Journal of Pathology
J Pathol 2017; **242:** 10–15
Published online 29 March 2017 in Wiley Online Library (wileyonlinelibrary.com) DOI: 10.1002/path.4880

BRIEF DEFINITIVE REPORT

Mutational signature analysis identifies *MUTYH* deficiency in colorectal cancers and adrenocortical carcinomas

Camilla Pilati[1†], Jayendra Shinde[2,3,4,5†], Ludmil B Alexandrov[6,7†], Guillaume Assié[8,9], Thierry André[10,11], Zofia Hélias-Rodzewicz[12,13], Romain Ducoudray[12,13], Delphine Le Corre[1], Jessica Zucman-Rossi[2,3,4,5], Jean-François Emile[12,13], Jérôme Bertherat[8,9‡], Eric Letouzé[2,3,4,5‡] and Pierre Laurent-Puig[1‡*]

[1] *INSERM UMR-S1147, Personalized Medicine, Pharmacogenomics, Therapeutic Optimization, Université Paris Descartes, Paris, France*
[2] *INSERM, Unité Mixte de Recherche (UMR) 1162, Génomique Fonctionnelle des Tumeurs Solides, Equipe Labellisée Ligue contre le Cancer, Paris, France*
[3] *Université Paris Descartes, Labex Immuno-Oncology, Sorbonne Paris Cité, Paris, France*
[4] *Université Paris 13, Sorbonne Paris Cité, Unité de Formation et de Recherche (UFR) Santé, Médecine, Biologie Humaine (SMBH), Bobigny, France*
[5] *Université Paris Diderot, Institut Universitaire d'Hématologie, Paris, France*
[6] *Theoretical Biology and Biophysics (T-6), Los Alamos National Laboratory, Los Alamos, NM, USA*
[7] *Center for Nonlinear Studies, Los Alamos National Laboratory, Los Alamos, NM, USA*
[8] *INSERM U1016, CNRS UMR 8104, Paris Descartes University, Institut Cochin, Paris, France*
[9] *Center for Rare Adrenal Diseases, Department of Endocrinology, Assistance Publique-Hôpitaux de Paris, Hôpital Cochin, Paris, France*
[10] *Department of Medical Oncology AP-HP, Hospital Saint-Antoine, Paris, France*
[11] *Université Pierre et Marie Curie (UMPC) Paris VI, Paris, France*
[12] *Department of Pathology AP-HP, Hôpital Ambroise Paré, Paris, France*
[13] *EA 4340, Université de Versailles, Versailles, France*

**Correspondence to: P Laurent-Puig, 45 Rue des Saints Pères, 75006 Paris, France. E-mail: pierre.laurent-puig@parisdescartes.fr*

†*These authors contributed equally to this work.*
‡*These authors also contributed equally to this work.*

Abstract

Germline alterations in DNA repair genes are implicated in cancer predisposition and can result in characteristic mutational signatures. However, specific mutational signatures associated with base excision repair (BER) defects remain to be characterized. Here, by analysing a series of colorectal cancers (CRCs) using exome sequencing, we identified a particular spectrum of somatic mutations characterized by an enrichment of C > A transversions in NpCpA or NpCpT contexts in three tumours from a MUTYH-associated polyposis (MAP) patient and in two cases harbouring pathogenic germline *MUTYH* mutations. In two series of adrenocortical carcinomas (ACCs), we identified four tumours with a similar signature also presenting germline *MUTYH* mutations. Taken together, these findings demonstrate that *MUTYH* inactivation results in a particular mutational signature, which may serve as a useful marker of BER-related genomic instability in new cancer types.

Keywords: mutational signatures; MUTYH; colorectal cancer; adrenocortical carcinomas

Received 10 October 2016; Revised 9 January 2017; Accepted 16 January 2017

No conflicts of interest were declared

Introduction

Inherited genetic defects affecting several DNA repair pathways have been associated with cancer predisposition. Defects in mismatch repair (mainly *MLH1*, *MSH2* or *MSH6* mutations associated with hereditary non-polyposis colorectal cancer), base excision repair [*MUTYH* mutations in MUTYH-associated polyposis (MAP)], nucleotide excision repair (xeroderma pigmentosum genes in skin cancer), homologous repair (mainly *BRCA1/2* in breast and other cancers), and the Fanconi anaemia pathway (*FANC* genes in acute myeloid leukaemia and squamous cell carcinomas) lead to different types of genetic instability that should be recognizable as distinct mutational signatures in cancer genomes [1]. At present, at least 30 mutational signatures have been identified by examining the cancers of more than 12 000 patients across 40 distinct human cancer types [2–4] and these are reported in the COSMIC database. Specific signatures have been related to *BRCA1/2* mutations, mismatch repair, and nucleotide-excision repair deficiency [4,5], but the mutational signatures of defects affecting the other DNA repair pathways remain to be characterized.

MUTYH is a DNA glycosylase involved in the repair of oxidative DNA damage, in particular of 8-oxoG, playing a critical role in the base excision repair (BER) pathway. Germline biallelic loss-of-function mutations in the *MUTYH* gene cause a condition predisposing to familial colorectal cancer (CRC) termed *MUTYH* association polyposis (MAP). Targeted and whole-exome sequencing (WES) studies revealed an excess of C > A mutations in MAP tumours [6,7]. Here, we describe a strong association between a particular spectrum of C > A single base substitutions, similar to the previously reported signature 18 in the COSMIC database, and *MUTYH* mutations in a series of CRCs and in two series of adrenocortical carcinomas (ACCs) analysed by WES.

Materials and methods

We performed whole-exome sequencing from fresh frozen tumour samples from 37 patients with advanced CRC (stage 4, treated with anti-EGFR antibodies), and two adenomas (P1 and P2) and a carcinoma (TU) from one patient with MAP (supplementary material, Table S1). Analysis of the triple knockout mouse model [8] and the ACC cohorts [9,10] were based on the mutation lists provided in the original publications. We used the Wellcome Trust Sanger Institute mutational signatures framework [2] to extract signatures from the catalogues of somatic mutations identified in CRC and ACC samples. Detailed methods are presented in supplementary material, Supplementary materials and methods.

Results and discussion

We performed a *de novo* analysis of the mutational spectrum of 37 CRCs using the Wellcome Trust Sanger Institute (WTSI) mutational signatures framework [2]. Mutational signature analysis identified three distinct signatures closely related to the previously described signatures 1, 5, and 18 [3] (supplementary material, Figure S1A). As previously reported [11], the clock-like signature 1 (but not signature 5) was positively correlated with age (Pearson coefficient = 0.2, $p = 0.0042$; supplementary material, Figure S1B). In contrast, the aetiology of signature 18 remains unknown. Signature 18, characterized by an enrichment of C > A transversions in NpCpA or NpCpT contexts (Figure 1A), has been described in neuroblastoma and in incidental samples of other cancer types, such as ACC, breast, and stomach carcinomas [11], but not in CRC. In our series, signature 18 was the dominant signature in two samples (CRC103 and CRC355, Figure 1B, C). Strikingly, these two cases were the only cases harbouring germline mutations (i.e. germline variants with a reported frequency ≤ 1% in the ExAC database [12]) in the *MUTYH* gene ($p = 0.0015$, Fisher's exact test). CRC103 displayed a nonsense mutation (rs762307622, W156*, ExAC frequency in non-Finnish Europeans = 0%) potentially leading to a truncated protein lacking most functional domains (supplementary material, Table S2). CRC355 harboured the known pathogenic variant rs36053993 (ExAC frequency = 0.4%) leading to the amino acid substitution G396D [13]. These two variants were heterozygous in the normal samples but became homozygous in the two tumours following loss of the wild-type allele (Figure 1D). To confirm the link between signature 18 and *MUTYH* mutations, we analysed by WES two adenomas (P1 and P2) and a carcinoma (TU) from one patient with MAP harbouring two different missense mutations of *MUTYH*: a Y179C mutation (rs34612342, ExAC frequency = 0.25%) known to be associated with polyposis [14], and an R495P mutation (ExAC frequency = 0%) predicted to be possibly damaging by using SNP&Go and SIFT software (supplementary material, Table S2). Signature 18 was the dominant signature in these three samples, accounting for more than 70% of mutations (Figure 1B). As a result, the mutational spectrum in *MUTYH*-mutated cases displayed a dramatic increase of C > A mutations at trinucleotide contexts characteristic of signature 18 (Figure 1C). The somatic mutations identified by Rashid *et al* in eight colorectal adenomas from two MAP patients [6] also displayed a trinucleotide spectrum very similar to signature 18 (cosine similarity = 0.85; supplementary material, Figure S2). *KRAS* G12C mutations are particularly frequent in MAP-related colorectal tumours [15] and were observed in our three MAP samples. These mutations result from C > A substitutions occurring in a CCA trinucleotide context showing a high mutation probability in signature 18 (Figure 1A). Thus, BER deficiency makes *MUTYH*-mutated colorectal cells particularly prone to acquiring C > A mutations at CCA sites, explaining the high frequency of *KRAS* G12C mutations in these patients.

To investigate further the link between *MUTYH* deficiency and the mutational signature 18, we analysed the mutations reported by Ohno *et al* in an *Mth1/Ogg1/Mutyh* triple knockout mouse model (TOY-KO), which accumulates 8-oxoG in the nuclear DNA of gonadal cells [8]. Using WES, Ohno *et al* identified 262 germline mutations acquired by the offspring of these mice. The trinucleotide mutation pattern of these mutations was highly consistent with the mutational pattern observed in our *MUTYH*-mutated CRC cases (Figure 2A), and a hierarchical clustering with the 30 signatures described in COSMIC showed that the TOY-KO mutational signature clustered together with signature 18 (Figure 2B). Indeed, the trinucleotide context frequencies of TOY-KO mice mutations and signature 18 displayed a cosine similarity of 0.94, where a similarity of 1.00 reflects a perfect match (Figure 2C).

In a recent study, we analysed the whole-exome sequences of 45 ACCs and reported two ACCs with a hypermutator phenotype dominated by C > A transversions [9]. Mutational signature analysis of these 45 ACCs identified signatures 1 and 5 in most samples and signatures 2 and 13 (related to the activity of AID/APOBEC cytidine deaminases) in three cases.

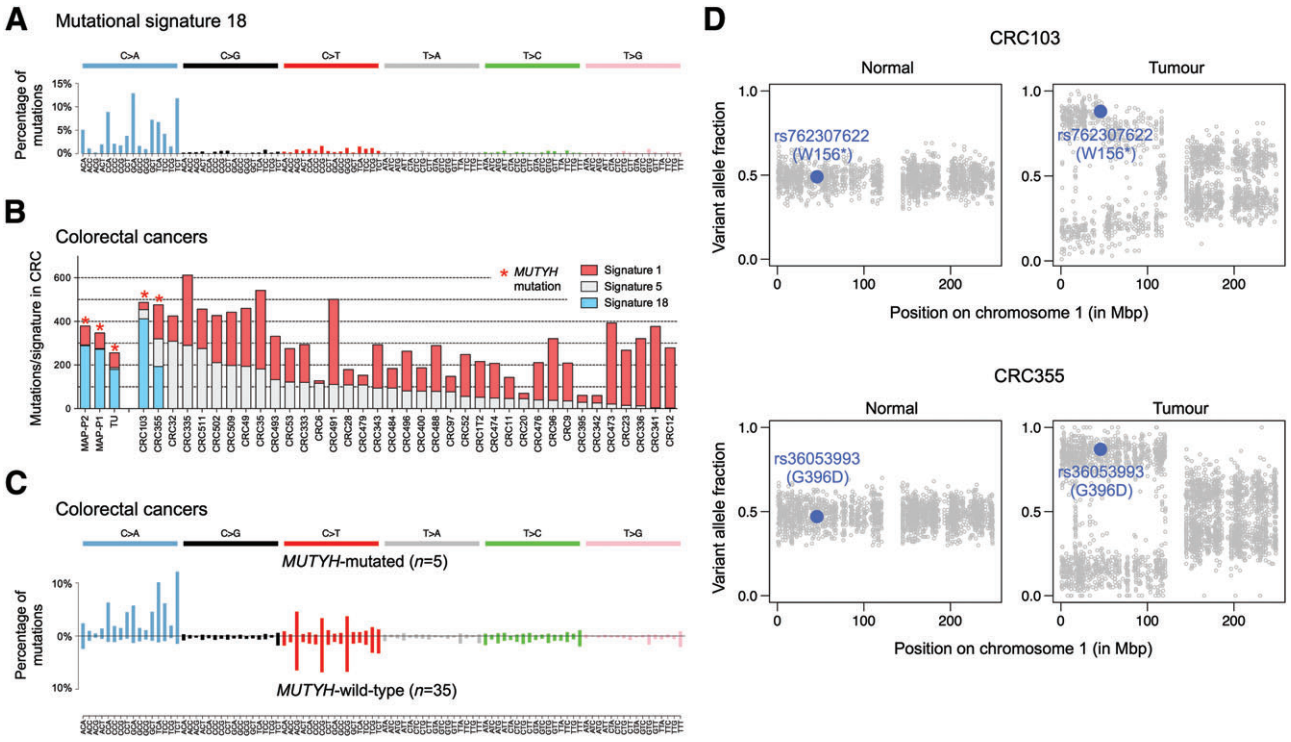

Figure 1. Mutational signatures identified in a cohort of 40 colorectal cancers (CRCs). (A) Mutational signature 18 is depicted using a 96-substitution classification defined by the substitution type and sequence context immediately 5′ and 3′ to the mutated base. (B) Contribution of the three signatures identified using the Wellcome Trust Sanger Institute mutational signatures framework to each tumour of the cohort of 40 CRCs. (C) Average mutational profile of colorectal tumours harbouring (top) or not (bottom) inactivating mutations of the *MUTYH* gene. (D) Loss of the wild-type *MUTYH* allele in the two cases harbouring monoallelic germline variants. The variant allele fraction is shown for all germline polymorphisms along chromosome 1 in the normal (left) and tumour (right) samples. Both *MUTYH* variants become homozygous in the tumours due to loss of heterozygosity (LOH) events.

Strikingly, signature 18 accounted for almost all mutations in the two hypermutated tumours (>10 mutations/Mb as opposed to a maximum of 1.2 mutations/Mb in the 43 others; Figure 3A, B). These two tumours were the only ones in the cohort harbouring germline *MUTYH* mutations ($p = 0.0010$, Fisher's exact test): ACC33 harboured the missense G396D variant (rs36053993), whereas ACC39 harboured the R241W variant (rs34126013, ExAC frequency = 0.005%) (supplementary material, Table S2). These two *MUTYH* mutations are also known to be pathogenic [13,16] and were found in cancers with deletion of the wild-type allele by LOH of chromosome 1 (Figure 3C). In an independent series of 91 ACCs from the TCGA [10], we identified two tumours with signatures similar to signature 18 (supplementary material, Figure S3). Both tumours displayed rare pathogenic *MUTYH* mutations (Y179C and W156*; supplementary material, Table S2), with loss of the wild-type allele in the tumour (supplementary material, Figure S3), versus none of the 89 other samples ($p = 0.00024$, Fisher's exact test).

To confirm the similarity of signatures related to *MUTYH* mutations in different biological contexts, we then performed an unsupervised classification of the mutational profiles observed in CRC and ACC cases, together with the 30 signatures currently described in COSMIC and the mutational signature of the TOY-KO mouse model. The hierarchical clustering (cosine distance, Ward method) revealed a *MUTYH* cluster comprising the three MAP tumours from our series, the two MAP patients from Rashid *et al* [6], CRC103, ACC33, and ACC39, together with the signature of TOY-KO mice mutations and COSMIC signature 18 (Figure 4). One *MUTYH*-mutated CRC (CRC355) clustered separately, due to the high prevalence of signature 1 in addition to signature 18 in this sample.

Taken together, our data suggest that mutational signature 18 reflects the accumulation of C > A mutations after incorrect replication of 8-oxoG. These mutations may accumulate in the context of defective base excision repair, e.g. in *MUTYH*-deficient cancers, or potentially by excessive generation of reactive oxygen species (ROS). Signature 18 has also been described in breast, stomach cancer, and neuroblastoma. We examined the whole-genome sequences of 560 breast cancers [17], 478 stomach cancers (TCGA-STAD), and 56 neuroblastomas [18], but we did not find any association between signature 18 and *MUTYH* mutations in these cancer types. Increased generation of ROS or other genetic defects or epigenetic alterations of BER pathway genes may explain signature 18 in these cancer types. Consistently, signature 18 is highly prevalent in neuroblastoma, a cancer type in which reactive oxygen species are known biological stimuli [19]. Finally, this study shows that mutational signatures can be used to detect

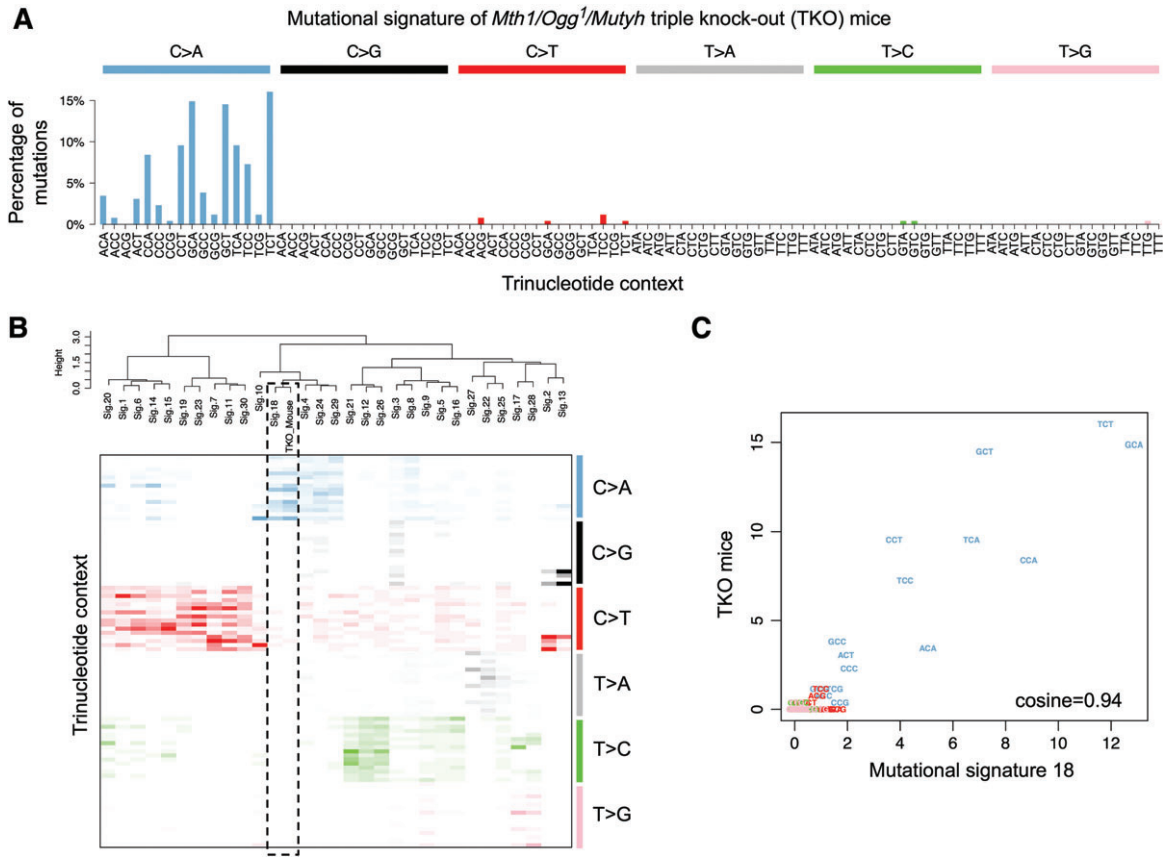

Figure 2. Mutations identified in a model of *MUTYH*-deficient mice are highly similar to mutational signature 18. (A) Mutational profile of 262 germline mutations identified by whole-exome sequencing in the offspring of *Mth1/Ogg1/Mutyh* triple knockout mice (TOY-KO mice). (B) Unsupervised classification of trinucleotide patterns corresponding to the 30 known mutational signatures [4] and mutations identified in TOY-KO mice [8]. The dendrogram (top) shows the grouping of signatures obtained by hierarchical clustering. The heatmap below shows the prevalence of the six substitution types (identified by a colour code) broken down according to the trinucleotide context. (C) Cosine similarity between the mutational patterns of TKO mice mutations and mutational signature 18. The proportion of mutations corresponding to each trinucleotide context in mutational signature 18 (*x*-axis) and TKO mice mutations (*y*-axis) is represented with the same colour code as in panel B (e.g. C > A mutations are indicated in blue).

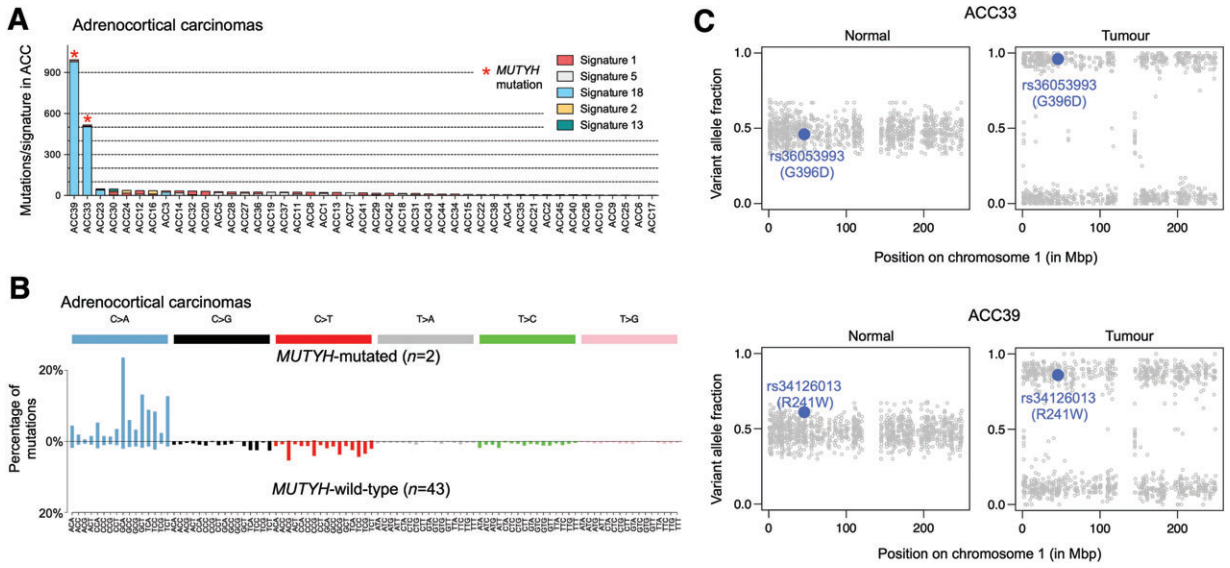

Figure 3. Mutational signatures identified in a cohort of 45 adrenocortical carcinomas (ACCs). (A) Contribution of the five signatures identified using the Wellcome Trust Sanger Institute mutational signatures framework to each tumour of the cohort of 45 ACCs. (B) Average mutational profile of ACC harbouring (top) or not (bottom) inactivating mutations of the *MUTYH* gene. (C) Loss of the wild-type *MUTYH* allele in the two ACC cases harbouring germline variants. The variant allele fraction is shown for all germline polymorphisms along chromosome 1 in the normal (left) and tumour (right) samples. Both *MUTYH* variants become homozygous in the tumours, due to loss of heterozygosity (LOH) events.

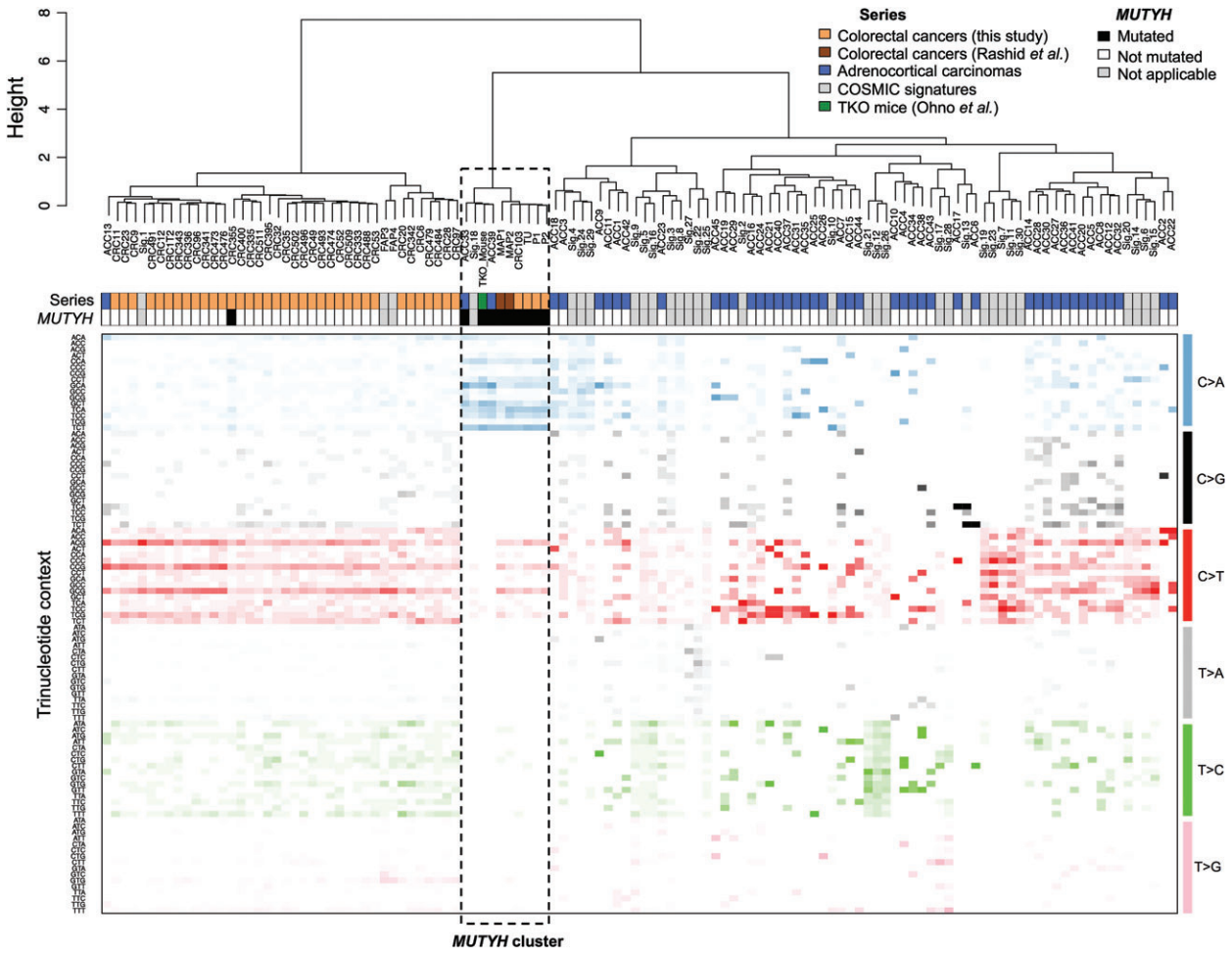

Figure 4. Unsupervised classification of the mutational patterns of 40 CRCs, two MAP patients (five and three colorectal adenomas), 45 ACCs, TKO mice mutations, and COSMIC mutational signatures. The dendrogram (top) shows the result of hierarchical clustering, highlighting a cluster of *MUTYH*-related signatures (dashed box). The heatmap below shows the prevalence of the six substitution types (identified by a colour code) broken down according to the trinucleotide context.

DNA damage defects in tumours, even before a causal mutation is identified. This approach led us to identify, for the first time to our knowledge, pathogenic *MUTYH* mutations in ACC patients. This finding has potential clinical utility since these patients may benefit from genetic counselling and treatments (e.g. immunotherapy) taking advantage of their high mutation load. Mutational signatures are thus a useful tool that may be used in the clinics to identify patients who may benefit from specific therapies [20].

Acknowledgements

CP was supported by a fellowship from ARC. JS was funded by the ICE (Interpretation of Clinical Exome) project (FUI, MEDICEN, Région Ile de France). LBA was supported through a J Robert Oppenheimer Fellowship at Los Alamos National Laboratory. The INSERM UMR-S1147 was supported by the Ministère de l'Enseignement Supérieur et de la Recherche, the Université Paris-Descartes, the Centre National de la Recherche Scientifique (CNRS), the Institut National de la Santé et de la Recherche Médicale (INSERM), the Agence Nationale de la Recherche (ANR Nanobiotechnologies; No ANR-10-NANO-0002-09), the SIRIC CARPEM, the Canceropole funding (No 2011-1-LABEL-UP5-2), and the Ligue Nationale contre le Cancer (Program 'Equipe labellisée LIGUE'; No EL2016.LNCC/VaT). The INSERM U1162 team was supported by the Ligue Nationale contre le Cancer. The INSERM U1016 team was supported by the PHRC COMETE-Tactic. This research used resources provided by the Los Alamos National Laboratory Institutional Computing Program, which is supported by the US Department of Energy National Nuclear Security Administration under Contract No DE-AC52-06NA25396. Research performed at Los Alamos National Laboratory was carried out under the auspices of the National Nuclear Security Administration of the United States Department of Energy.

Author contributions statement

EL and PL-P directed the research. CP, LBA, EL, and PL-P wrote the manuscript. DLC, ZH-R, and RD performed the experiments. CP, JS, LBA, EL, and PL-P analysed and interpreted the data. JS, LBA, and EL

carried out signatures and/or statistical analyses. GA, TA, JZ-R, J-FE, and JB provided essential biological resources and collected clinical data. All authors approved the final manuscript and contributed to critical revisions to its intellectual context.

References

1. Curtin NJ. DNA repair dysregulation from cancer driver to therapeutic target. *Nature Rev Cancer* 2012; **12:** 801–817.
2. Alexandrov LB, Nik-Zainal S, Wedge DC, *et al.* Deciphering signatures of mutational processes operative in human cancer. *Cell Rep* 2013; **3:** 246-259.
3. Alexandrov LB, Nik-Zainal S, Wedge DC, *et al.* Signatures of mutational processes in human cancer. *Nature* 2013; **500:** 415–421.
4. Alexandrov LB. Understanding the origins of human cancer. *Science* 2015; **350:** 1175.
5. Kim J, Mouw KW, Polak P, *et al.* Somatic *ERCC2* mutations are associated with a distinct genomic signature in urothelial tumors. *Nature Genet* 2016; **48:** 600–606.
6. Rashid M, Fischer A, Wilson CH, *et al.* Adenoma development in familial adenomatous polyposis and MUTYH-associated polyposis: somatic landscape and driver genes. *J Pathol* 2016; **238:** 98–108.
7. Sampson JR, Jones S, Dolwani S, *et al.* MutYH (MYH) and colorectal cancer. *Biochem Soc Trans* 2005; **33:** 679–683.
8. Ohno M, Sakumi K, Fukumura R, *et al.* 8-Oxoguanine causes spontaneous *de novo* germline mutations in mice. *Sci Rep* 2014; **4:** 4689.
9. Assie G, Letouze E, Fassnacht M, *et al.* Integrated genomic characterization of adrenocortical carcinoma. *Nature Genet* 2014; **46:** 607–612.
10. Zheng S, Cherniack AD, Dewal N, *et al.* Comprehensive pan-genomic characterization of adrenocortical carcinoma. *Cancer Cell* 2016; **29:** 723–736.
11. Alexandrov LB, Jones PH, Wedge DC, *et al.* Clock-like mutational processes in human somatic cells. *Nature Genet* 2015; **47:** 1402–1407.
12. Lek M, Karczewski KJ, Minikel EV, *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 2016; **536:** 285–291.
13. Lubbe SJ, Di Bernardo MC, Chandler IP, *et al.* Clinical implications of the colorectal cancer risk associated with *MUTYH* mutation. *J Clin Oncol* 2009; **27:** 3975–3980.
14. Al-Tassan N, Chmiel NH, Maynard J, *et al.* Inherited variants of *MYH* associated with somatic G:C→T:A mutations in colorectal tumors. *Nature Genet* 2002; **30:** 227–232.
15. Jones S, Lambert S, Williams GT, *et al.* Increased frequency of the k-*ras* G12C mutation in MYH polyposis colorectal adenomas. *Br J Cancer* 2004; **90:** 1591–1593.
16. Bai H, Jones S, Guan X, *et al.* Functional characterization of two human MutY homolog (hMYH) missense mutations (R227W and V232F) that lie within the putative hMSH6 binding domain and are associated with hMYH polyposis. *Nucleic Acids Res* 2005; **33:** 597–604.
17. Nik-Zainal S, Davies H, Staaf J, *et al.* Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature* 2016; **534:** 47–54.
18. Peifer M, Hertwig F, Roels F, *et al.* Telomerase activation by genomic rearrangements in high-risk neuroblastoma. *Nature* 2015; **526:** 700–704.
19. Marengo B, Raffaghello L, Pistoia V, *et al.* Reactive oxygen species: biological stimuli of neuroblastoma cell response. *Cancer Lett* 2005; **228:** 111–116.
20. Alexandrov LB, Nik-Zainal S, Siu HC, *et al.* A mutational signature in gastric cancer suggests therapeutic strategies. *Nature Commun* 2015; **6:** 8683.

*21. Guichard C, Amaddeo G, Imbeaud S, *et al.* Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. *Nat Genet* 2012; **44:** 694–698.

*22. Cancer Genome Atlas N. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* 2012; **487:** 330–337.

*Cited only in supplementary material.

SUPPLEMENTARY MATERIAL ONLINE

Supplementary materials and methods

Supplementary figure legends

Figure S1. Mutational signatures and correlations for 38 CRC patients

Figure S2. Mutational profiling of MAP patient samples and their similarity to mutational signature 18

Figure S3. Comparison of mutational patterns in two ACC series and COSMIC mutational signature 18

Table S1. Somatic single nucleotide variants (SNVs) among the 40 colorectal tumours

Table S2. MUTYH SNPs' characteristics