions (0.05–0.95, the remainder contribution being SBS5) across a range of mutation counts (5–250) to obtain the 99th percentile

GCV<sup>sig</sup> contribution score from the simulated SBS18 or SBS38 spectra (Additional file 3: Fig. S1B–C). Using these values, we then used the following regression model to estimate the 1% FDR GCV<sup>sig</sup> score cutoff.

$$GCV^{sig}_{score} = \beta_0 + \beta_1 X_{count} + \beta_2 \log(X_{count}) + \beta_3 X_{contrib} + \beta_4 X_{contrib} \log(X_{count})$$

where  $X_{contrib}$  indicates the SBS contribution fraction and  $X_{count}$  is the mutation burden. The model provides a very good estimation of the actual GCV<sup>sig</sup> score ( $R^2=0.98$  and  $0.85$  for SBS18 and SBS38 respectively, Additional file 3: Fig. S1D–E) and enables the 99th percentile GCV<sup>sig</sup> cutoff score to be estimated without a computationally costly simulation for each sample. For individual samples with sufficient CA > AA mutations (>10), the GCV<sup>sig</sup> contribution is estimated based on the fraction of CA > AA mutations and, along with the total mutation count, is input to the regression model for

SBS18 and SBS38 to determine if the observed  $GCV^{sig}$  score is above the estimated 99th percentile  $GCV^{sig}$  score cutoff. These cutoff values are listed for each sample in Additional file 1: Table S5. To test if the  $GCV^{sig}$  of the sample passed this cutoff, the SBS contribution is estimated using the fraction of CA>AA mutations in the sample.

To ensure that our FDR cutoff is robust, particularly for samples with low number of mutations, we further estimated the FDR of  $GCV^{sig}$ -positive samples for the cohort. To do this, we randomly sampled mutational spectra from the cohort-wide average and calculated the  $GCV^{sig}$  contribution score and whether it is above 1% SBS18 or SBS38 FDR as described above. This whole process was repeated 10 times.

Finally, to further account for multiple testing correction, we calculated  $p$ -values for each sample by comparing the difference between the expected and the observed frequency of mutations attributed to  $GCV^{sig}$ . This frequency and the frequency of non- $GCV^{sig}$  mutations are calculated respectively as,

$$X_{GCV^{sig}} = GCV^{sig}_{score} X_{count} \text{ and } X_{non-GCV^{sig}} = (1 - GCV^{sig}_{score}) X_{count}$$

Under the null hypothesis, the expected frequency of  $GCV^{sig}$  is based on the cohort-wide average mutational spectrum simulated 1000 times. The chi-square test is then used to compute the  $p$ -value, which is then multiple testing corrected by the Benjamini–Hochberg procedure.

For analysis of significance relating to cell line and organoid models, two-sided Student's test, ratio  $t$ -test, or Fisher's exact test was used as appropriate.

## Results

### Ganciclovir (GCV)-associated mutagenesis is detected across a broad range of cancers

To identify the presence of  $GCV^{sig}$  across different cancer types, we analysed genomics data for 121,771 patients who had undergone targeted sequencing compiled by the Genomics Evidence Neoplasia Information Exchange (GENIE) or Foundation Medicine (FM). Despite the relatively small region of the genome covered by targeted sequencing platforms, we were confident that the  $GCV^{sig}$  could be detected as the signature is highly distinctive, and we had previously found it associated with hypermutation (>100 mut/Mb) [8] (Fig. 1A). We identified 22 (0.0181% of 121,771) patients with strong evidence of  $GCV^{sig}$  (FDR < 1%, see the “Methods” section) (Fig. 1B, C and Additional file 3: Figs. S2 and S3, Additional file 1: Table S5). After multiple testing correction (see the “Methods” section, Additional file 1: Table S5), 17/22 remained significant (adjusted  $p$ -val < 0.05), generally with the samples with

lower mutations becoming insignificant. Previously, C>A mutations from  $GCV^{sig}$  had been described to show replication strand asymmetry with more C>A mutations on the leading strand (or G>T mutations on the lagging strand) [6]. Despite the very low mutation count, 15/22 samples showed more C>A mutations on the leading strand and this bias was also evident in aggregate (Additional file 3: Fig. S4), consistent with the mutations being linked to GCV exposure. Of the 5 samples that were below significance after multiple testing correction, 3 showed leading strand bias; thus, we included all 22 samples for the subsequent analysis.

We were able to obtain detailed clinical information for four samples (of the 22), and all had a history of solid organ transplantation and treatment with valganciclovir, the prodrug of GCV (Additional file 1: Table S1). Of note, the 22 patients spanned a diverse range of malignancies (Fig. 1D), with the highest occurrence in epithelial cancers of diverse organs ( $n=16$ ), such as the gastrointestinal tract ( $n=5$ ), skin ( $n=3$ ), head and neck ( $n=2$ ), and bladder ( $n=2$ ); sarcoma ( $n=3$ ); and haematolymphoid malignancy ( $n=3$ ), which is consistent with a range of cancers where  $GCV^{sig}$  had been recently observed [6–8, 20]. Even though our analysis was only powered to detect  $GCV^{sig}$  in samples with 10 or more mutations, those with  $GCV^{sig}$  detected all had high mutation burden, typically among the top 1–5% of respective cancer types (Additional file 3: Fig. S5).

To evaluate the potential of GCV-induced damage directly leading to driver mutations, we calculated the ability of the signature to access known drivers from COSMIC census genes. The gene most accessible by  $GCV^{sig}$  in our analysis was *KRAS*, which has two hotspot variants in the CCA context. GCV can access multiple hotspot sites in the RAS family, including *KRAS* (G12C, G13C), *HRAS* (Q61K), and *NRAS* (Q61K). Other important hotspots include mutations linked to myeloproliferative disorders, including *JAK2* (V617F) and *MPL* (W515L) (Fig. 1E). This is consistent with the observation that 9 of the 22 patients from the GENIE and FM cohorts have a mutation in one of the *RAS* genes in the CA>AA context (Fig. 1D) and also in line with previous observations in  $GCV^{sig}$ -associated cancers [6]. Furthermore, all the samples have at least one potential CA>AA protein-altering driver mutation. This suggests that GCV can play a direct role in carcinogenesis.

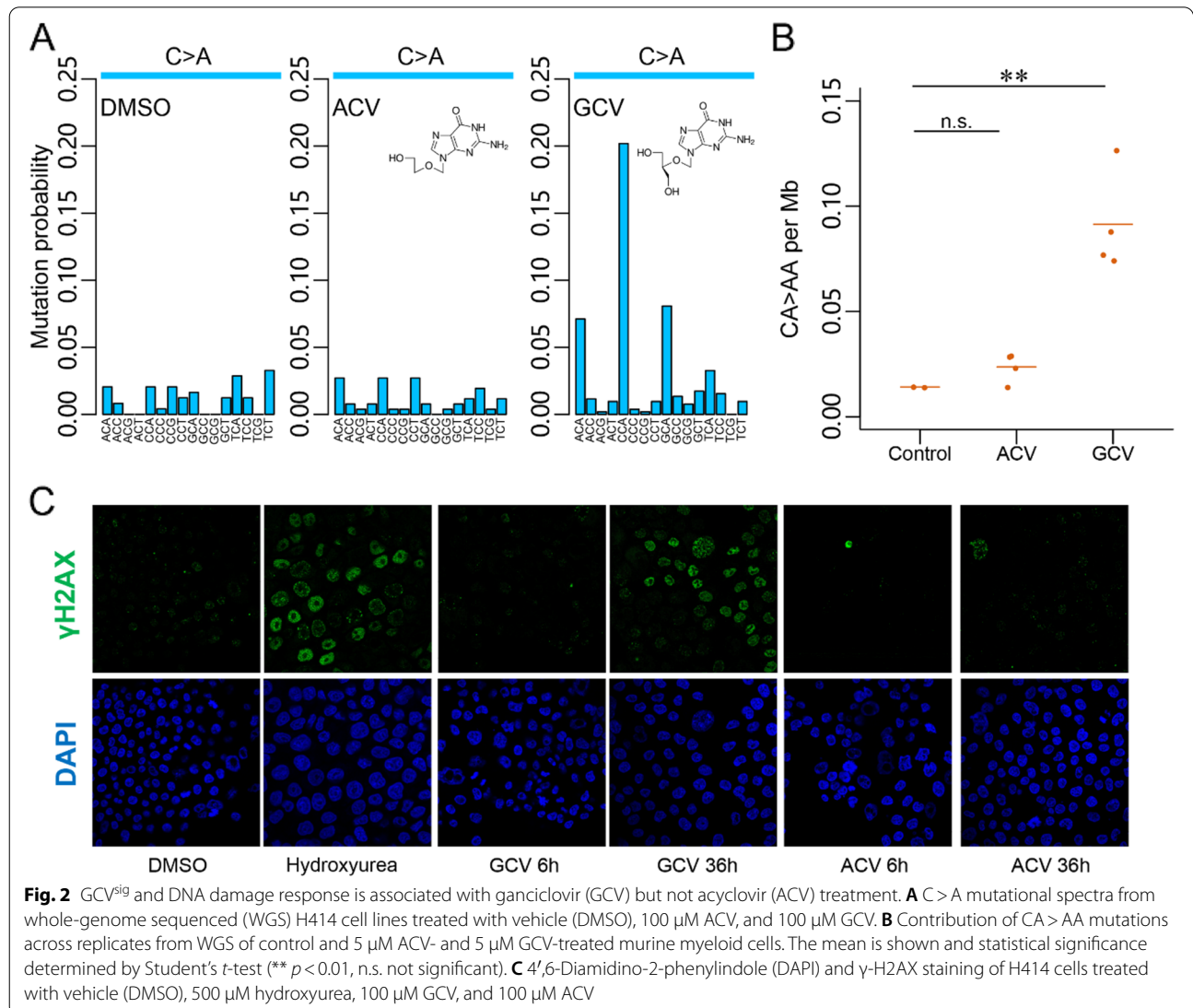
### GCV but not acyclovir (ACV) induces DNA damage response and $GCV^{sig}$ in cell line models

We reviewed the clinical and drug history of our two patients with HSCT and colorectal neoplasms carrying  $GCV^{sig}$  (H015 and WEHI-2) and noted that aside from GCV, they had also received acyclovir (ACV), which

could also potentially perturb DNA synthesis [21]. To investigate whether GCV<sup>sig</sup> is uniquely associated with GCV treatment, we performed WGS of cloned CRC cell line, H414, treated with DMSO, GCV, or ACV. The cell line was generally tolerant of drug treatment (Additional file 3: Fig. S6A) and 100 μM GCV and ACV were selected for the WGS experiments. GCV treatment resulted in the distinctive CA>AA mutational signature, whereas samples treated with ACV showed no such signature (Fig. 2A, Additional file 3: Fig. S6B, Additional file 1: Table S6). To validate the specificity of GCV-associated mutagenesis, treatment and WGS were repeated in murine myeloid cell lines (*HoxA9-Meis1*-transformed primary cells), which were less tolerant of the antivirals than H414 (Additional file 3: Fig. S6C). Despite the substantially lower dose

(5 μM), only GCV-treated cells acquired CA>AA mutations compared with ACV and controls ( $p < 0.01$ , Student's *t*-test, Fig. 2B, Additional file 1: Table S7).

GCV-triphosphate (GCV-TP) has been reported to be misincorporated into genomic DNA in place of guanosine triphosphate [22]. Consistently, staining of DNA damage response marker γH2AX in H414 showed that accumulation of damage was most evident only after 36 h of GCV treatment when most cells have replicated, while γH2AX signal was largely absent in ACV-treated cells (Fig. 2C). Flow cytometry analysis of the treated cells did not show strong evidence of cell cycle arrest after 36 h of GCV or ACV treatment, supporting the view that H414 cells are generally tolerant of treatment (Additional file 3: Fig. S6D).



### Mycophenolate mofetil treatment potentiates GCV-induced mutagenesis in cells

While GCV is widely used in post-transplant patients, it is unclear what other factors, such as dosage, drug interactions, genetics, or environmental factors, contribute to its penetrance and mutational burden. MMF is an immunosuppressant that disrupts guanosine synthesis and is frequently used in transplant patients [23], including those found with GCV<sup>sig</sup> from this study and our two patients (Additional file 1: Table S1). To test if MMF might influence GCV-induced mutagenesis, we performed combination drug treatments followed by WGS in H414 cells. The fraction of mutations attributed to CA>AA was higher with combination treatment than GCV alone ( $p < 0.0001$ , Fisher's exact test, 0.728 versus 0.387, Fig. 3A, Additional file 3: Fig. S6B, Additional file 1: Table S6). MMF alone did not induce any proportional increase in CA>AA mutations when compared with the DMSO control ( $p = 0.878$ , Fisher's exact test, 0.091 versus 0.086, Fig. 3A, Additional file 3: Fig. S6B, Additional file 1: Table S6). The fraction of CA>AA mutations was also higher in organoids treated with MMF and GCV, compared with GCV alone ( $p < 0.0001$ , Fisher's exact test, 0.269 versus 0.129, Fig. 3B, Additional file 3: Fig. S7A, Additional file 1: Table S6). In this system, increasing GCV dosage from 20 to 40  $\mu\text{M}$  did not further increase the fraction of CA>AA mutations ( $p = 0.7721$ , Fisher's exact test, 0.269 versus 0.265, Fig. 3B, Additional file 3: Fig. S7A, Additional file 1: Table S6).

We further assessed the impact of GCV-induced mutations with and without MMF in cells with DNA mismatch repair deficiency (MMRd). Despite the very high background mutation rate arising from MMRd in HCT116 cells, an increase in the portion of CA>AA mutations could be observed with GCV treatment compared with DMSO control ( $p < 0.0001$ , Fisher's exact test, 0.062 versus 0.035, Fig. 3C, Additional file 3: Fig. S7B, Additional file 1: Table S6). This was further increased when used in combination with MMF ( $p = 0.03$ , Fisher's exact test, 0.074 versus 0.062, Fig. 3C, Additional file 3: Fig. S7B, Additional file 1: Table S6). Using sigfit, the GCV<sup>sig</sup> contribution score significantly increased between GCV alone and GCV with MMF across the various treatments in cell lines and organoids ( $p = 0.0134$ , ratio  $t$ -test, Fig. 3D).

Although MMF treatment strongly increased the doubling time of H414 cells ( $\alpha = 1.51$ ,  $R^2 = 0.8869$ ,  $p < 0.001$ , Fig. 3E, Additional file 3: Fig. S7C), prolonged exposure of cells to GCV prior to cell replication is unlikely to increase mutagenesis as GCV-associated DNA damage occurs during DNA replication [22]. To test whether MMF may have a synergistic effect on GCV-TP misincorporation, we used a mass spectrometry-based assay

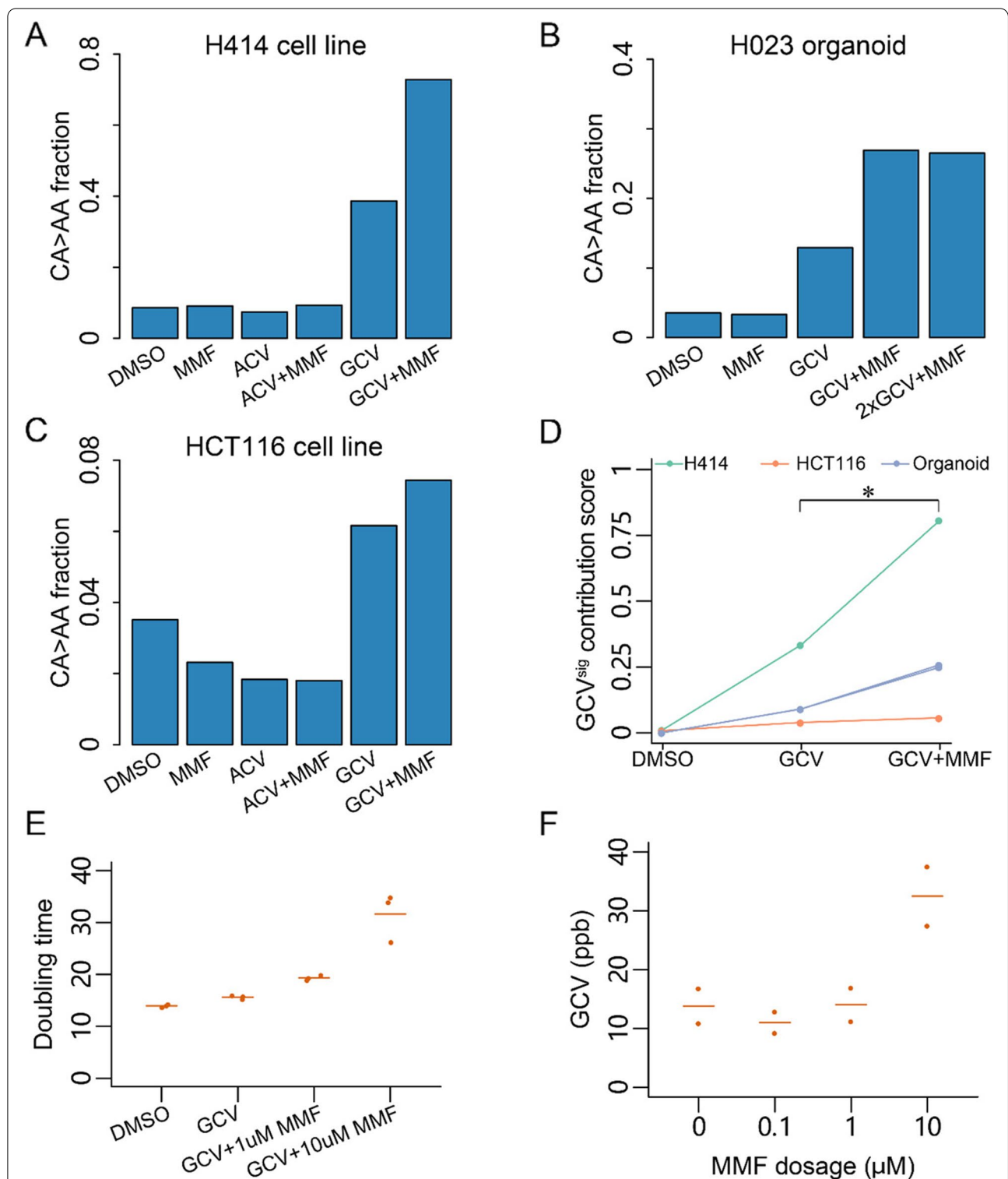
to quantify GCV in genomic DNA [24] (Additional file 3: Fig. S7D). The level of GCV in DNA was found to increase from 13.8 ppb without MMF to 32.5 ppb at 10  $\mu\text{M}$  MMF (Fig. 3F, Additional file 3: Fig. S7E). This suggests that MMF likely increases the mutagenic potential of GCV by increasing its rate of misincorporation into genomic DNA.

### Discussion

Post-transplant malignancy has typically been thought of as a natural consequence of sustained immunosuppression, but our results point to an active role for the antiviral GCV in causing mutations that may lead to cancer. Cancers with high levels of GCV<sup>sig</sup> were rare in this group of unselected cancer patients, occurring at approximately 1 in 5500 individuals. This may underestimate the impact of GCV, because of the high level of signal required for detection using a targeted panel. GCV<sup>sig</sup> is also likely to be far more common in transplant recipients, who regularly receive antiviral therapies. We show that these cancers harbour driver mutations that are accessible by GCV-associated mutagenesis, implicating DNA damage from GCV as a causative factor in initiating the disease. Furthermore, we found that MMF may potentiate the mutagenic effect of GCV in cell line and organoid models.

The finding of GCV-associated mutagenesis across diverse cancer types spanning epithelial, mesenchymal, and haematolymphoid lineages suggests that GCV could exert a broad mutagenic effect on most cells in the body during the course of GCV treatment. In support of this, we previously observed GCV<sup>sig</sup> not only in the tumour tissue and tumour-derived organoid, but it was also found in an organoid generated from the normal colon of patient H015, albeit at lower levels [8]. Furthermore, GCV<sup>sig</sup> was observed in normal blood stem and progenitor cells after bone marrow transplantation [6]. As GCV is widely prescribed in post-transplant patients, both as prophylaxis and treatment for CMV, most post-transplant patients are likely to have some level of exposure to the drug. Future studies of somatic mutations in normal cells and tissues from GCV-treated individuals will help establish the prevalence and cell types most vulnerable to GCV-induced mutagenesis. This may ultimately provide important clues in the potential monitoring of genotoxicity of GCV in patients undergoing treatment.

GCV is a guanine analogue that is phosphorylated and misincorporated into genomic DNA [22]. Given that the rate of GCV misincorporation is likely to be influenced by the composition of the nucleotide pool, we hypothesised that MMF, an inhibitor of inosine monophosphate dehydrogenase, might potentiate the mutagenic effect



**Fig. 3** Mycophenolate mofetil (MMF) treatment enhances ganciclovir (GCV)-associated mutagenesis. Fraction of CA > AA mutations in the H414 cell line (A), H023 organoid (B), and the HCT116 cell line (C) treated with vehicle (DMSO), MMF, ACV, ACV + MMF, GCV, or GCV + MMF. The concentration of MMF is 1 μM in all treatments. For H414 and HCT116 cells, the concentration of ACV and GCV is 100 μM. For H023, the concentration of GCV is 20 μM and 40 μM (2 ×). D GCV<sup>sig</sup> contribution score computed using sigfit for the mutational spectra of the treated cell lines and organoids in A–C. E Doubling time of H414 cells under different treatment conditions. GCV concentration was constant at 100 μM. The line indicates the mean of replicates. F Liquid chromatograph-tandem mass spectrometry-based quantification of GCV incorporation into genomic DNA under increasing MMF dosage in 100 μM GCV-treated H414 cells. The line indicates the mean of replicates



of GCV. Indeed, using cell line and organoid models, MMF was found to increase the incorporation of GCV in genomic DNA and the proportion of CA>AA mutations. The finding has pronounced clinical relevance since MMF is commonly used long-term in solid organ transplant patients as part of a steroid-sparing regime and for treating graft versus host disease in HSCT recipients. Since co-existing CMV infection is common in these patients, the risk of inducing GCV mutations could vary depending on the dosage and duration of drug usage. If the interaction between GCV and MMF is verified *in vivo*, then the combined long-term use of the two drugs should be cautioned. We note that patient H015, who displayed an extreme level of GCV<sup>sig</sup> in her CRC and normal colon [8], had concurrently received both drugs for over 5 years prior to colon cancer onset. We did not see elevated mutation rates in response to treatment with ACV, but ACV is much less effective against CMV compared to GCV. Further well-controlled population studies or clinical trials will be needed to document the cancer risk of GCV (and valganciclovir) alone or in combination with mycophenate, and whether other anti-CMV agents may be used instead to limit GCV exposure. For example, foscarnet, which has been shown to be non-mutagenic in cell culture [6], has been shown to have comparable efficacy and toxicity to GCV both as preemptive and first-line CMV treatment [25, 26]. However, foscarnet is administered intravenously, which will limit its use in some instances.

Another interesting observation from our study is that GCV<sup>sig</sup> is associated with a broad range of transplantation types. Cancers with GCV<sup>sig</sup> were previously reported in patients who had HSCT or kidney transplants [6]. In the four target panel sequenced patients where we have detailed clinical information, three had received a kidney or kidney and pancreas transplant, and the other had a lung transplant. Meanwhile, our two patients both received HSCT [7, 8]. This suggests that GCV treatment induces mutations across a range of transplant protocols. It would be beneficial to look further at how this relates to CMV status and clinical history to evaluate the level of GCV exposure. CMV infection and the prescription of GCV are generally common in immunocompromised individuals. It would be important to establish whether GCV-associated cancers are also present in other immunocompromised patients, such as those with acquired immunodeficiency syndrome (AIDS) due to infection of human immunodeficiency virus (HIV). As MMF is not part of the standard of care in HIV/AIDS, such a study could help further define interactions between treatments and the conditions under which GCV-associated cancers are more likely to develop.

In our cell line models, we sought to compare GCV<sup>sig</sup> in MMR-deficient (HCT116) and MMR-proficient (H414) cell lines. We observed that the relative contribution of GCV<sup>sig</sup> was lower in HCT116 due to increased background C>T and T>C mutations resulting from replication error (Fig. 3A, C). Comparing absolute counts in our data (Additional file 3: Fig. S8) may not reflect differences in the rate of GCV-induced mutation formation as the cancer cells have some level of polyploidy and showed variable growth characteristics, cloning efficiency, and drug response (Additional file 3: Fig. S6A). To control for polyploidy, we evaluated the variant allele frequency of diploid regions (Additional file 3: Figs. S9 and 10) and found that there is still variation of clonality across the samples. Further experiments will need to be conducted in a wider range of MMR-proficient and MMR-deficient cell lines along with confirmation using an *in vitro* mismatch repair and excision assay to fully evaluate the impact of MMR on GCV-induced mutagenesis [27].

A caveat of using target panel sequencing to detect mutational signatures is that the number of mutations observed is limited compared to whole genomes and exomes. In our analysis, we found that FDR estimation is unreliable for samples with less than 10 mutations. As such, it is possible that our analysis underestimates the true prevalence of GCV-associated mutagenesis. Nevertheless, consistent with our prevalence estimate, we note that there were only two further cancers potentially with GCV<sup>sig</sup> from a recently published cohort of 12,222 whole cancer genomes [28]. It remains to be established whether GCV<sup>sig</sup> is more common as a background mutational process in the normal cells of transplant recipients who have been treated with GCV. Another limitation relates to using cell line and organoid models to examine GCV- and MMF-associated mutagenesis. It will be important to establish the contribution of MMF to GCV-induced mutagenesis using patient material or *in vivo* animal models or through pharmacoepidemiology investigations.

## Conclusions

This study provides further evidence that GCV treatment influences cancer development in diverse organs. While the prevalence is relatively low, at 1 in 5500 in an unselected cohort of 121,771 cancer patients, it is likely much more common in patients who have undergone organ transplantation. Our cell line and organoid models further suggest that co-treatment with MMF increases the mutagenicity of GCV. Further studies should be performed to characterise the safety profile of GCV and combination treatments in humans.

## Abbreviations

GCV: Ganciclovir; MMF: Mycophenolate mofetil; GENIE: Genomics Evidence Neoplasia Information Exchange; FM: Foundation Medicine; HSCT: Haematopoietic stem cell transplants; CRC: Colorectal cancer; CMV: Cytomegalovirus; GCV<sup>sig</sup>: GCV-associated signature; FDR: False discovery rate; GCV-TP: GCV-triphosphate; MMRd: Mismatch repair deficiency; HIV: Human immunodeficiency virus; AIDS: Acquired immunodeficiency syndrome.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13073-022-01131-w>.

**Additional file 1: Table S1.** Clinical and mutational information of patients with GCV<sup>sig</sup> identified in this study and from our previous patients. **Table S2.** Patient and panel characteristics of GENIE and FM cohorts. **Table S3.** Concentrations used for the treatment of respective cell lines for WGS mutation experiments. **Table S4.** Cosine similarity of COSMICv3 signatures with GCV<sup>sig</sup>. **Table S5.** GENIE and FM samples analysed in the study with their trinucleotide mutational signature and similarity to GCV<sup>sig</sup> shown. **Table S6.** Mutational spectrum and sigfit contribution score to COSMIC signatures for H414, HCT116 and normal colon organoid H023. **Table S7.** Mutational spectrum of mouse myeloid cell line.

**Additional file 2.** Supplementary Methods.

**Additional file 3: Figure S1.** Determination of sensitivity and false discovery rate based on SBS18 and SBS38. **Figure S2.** Trinucleotide mutational spectra of the additional 21 samples with GCVsig detected from AACR Project GENIE and Foundation Medicine cohorts. **Figure S3.** Permutation test to determine cohort-wide false discovery rate. **Figure S4.** Replication strand bias for cell lines and cancer samples. **Figure S5.** Mutation burden of GCV positive samples relative to other samples by cancer type. **Figure S6.** Treatment of cells with ganciclovir and acyclovir. **Figure S7.** Treatment of cells with ganciclovir (GCV) and mycophenolate mofetil (MMF). **Figure S8.** Absolute mutations for GCV-induced mutations(CA>AA), C>T and T>C mutations across H414, HCT116 cell line and H023 organoid models. **Figure S9.** Copy number variation across the genomes of H414 and HCT116 clones. The ploidy values are calculated using Control-FREEC from each sample's whole genome sequencing data. **Figure S10.** Histogram showing the distribution variant allele frequency (VAF) of mutations from diploid regions in each clone of H414 (A) and HCT116 (B).

## Acknowledgements

The authors thank David Shih for providing valuable advice for statistical analysis. The authors also thank the Centre for PanorOmic Sciences (CPOS) of the University of Hong Kong Li Ka Shing Faculty of Medicine and the Australian Genome Research Facility for performing mass spectrometry and next-generation sequencing. The authors would like to acknowledge the American Association for Cancer Research and its financial and material support in the development of the AACR Project GENIE registry, as well as members of the consortium for their commitment to data sharing. Interpretations are the responsibility of the study authors.

## Authors' contributions

Concept and design: JWHW, SYL, and IJM. Initial in vitro experiments identifying the association of GCV and MMF with GCV<sup>sig</sup> in organoids: HHNY. Cell line experiments: RAB, HHNY, JAB, ZX, MX, JL, and MSYH. Bioinformatics and statistical analysis: HF, RAB, CF, HCS, and JWHW. Mass spectrometry analysis: STL, RS, and JWHW. Clinical data analysis: MT, EC, HWI, CYMC, EAS-C, and SYL. Drafting of the manuscript: JWHW and HF. Critical revision of the manuscript for important intellectual content: SYL and IJM. Obtained funding: JWHW, SYL, and IJM. Administrative, technical, or material support: JWHW. Supervision: JWHW, SYL, and IJM. The authors read and approved the final manuscript.

## Funding

The project was supported by the Research Grants Council, HK (17100920 and C7028-19G), to JWHW; a theme-based research grant from the Research Grants Council of the Hong Kong SAR (Project No. T12-710/16-R) to SYL; the Centre for Oncology and Immunology under the Health@InnoHK Initiative funded by the Innovation and Technology Commission, The Government of

Hong Kong SAR, China; donations from the Hong Kong Jockey Club Charities Trust and the Kadoorie Charitable Foundation; and the Victorian State Government Operational Infrastructure Support and Australian Government NHMRC IRIISS.

## Availability of data and materials

Raw sequencing reads are available from NCBI SRA PRJNA830636 (<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA830636>) [29] for H414 and HCT116 cell lines and PRJNA837717 (<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA837717>) [30] for murine myeloid cell lines and EGA EGAS00001006707 (<https://ega-archive.org/studies/EGAS00001006707>) [31] for organoids.

## Declarations

### Ethics approval and consent to participate

This study was approved by the Institutional Review Board (IRB) of The University of Hong Kong/Hospital Authority Hong Kong West Cluster (UW14-257 and UW18-599) and by the Human Research Ethics Committees from the Walter & Eliza Hall Institute (project 13/01) and Melbourne Health (project 2012.274). All patients involved consented to the use of clinical information and tissue for research. In accordance with the IRB approvals, all the research conducted conformed to the principles of the Helsinki Declaration.

### Consent for publication

Not applicable.

### Competing interests

SYL has received research sponsorship from Pfizer, Merck, Servier, and Curegenix. The remaining authors declare that they have no competing interests.

### Author details

<sup>1</sup>School of Biomedical Sciences, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China. <sup>2</sup>Department of Research & Development, South China Hospital, Health Science Center, Shenzhen University, Shenzhen 518116, China. <sup>3</sup>Department of Pathology, School of Clinical Medicine, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China. <sup>4</sup>Centre for Oncology and Immunology, Hong Kong Science Park, Hong Kong SAR, China. <sup>5</sup>The Walter and Eliza Hall Institute of Medical Research, 1G Royal Parade, Parkville, VIC 3052, Australia. <sup>6</sup>Department of Medical Biology, University of Melbourne, 1G Royal Parade, Parkville, VIC 3052, Australia. <sup>7</sup>Division of Hematology/Oncology, Department of Pediatrics, University of California San Francisco, San Francisco, CA, USA. <sup>8</sup>Centre for PanorOmic Sciences, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China. <sup>9</sup>Department of Pathology, Queen Mary Hospital, Hong Kong SAR, China. <sup>10</sup>Department of Medicine, Queen Mary Hospital, Hong Kong SAR, China. <sup>11</sup>The Jockey Club Centre for Clinical Innovation and Discovery, LKS Faculty of Medicine, The University of Hong Kong, Pokfulam, Hong Kong SAR, China.

Received: 1 July 2022 Accepted: 20 October 2022

Published online: 31 October 2022

## References

- Baker KS, DeFor TE, Burns LJ, Ramsay NK, Neglia JP, Robison LL. New malignancies after blood or marrow stem-cell transplantation in children and adults: incidence and risk factors. *J Clin Oncol.* 2003;21(7):1352–8.
- Engels EA, Pfeiffer RM, Fraumeni JF Jr, Kasiske BL, Israni AK, Snyder JJ, Wolfe RA, Goodrich NP, Bayakly AR, Clarke CA, et al. Spectrum of cancer risk among US solid organ transplant recipients. *JAMA.* 2011;306(17):1891–901.
- Inman GJ, Wang J, Nagano A, Alexandrov LB, Purdie KJ, Taylor RG, Sherwood V, Thomson J, Hogan S, Spender LC, et al. The genomic landscape of cutaneous SCC reveals drivers and a novel azathioprine associated mutational signature. *Nat Commun.* 2018;9(1):3667.
- Dierickx D, Habermann TM. Post-transplantation lymphoproliferative disorders in adults. *N Engl J Med.* 2018;378(6):549–62.
- Hosseini-Moghaddam SM, Soleimanirahbar A, Mazzulli T, Rotstein C, Husain S. Post renal transplantation Kaposi's sarcoma: a review of its

- epidemiology, pathogenesis, diagnosis, clinical aspects, and therapy. *Transpl Infect Dis.* 2012;14(4):338–45.
6. de Kanter JK, Peci F, Bertrums E, Rosendahl Huber A, van Leeuwen A, van Roosmalen MJ, Manders F, Verheul M, Oka R, Brandsma AM, et al. Antiviral treatment causes a unique mutational signature in cancers of transplantation recipients. *Cell Stem Cell.* 2021;28(10):1726–1739 e1726.
  7. Palles C, West HD, Chew E, Galavotti S, Flensburg C, Grolleman JE, Jansen EAM, Curley H, Chegwidden L, Arbe-Barnes EH, et al. Germline MBD4 deficiency causes a multi-tumor predisposition syndrome. *Am J Hum Genet.* 2022;109(5):953–60.
  8. Yan HHN, Siu HC, Ho SL, Yue SSK, Gao Y, Tsui WY, Chan D, Chan AS, Wong JWH, Man AHY, et al. Organoid cultures of early-onset colorectal cancers reveal distinct and rare genetic profiles. *Gut.* 2020;69(12):2165–79.
  9. Consortium APG. AACR Project GENIE: powering precision medicine through an international consortium. *Cancer Discov.* 2017;7(8):818–31.
  10. Hartmaier RJ, Albacker LA, Chmielecki J, Bailey M, He J, Goldberg ME, Ramkissoon S, Suh J, Elvin JA, Chiacchia S, et al. High-throughput genomic profiling of adult solid tumors reveals novel insights into cancer pathogenesis. *Cancer Res.* 2017;77(9):2464–75.
  11. Sondka Z, Bamford S, Cole CG, Ward SA, Dunham I, Forbes SA. The COSMIC Cancer Gene Census: describing genetic dysfunction across all human cancers. *Nat Rev Cancer.* 2018;18(11):696–705.
  12. Lee BCH, Robinson P, Coorens T, Yan HHN, Olafsson S, Lee-Six H, Sanders M, Siu HC, Hewinson J, Yue SSK et al: Mutational landscape of normal epithelial cells in Lynch syndrome patients. *Nat Commun* 2022, In press.
  13. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009;25(14):1754–60.
  14. Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, Gabriel S, Meyerson M, Lander ES, Getz G. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol.* 2013;31(3):213–9.
  15. Kim S, Scheffler K, Halpern AL, Bekritsky MA, Noh E, Kallberg M, Chen X, Kim Y, Beyter D, Krusche P, et al. Strelka2: fast and accurate calling of germline and somatic variants. *Nat Methods.* 2018;15(8):591–4.
  16. Flensburg C, Sargeant T, Oshlack A, Majewski IJ. SuperFreq: integrated mutation detection and clonal tracking in cancer. *PLoS Comput Biol.* 2020;16(2):e1007603.
  17. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, Bignell GR, Bolli N, Borg A, Borresen-Dale AL, et al. Signatures of mutational processes in human cancer. *Nature.* 2013;500(7463):415–21.
  18. Gori K, Baez-Ortega A: sigfit: flexible Bayesian inference of mutational signatures. *bioRxiv* 2020:372896. <https://doi.org/10.1101/372896>.
  19. Alexandrov LB, Kim J, Haradhvala NJ, Huang MN, Tian Ng AW, Wu Y, Boot A, Covington KR, Gordenin DA, Bergstrom EN, et al. The repertoire of mutational signatures in human cancer. *Nature.* 2020;578(7793):94–101.
  20. Nadeu F, Royo R, Massoni-Badosa R, Playa-Albinyana H, Garcia-Torre B, Duran-Ferrer M, Dawson KJ, Kulis M, Diaz-Navarro A, Villamor N, et al. Detection of early seeding of Richter transformation in chronic lymphocytic leukemia. *Nat Med.* 2022;28(8):1662–71.
  21. Elion GB. The biochemistry and mechanism of action of acyclovir. *J Antimicrob Chemother.* 1983;12(Suppl B):9–17.
  22. Thust R, Tomicic M, Klocking R, Voutilainen N, Wutzler P, Kaina B. Comparison of the genotoxic and apoptosis-inducing properties of ganciclovir and penciclovir in Chinese hamster ovary cells transfected with the thymidine kinase gene of herpes simplex virus-1: implications for gene therapeutic approaches. *Cancer Gene Ther.* 2000;7(1):107–17.
  23. Allison AC, Eugui EM. Mycophenolate mofetil and its mechanisms of action. *Immunopharmacology.* 2000;47(2–3):85–118.
  24. Billat PA, Sauvage FL, Picard N, Tafzi N, Alain S, Essig M, Marquet P, Saint-Marcoux F. Liquid chromatography tandem mass spectrometry quantitation of intracellular concentrations of ganciclovir and its phosphorylated forms. *Anal Bioanal Chem.* 2015;407(12):3449–56.
  25. Miyao K, Terakura S, Ozawa Y, Sawa M, Kohno A, Kasahara S, Iida H, Ino K, Kusumoto S, Kasai M, et al. Comparison of transplantation outcomes after foscarnet and ganciclovir administration as first-line anti-cytomegalovirus preemptive therapy. *Transplant Cell Ther.* 2021;27(4):342 e341–342 e310.
  26. Reusser P, Einsele H, Lee J, Volin L, Rovira M, Engelhard D, Finke J, Cordonnier C, Link H, Ljungman P, et al. Randomized multicenter trial of foscarnet versus ganciclovir for preemptive therapy of cytomegalovirus infection after allogeneic stem cell transplantation. *Blood.* 2002;99(4):1159–64.
  27. Zou X, Koh GCC, Nanda AS, Degasperis A, Urgo K, Roumeliotis TI, Agu CA, Badja C, Momen S, Young J, et al. A systematic CRISPR screen defines mutational mechanisms underpinning signatures caused by replication errors and endogenous DNA damage. *Nat Cancer.* 2021;2(6):643–57.
  28. Degasperis A, Zou X, Amarante TD, Martinez-Martinez A, Koh GCC, Dias JML, Heskin L, Chmelova L, Rinaldi G, Wang VYW, et al. Substitution mutational signatures in whole-genome sequenced cancers in the UK population. *Science.* 2022;376(6591):abl9283.
  29. Fang H, Yan HHN, Bilardi RA, et al. Ganciclovir-induced mutations are present in a diverse spectrum of post-transplant malignancies. PRJNA830636, NCBI Sequence Read Archive. 2022. <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA830636>.
  30. Fang H, Yan HHN, Bilardi RA, et al. Ganciclovir-induced mutations are present in a diverse spectrum of post-transplant malignancies. PRJNA837717, NCBI Sequence Read Archive. 2022. <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA837717>
  31. Fang H, Yan HHN, Bilardi RA, et al: Ganciclovir-induced mutations are present in a diverse spectrum of post-transplant malignancies. EGAS00001006707, EGA. 2022. <https://ega-archive.org/studies/EGAS00001006707>.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

