



## ARTICLE

## Cellular and Molecular Biology

# A common polymorphism in the retinoic acid pathway modifies adrenocortical carcinoma age-dependent incidence

Mirvat Surakhy<sup>1</sup>, Marsha Wallace<sup>1</sup>, Elisabeth Bond<sup>1</sup>, Lukasz Filip Grochola<sup>2,3</sup>, Husein Perez<sup>4</sup>, Matteo Di Giovannantonio<sup>1</sup>, Ping Zhang<sup>1</sup>, David Malkin<sup>5</sup>, Hannah Carter<sup>6</sup>, Ivy Zortea S. Parise<sup>7</sup>, Gerard Zambetti<sup>8</sup>, Heloisa Komechen<sup>7</sup>, Mariana M. Paraizo<sup>7</sup>, Meghana S. Pagadala<sup>6</sup>, Emilia M. Pinto<sup>8</sup>, Enzo Lalli<sup>9</sup>, Bonald C. Figueiredo<sup>7,10,11</sup> and Gareth L. Bond<sup>1</sup>

**BACKGROUND:** Genome-wide association studies (GWASs) have enriched the fields of genomics and drug development.

Adrenocortical carcinoma (ACC) is a rare cancer with a bimodal age distribution and inadequate treatment options. Paediatric ACC is frequently associated with *TP53* mutations, with particularly high incidence in Southern Brazil due to the *TP53* p.R337H (R337H) germline mutation. The heterogeneous risk among carriers suggests other genetic modifiers could exist.

**METHODS:** We analysed clinical, genotype and gene expression data derived from paediatric ACC, R337H carriers, and adult ACC patients. We restricted our analyses to single nucleotide polymorphisms (SNPs) previously identified in GWASs to associate with disease or human traits.

**RESULTS:** A SNP, rs971074, in the alcohol dehydrogenase 7 gene significantly and reproducibly associated with allelic differences in ACC age-of-onset in both cohorts. Patients homozygous for the minor allele were diagnosed up to 16 years earlier. This SNP resides in a gene involved in the retinoic acid (RA) pathway and patients with differing levels of RA pathway gene expression in their tumours associate with differential ACC progression.

**CONCLUSIONS:** These results identify a novel genetic component to ACC development that resides in the retinoic acid pathway, thereby informing strategies to develop management, preventive and therapeutic treatments for ACC.

*British Journal of Cancer* (2020) 122:1231–1241; <https://doi.org/10.1038/s41416-020-0764-3>

**BACKGROUND**

Adrenocortical carcinoma (ACC) is a rare cancer that usually occurs in the first or after the fifth decade of life.<sup>1</sup> The incidence of most paediatric carcinomas increases progressively with age, but ACC has a peak incidence between birth and 4 years of age.<sup>2</sup> Li-Fraumeni syndrome (LFS) families have a higher risk of developing paediatric ACC, with more than 50% of all children with ACC having a heritable *TP53* mutation.<sup>1</sup> The germline R337H mutation was spread through a founder effect.<sup>3,4</sup> Newborn screening for the R337H in Southern and South-eastern Brazil showed that this mutation has a frequency of 0.21–0.30%.<sup>5–7</sup> This mutation is responsible for the increased incidence of paediatric ACC in Paraná state (3.4 per million before the age 15 years per year compared to 0.3 per million in the USA, or 6.4/million children younger than 10 years per year).<sup>2,5</sup> A hospital-based cohort estimated that the penetrance of this mutation for paediatric ACC is 10%.<sup>8</sup> However in a large population-based cohort, the penetrance for paediatric ACC is less than 5%, whilst the overall

cancer risk among females and males is less than 50% before the seventh decade.<sup>5,7</sup> The frequency of R337H prevalence among paediatric ACC varies from 75–96%, and is nearly 15% among adult ACC in Southern and South-eastern Brazil.<sup>7,9,10</sup> The phenotypic heterogeneity among R337H carriers in terms of cancer risk and the variability in age-of-onset may indicate the involvement of environmental and/or genetic variants that are able to modulate cancer development.<sup>5,7</sup> The enormous heterogeneity in cancer development noted in families carrying the germline R337H mutation<sup>5</sup> has prompted us to gain further insight into potential genetic factors cooperating to modify the age-dependent incidence of cancer onset in the adrenal cortex.

ACC requires more effective treatment strategies. The prognosis of children with residual or metastatic ACC is dismal<sup>11</sup> and can be worse in older children<sup>12</sup> and adult patients.<sup>13,14</sup> Knowing which gene(s) play important roles in the multistep tumorigenesis process in adrenocortical cells will certainly inform strategies to develop preventive and/or therapeutic protocols. Growing

<sup>1</sup>Ludwig Institute for Cancer Research, Nuffield Department of Clinical Medicine, University of Oxford, Oxford, UK; <sup>2</sup>Institute for Regenerative Medicine (IREM), University of Zurich, Zurich, Switzerland; <sup>3</sup>Department of Surgery, Cantonal Hospital Winterthur, Winterthur, Switzerland; <sup>4</sup>Faculty of Technology, Design and Environment, Oxford Brookes University, Oxford, UK; <sup>5</sup>Division of Hematology/Oncology, The Hospital for Sick Children, Department of Pediatrics, University of Toronto, Toronto, Canada; <sup>6</sup>Division of Medical Genetics, Department of Medicine, University of California, San Diego, USA; <sup>7</sup>Instituto de Pesquisa Pelé Pequeno Príncipe, Faculdades Pequeno Príncipe, Curitiba, Brazil; <sup>8</sup>Department of Pathology, St. Jude Children's Research Hospital, Memphis, TN, USA; <sup>9</sup>Institut de Pharmacologie Moléculaire et Cellulaire CNRS, Université Côte D'Azur, Inserm, Valbonne, France; <sup>10</sup>Departamento de Saúde Coletiva, Universidade Federal do Paraná, Curitiba, PR, Brazil and <sup>11</sup>Centro de Genética Molecular e Pesquisa do Câncer em Crianças (CEGEMPAC), Curitiba, PR, Brazil

Correspondence: Bonald C. Figueiredo ([bonald@ufpr.br](mailto:bonald@ufpr.br)) or Gareth L. Bond ([gareth.bond@ludwig.ox.ac.uk](mailto:gareth.bond@ludwig.ox.ac.uk))

These authors contributed equally: Mirvat Surakhy, Marsha Wallace.

Received: 18 April 2019 Revised: 9 January 2020 Accepted: 4 February 2020

Published online: 9 March 2020

genomic data has allowed recent studies to highlight the importance of heritable genetic determinants in identifying potential drug targets and disease treatments.<sup>15–17</sup> For example, one analysis showed that if a drug target has heritable genetic data supporting its role in a given disease, the chance of successful clinical development of a drug is doubled.<sup>16</sup> Through advances in genome genotyping and sequencing technologies, the identification of genetic variations associating with disease traits has rapidly increased. For example, GWASs have identified over 900 SNPs significantly associated with cancer susceptibility traits.<sup>18</sup> However, the necessity of large cohort sizes in GWASs to overcome the burden of multiple hypothesis testing has resulted in a relative deficit in our knowledge of the genetic variants underlying rare cancers like ACC.<sup>19,20</sup>

The genetic data supporting the role of p53 in ACC has improved the early diagnosis of ACC in *TP53* mutation carriers through asymptomatic screening.<sup>5,21</sup> However, targeting p53 in either the preventative or therapeutic setting remains a challenge.<sup>1</sup> The identification of other more druggable gene targets with genetic associations with ACC could prove useful, whereby SNPs in genes that associate with differential ACC cancer risk in *TP53* mutation carriers could help to identify novel druggable genes and/or pathways. SNPs in the p53 pathway have been clearly shown to associate with differential cancer risk. Indeed, SNPs in well-described p53 pathway genes have been frequently identified in GWASs of seemingly sporadic cancers.<sup>18</sup> In fact, one p53 pathway SNP, *MDM2* SNP309 that associates with differential expression levels of a key regulator of p53, has been shown to modify cancer risk in *TP53* mutation carriers with LFS.<sup>22</sup> Specifically, LFS families with more penetrant *TP53* mutations relative to the Brazilian R337H, are at higher risks of developing many types of cancer, such as sarcomas, breast, and brain cancers, as well as ACC. *MDM2* SNP309 has been shown to modify this cancer risk in multiple cohorts of *TP53* mutation carriers in an age- and gender-dependent manner.<sup>22–26</sup> Importantly, genetically engineered mouse models carrying either alleles of *MDM2* SNP309 and a highly penetrant LFS p53 mutation demonstrated similar allele-specific differences in age-dependent cancer incidence.<sup>26</sup> Together, observations such as these suggest that SNPs could potentially modify ACC cancer risk. Here, we studied two independent cohorts of ACC patients (paediatric and adult) to demonstrate that a SNP in a druggable gene and pathway significantly associates with ACC age-dependent incidence.

## METHODS

### Study cohorts

**TP53 p.R337H paediatric ACC cohorts.** A cohort of children diagnosed with ACC with the R337H germline mutation was used in this study. We divided this cohort into a discovery cohort (BRZ1,  $N = 26$ ) and a validation cohort (BRZ2,  $N = 16$ ) based on when the DNAs were received from the Pathology unit. The patients were treated in the Pequeno Príncipe Hospital, Brazil. This study was approved by the Ethics Committee from Pequeno Príncipe Hospital, and authorisation was obtained by signing a consent form. Clinical data that included the age at diagnosis for each patient, gender, and *TP53* mutation status were available.

**Adult ACC cohort.** A retrospective database for primarily sporadic ACC patients ( $N = 92$ ) was analysed from the Cancer Genome Atlas.<sup>27</sup> These patients were diagnosed with adult ACC and blood-derived genotype data from the Affymetrix™ Genome-Wide Human SNP Array 6.0 was available together with the relevant clinical information. Approval from the National Institute of Health (NIH) Data Access Committee was obtained (request number 34594–5). Somatic mutations of *TP53* for this data set were retrieved from cBioPortal (<http://www.cbioportal.org/>) website (accessed December 2018).

### Genomic DNA preparation

Genomic DNAs from the BRZ patients were extracted from peripheral blood using the ReliaPrep™ gDNA kit as per the manufacturer's recommendations (A2051, Promega).

### SNP genotyping

Genomic DNAs from the BRZ cohort were genotyped using the Infinium® LCG assay on the Human Omni 2.5–8 v1.3 BeadChip (Illumina platform), according to the manufacturer's instructions and at the High-Throughput Genomics at the Wellcome Trust Centre for Human Genetics, University of Oxford.

### Cancer GWAS and GWAS SNPs

The GWAS Catalog was downloaded from the European Bioinformatics Institute (EMBL-EBI) website (<https://www.ebi.ac.uk/gwas/>) on February 2018. We found 35,705 GWAS lead SNPs in studies with individuals with European (EUR) ancestry. We then retrieved all closely linked SNPs (proxies) based on the 1000 Genomes phase 3 data acquired through the web server: rAggr (<http://raggr.usc.edu>), and only considered the proxies that meet the following criteria: population: EUR; minimum MAF (minor allele frequency):  $\geq 0.01$ ;  $r^2 \geq 0.8$ ; maximum distance: 500 kb; maximum # Mendel error: 1; Hardy-Weinberg equilibrium (HWE)  $p$ -value:  $1e-6$ ; and minimum genotyping rate of 95%. Cancer GWAS proxy SNPs and GWAS proxy SNPs were then pulled to intersect with the significant SNPs genotyped on the Infinium array as described below.

SNPs on the Infinium genotyping array with  $MAF \geq 0.05$  were selected based on dbSNP150 hg19 genome build. Briefly, snp150Common.txt were accessed in August 2018 from <http://hgdownload.soe.ucsc.edu/goldenPath/hg19/database/>. Triallelic SNPs were removed. SNPs with  $MAF \geq 0.05$  (8,209,080 SNPs) were then pulled to identify the number of SNPs genotyped on the Infinium array with  $MAF \geq 0.05$  among cancer GWAS and GWAS SNP lists.

### SNP analyses

The following methods were used to analyse the BRZ cohort. First, SNPs mapped to chromosome: 0 (unknown), XY (pseudoautosomal), Y, or MT (mitochondria) and SNPs mapped to chromosome position 0 were removed. Only females were considered for the SNPs on the X Chromosome. A MAF threshold was set at 0.05 (SNPs with  $MAF < 0.05$  were removed). Nonpolymorphic SNPs and duplicated SNPs (SNPs with two rsIDs at the same position) and SNPs that failed genotyping for all patients were removed. For each SNP, heterozygous and homozygous minor allele genotype calls were combined into the same group for statistical analysis, to overcome the missing genotypes and/or low patient numbers per genotype within these two small cohort sizes. A Log-ranked non-parametric test, the Mann–Whitney  $U$  test, two-sided, was used to determine statistical differences between genotype and age-of-tumour onset. The analysis was performed in base R using the Wilcoxon test. SNPs in the BRZ1 cohort with a raw  $p$ -value  $< 0.05$  in the Mann–Whitney  $U$  test were then likewise tested for significance in the BRZ2 cohort and examined for consistent effects. The subset of SNPs having a raw  $p$ -value  $< 0.05$  in the Mann–Whitney  $U$  test in both cohorts and showing a consistent effect were considered for further examination and validation in the adult ACC cohort.

For the primarily adult ACC cohort, reference SNPs rsIDs were extracted from the Affymetrix Genome-Wide Human SNP 6.0 Array names according to <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=%20GPL6801> (accessed July 2018). The primarily adult ACC cohort genotyping was on the Affymetrix SNP array, while the paediatric BRZ cohorts were genotyped using the Infinium LCG array. Thus, when testing significant SNPs (raw  $p < 0.05$ ) identified in both BRZ cohorts, we either utilised the genotypes for the same SNP, if directly genotyped SNP on the Affymetrix SNP array, or we

utilised the genotypes of a closely linked proxy SNP genotyped on the Affymetrix SNP array. Proxy SNPs were determined using rAggr (<http://raggr.usc.edu>) for the EUR population. If multiple proxies for a given significant SNP were genotyped in the adult ACC cohort, the most closely linked SNP was utilised. The heterozygous and homozygous minor allele genotypes were combined for statistical analyses if the number of minor homozygotes were  $\leq 4$ . A Log-ranked non-parametric test, two-sided, the Mann–Whitney  $U$  test (2 groups) or the Kruskal–Wallis test (three groups), was applied to determine statistical differences between genotype and age of tumour onset.  $P$ -values were subsequently adjusted for multiple hypothesis correction (Bonferroni correction).

#### mRNA expression levels and survival analyses

RNA-seq data for the adult ACC cohort (illumine hiseq Level 3 RSEM normalised genes) were retrieved from <http://firebrowse.org/?cohort=ACC> accessed February 2019. Genes involved in the regulation of retinoid metabolism and those involved in retinoic acid receptor-mediated signalling were retrieved from [http://software.broadinstitute.org/gsea/msigdb/cards/GO\\_RETINOIC\\_ACID\\_METABOLIC\\_PROCESS.html](http://software.broadinstitute.org/gsea/msigdb/cards/GO_RETINOIC_ACID_METABOLIC_PROCESS.html) (GO:0042573) and [http://software.broadinstitute.org/gsea/msigdb/cards/PID\\_RETINOIC\\_ACID\\_PATHWAY.html](http://software.broadinstitute.org/gsea/msigdb/cards/PID_RETINOIC_ACID_PATHWAY.html), accessed 18 July 2019. The cut-off point between high and low mRNA expression levels for each gene was computed based on the maximally selected rank statistics using the “survminer” package in R. mRNA levels with values above the cut-off point were considered high and below were considered low.

Survival data for the adult ACC were retrieved from Liu et al.,<sup>28</sup> Table S1 Tab TCGA-CDR, accessed May 2019. The progression-free interval (PFI) endpoint was analysed (PFI, the length of time during and after the treatment of a disease when a patient lives with the disease but it does not get worse). The patients in the adult ACC cohort were randomly separated into Discovery and Validation cohort pairs at 50:50 ratios 10 times using the “R Base” package.<sup>29</sup> We first performed Kaplan–Meier survival analyses to estimate the survival probability as a function of time in relation to mRNA expression levels for each transcript (patients with low versus high expression levels in tumours). The Log-Mantel and Haenszel survival test (a Log-rank test) was used to determine the statistical differences between the groups. Genes that were significant with  $p < 0.05$  were tested in the paired Validation cohort. In the Validation cohort the analyses were undertaken in the exact same manner except for the final adjustment of the  $p$ -values for multiple hypothesis testing, which was performed using the Bonferroni method. Adjusted  $p < 0.05$  were considered significant. These analyses were performed in each of the 10 Discovery/Validation pairs and those genes whose transcript levels associated with PFI six times or more were considered hits.

Gene expression data from two independent paediatric ACC cohorts were used in this study: the Children’s Oncology Group (COG,  $n = 34$ , accession number; GSE76019) and International Pediatric Adrenocortical Tumor Registry group (IPACTR,  $n = 19$ , accession number; GSE76021). The data from the COG and the IPACTR were accessed via the Gene Expression Omnibus (GEO) in June 2019. For these cohorts, the progression-free survival (PFS) endpoint was considered (PFS, the time elapsed between protocol enrolment and the date of disease progression, death, or most recent follow-up).<sup>30</sup> mRNA expression data was available as a normalised log<sub>2</sub> signal. The data were generated on two different array platforms: IPACTR on the GLP96 platform, COG on the GPL13158 platform. In order to maximise consistency in the analyses, we considered only probes with the same Transcript ID design. The cut-off point between high and low mRNA expression levels, the statistical difference in survival, and the approach used to validate the significant genes were performed as described above.

All analyses were performed on R (version 3.5.1) using the following packages “data.table”, “dplyr”, “reshape2”, “tidyr”, “plyr”, “survminer” and “survival” in addition to base R. Plots were generated using Prism 7 (GraphPad) and R using the “qqman”, “gplots” and “ggplot2” packages. All statistical tests were two-sided except where otherwise stated. Pathway diagram was created with BioRender.com.

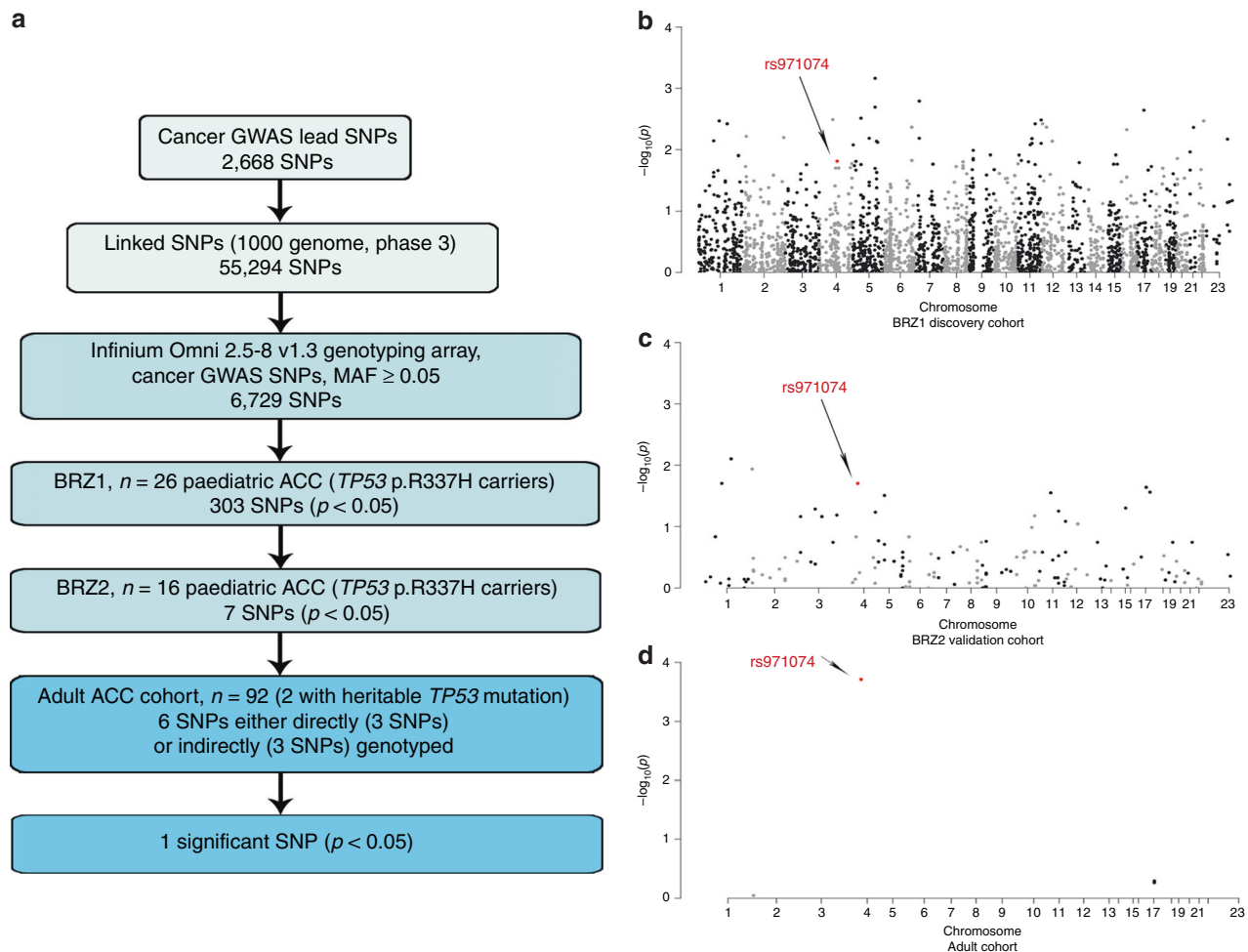
## RESULTS

In order to identify potential genetic modifiers underlying ACC development that could improve cancer risk management and treatment, we genotyped a total of 42 *TP53* R337H carriers with ACC, divided into a discovery cohort (BRZ1,  $N = 26$ ) and a validation cohort (BRZ2,  $N = 16$ ). All BRZ samples had a >98% genotype call rate, with an average overall call rate of 99.6%. The clinical characteristics of the paediatric patients of both cohorts are shown in Supplementary Table 1. All children (29 girls and 13 boys) were carriers of the R337H mutation and were diagnosed at ages ranging from 0.14 to 14.7 years. Family history showed that three children were from LFS families, 17 children were from Li-Fraumeni-Like (LFL) families and 22 with no LFS/LFL family histories (Supplementary Table 1) according to the most stringent criteria.<sup>31</sup>

#### Cancer GWAS SNPs and ACC age-of-onset in R337H mutation carries

The small cohort size and narrow age-of-onset window necessitated a focused analytical approach in the search for SNP modifiers. Therefore, we focused our search on SNPs that have been previously found to associate with differential cancer risk in GWAS (Fig. 1a). This process began by procuring the 2,668 GWAS lead SNPs in the GWAS catalog found to significantly associate with cancer risk among Europeans. In total, these SNPs were found to associate with differential risk of 33 different cancers. Next, proxy SNPs linked to these cancer GWAS tags were obtained. SNP linkage information is not available for the Brazilian population, but it is thought that the R337H mutation originates from a founder effect of European origin.<sup>3,4</sup> Therefore, we retrieved proxies that are linked in the EUR. This resulted in 55,294 cancer GWAS SNPs. Of the cancer GWAS SNPs, 6,729 SNPs were on the Infinium genotyping array with  $MAF \geq 0.05$ . 303 cancer GWAS SNPs showed significant allelic differences with ACC age-of-onset in BRZ1 ( $p < 0.05$ , Mann–Whitney  $U$  test). 282 SNPs were tested in BRZ2 ( $MAF \geq 0.05$  in BRZ2), of these, seven SNPs were also associated significantly ( $p < 0.05$ , Mann–Whitney  $U$  test) with differential age-of-onset in a consistent manner whereby the same allele associated with an earlier diagnosis (Fig. 1a).

Cancer GWAS SNPs and ACC age-of-onset in sporadic adult ACC  
The cohort of primarily adult ACC patients included 92 individuals (60 women and 32 men) diagnosed with a mean age of 47.16 years (range 14–83 years). Two patients were carriers of the R337H germline mutation (diagnosed at ages 23 and 30 years). For all patients, genotyping information was available from blood DNA (Affymetrix SNP 6.0 Array), while somatic mutation data from the tumours was available for 91 patients (Supplementary Table 2). Out of the seven SNPs that significantly associated with differential age-of-onset of paediatric ACC, six SNPs could be studied in the primarily adult ACC cohort, as three were successfully genotyped with the Affymetrix array, and three SNPs had successfully genotyped proxies. One of these six SNPs, rs971074, significantly associated with differential age-of-onset (adj.  $p = 0.00114$ , Kruskal–Wallis test). rs971074 (C/T, chr4:100341861, Genome Reference Consortium Human Build 37 (GRCh37)/hg19) is a synonymous SNP located in exon 6 of the alcohol dehydrogenase 7 gene (*ADH7*) gene (Fig. 1b–d). The *TP53* mutation carriers with rs971074 minor allele (CT) were diagnosed



**Fig. 1 Cancer GWAS approach to identify the genetic modifiers in ACC age-of-onset.** **a** A flow chart representing the focused analytical approach taken to identify potential cancer GWAS SNPs which associate with differential age of ACC onset. **b–d** Manhattan plot representing cancer GWAS SNPs tested in BRZ1 (discovery) cohort with p-value represented in  $-\log_{10}(p)$  for each SNP ( $n = 6059$ ) (**b**). Manhattan plot representing the significant cancer GWAS SNPs ( $p < 0.05$ ) in BRZ1 and their validation in the BRZ2 cohort ( $n = 282$ , out of 303 significant in BRZ1) (**c**). Manhattan plot representing significant the cancer GWAS SNPs in both BRZ1 and BRZ2 and validated in the adult ACC cohort ( $n = 6$ , out of 7 significant in BRZ1 and BRZ2) (**d**). The significant SNP in the 3 cohorts is marked in red and pointed with an arrow.

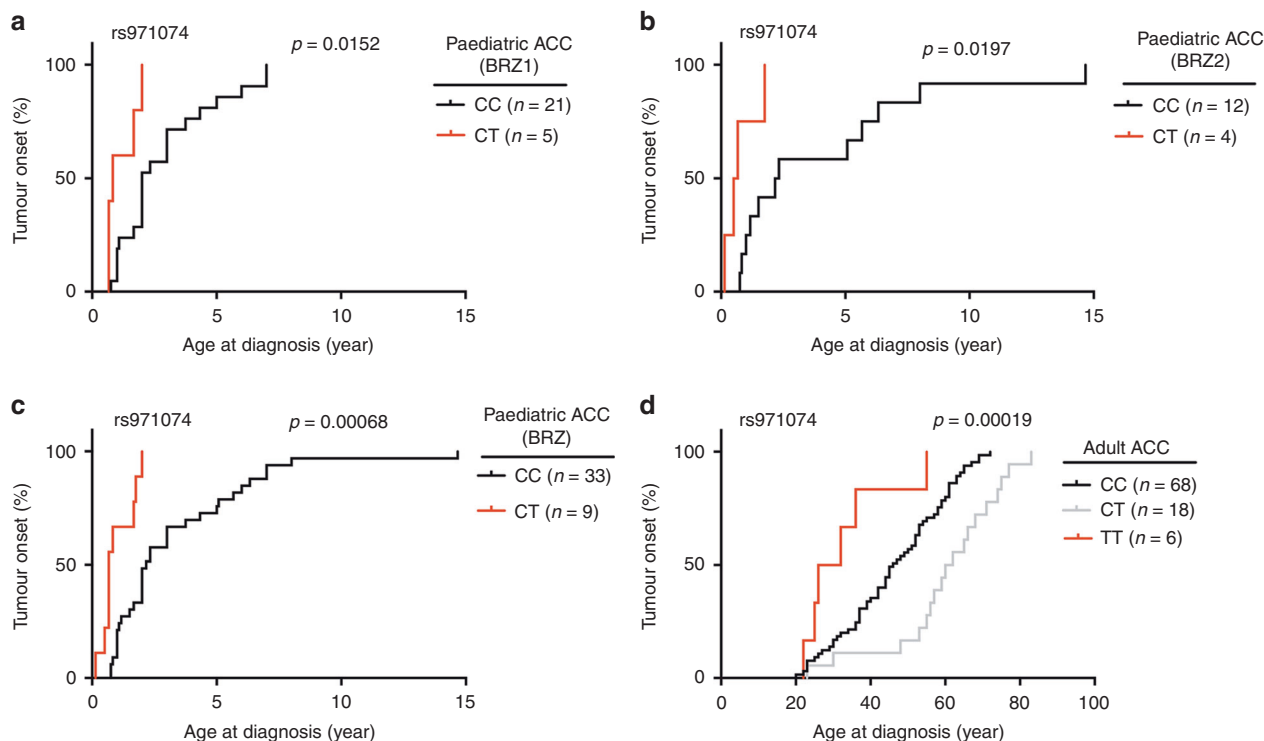
at a median age of 0.85 year in the BRZ1 and at a median of 0.585 year in BRZ2 cohort. However, carriers of the homozygous major allele (CC) were diagnosed at a median age of 2 years (BRZ1,  $p = 0.0153$ , Mann–Whitney  $U$  test), and at a median of 2.25 years (BRZ2,  $p = 0.0197$ , Mann–Whitney  $U$  test) Table 1, Fig. 2a, b. When the patients of BRZ1 and BRZ2 are combined (BRZ), CT patients were diagnosed at a median age of 0.67 years compared to a median age of 2.17 years for CC patients ( $p = 0.00068$ , Mann–Whitney  $U$  test) Table 1 and Fig. 2c.

In the primarily adult ACC cohort, CC patients were diagnosed 16 years earlier than TT patients (at a median age of 45 and 29 years, respectively.  $p = 0.0498$ , Mann–Whitney  $U$  test) Table 2 and Fig. 2d, in a manner similar to the paediatric cohort. In contrast to the paediatric cohort, patients heterozygous for rs971074 were diagnosed at older ages (16 years later than CC patients,  $p = 0.00038$ ; 32 years later than TT patients,  $p = 0.00414$ , Mann–Whitney  $U$  test).

Altogether, these data show that the rs971074 SNP associates with differential ACC age-of-onset in both *TP53* mutation carriers and non-carriers alike. However, in ACC patients with no inherited *TP53* mutations, the *TP53* gene is known to be somatically mutated (19.6%, cBioPortal). In the adult cohort, 18 patients (19.6%) had known pathogenic *TP53* mutations in their tumours,

Cohort		Patient numbers	Mean (years)	Median (years)
BRZ1	Genotype Major (CC)	21	2.9	2
	Minor (CT)	5	1.168	0.83
	MAF	0.0962		
	p-value (MWU)	0.0153		
BRZ2	Genotype Major (CC)	12	4.13	2.25
	Minor (CT)	4	0.765	0.585
	MAF	0.125		
	p-value (MWU)	0.0197		
BRZ	Genotype Major (CC)	33	3.345	2.17
	Minor (CT)	9	0.999	0.67
	MAF	0.1071		
	p-value (MWU)	0.000068		

MAF minor allele frequency, MWU Mann–Whitney  $U$  test.



**Fig. 2** rs971074 SNP modifies the age of ACC onset. Kaplan–Meier plots representing the age of tumour onset for both the major allele homozygous (CC) and the minor allele heterozygous (TC) and homozygous (TT). In **a** the ages of onset of the first paediatric ACC cohort is plotted (BRZ1) ( $p = 0.0152$ , two-sided Mann–Whitney  $U$  test). In **b** the second paediatric cohort is plotted (BRZ2) ( $p = 0.0197$ , two-sided Mann–Whitney  $U$  test). In **c** the paediatric cohort combined (BRZ) is plotted ( $p = 0.00063$ , two-sided Mann–Whitney  $U$  test). In **d** the patients in the primarily adult ACC cohort are plotted ( $n = 92$ ),  $p = 0.00019$ , Kruskal–Wallis test.

Cohort		Patient number	Mean (years)	Median (years)	
Adult ACC	Genotype	CC <sup>a</sup>	68	45.01	45
		CT <sup>b</sup>	18	60.1	61
		TT <sup>c</sup>	6	32.67	29
	MAF		0.163		
	$p$ -value (KW)		0.00019		
	adjusted $p$ -value (Bonferroni)		0.00114		
	$p$ -value (MWU) (CC vs TT)		0.0498		
	$p$ -value (MWU) (CC vs CT)		0.00038		
	$p$ -value (MWU) (TT vs CT)		0.0041		
	Adult ACC (somatic wild type p53)	Genotype	CC	50	45.34
	CT	16	60.19	61	
	TT	5	32	26	
	MAF		0.1831		
	$p$ -value (KW)		0.00075		

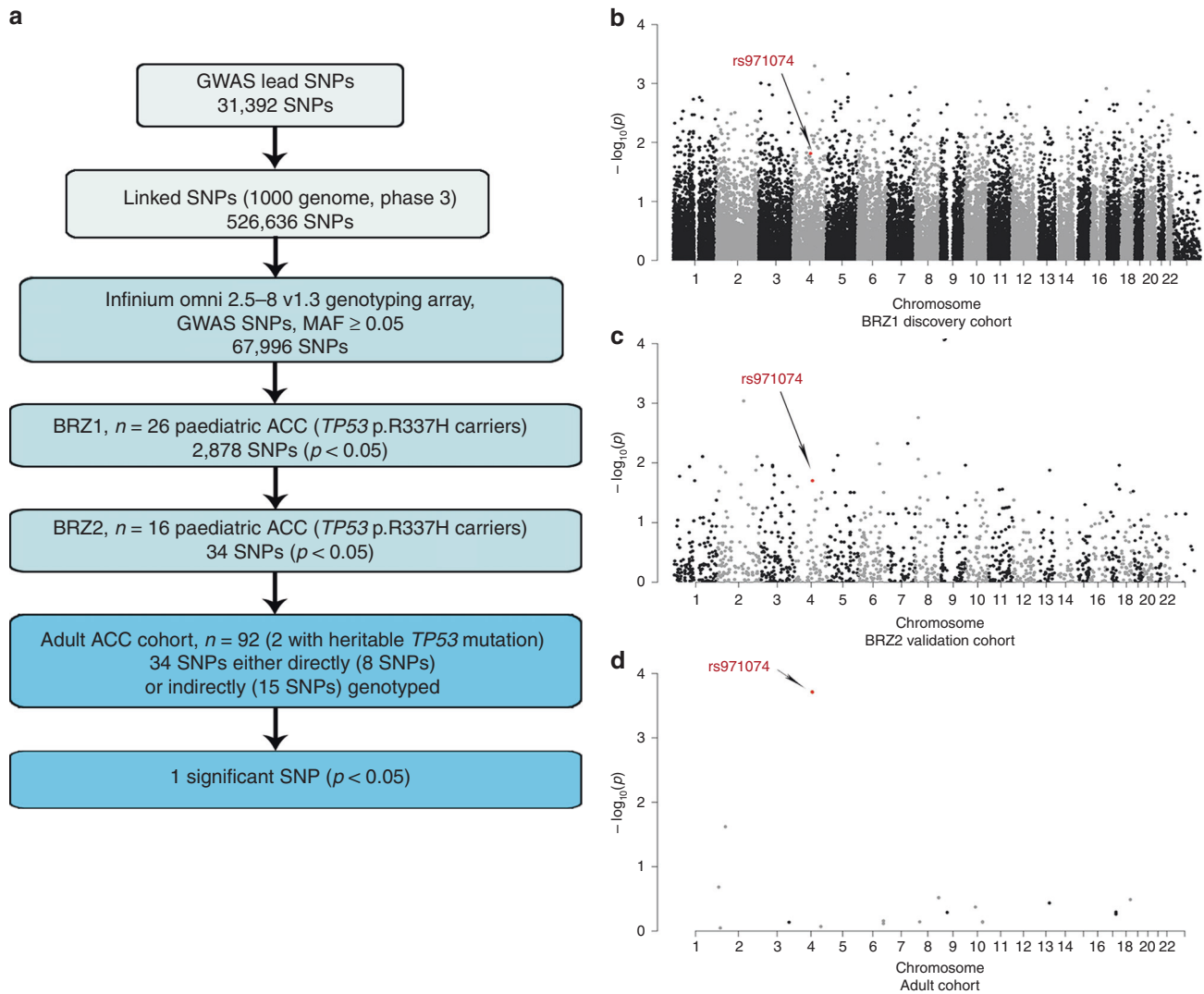
MAF minor allele frequency, KW Kruskal–Wallis, MWU Mann–Whitney  $U$  test.  
<sup>a</sup>CC: Homozygous major.  
<sup>b</sup>CT: Heterozygous.  
<sup>c</sup>Homozygous minor.

and 2 had the R337H germline mutation. When we restrict our analyses to those 71 patients without either inherited or somatic *TP53* mutations, and excluding the patient with no somatic mutation data, similar allelic differences in the age-of-onset were found ( $p = 0.00075$ , Kruskal–Wallis test, Table 2 and Supplementary Fig. 1), wherein TT patients were diagnosed at a median age of 26 years compared to a median age of 46.5 years for the CC patients. This corresponds to a 20.5 year earlier diagnosis for TT patients demonstrating the apparent *TP53* mutation-independence of these associations.

It is important to note that in a smaller paediatric ACC cohort ( $N = 21$ , The Hospital for Sick Children, Canada) of *TP53* mutation carriers with more penetrant mutations and diagnosed at ages ranging from 1 to 17.83 years (Supplementary Table 3), no significant difference in ACC age-of-onset with the rs971074 SNP genotype was observed (Supplementary Table 4).

#### GWAS SNPs and ACC age-of-onset

The rs971074 SNP was included in our list of cancer GWAS SNPs due to its association with differential cancer risk in the upper aerodigestive tract, particularly oesophageal cancers.<sup>32,33</sup> To further explore this association in ACC, we included GWAS SNPs known to associate not only with cancer, but also with all other diseases and traits that have been studied in GWAS. Specifically, in total, 31,392 lead SNPs can be found in the GWAS catalog. For these, we identified 526,636 closely linked proxy SNPs. Of these, 67,996 SNPs with  $MAF \geq 0.05$  were present on the genotyping array. Among these, 2,878 SNPs were associated with the ACC age-of-onset in BRZ1 ( $p < 0.05$ , Mann–Whitney  $U$  test). 2,644 SNPs were tested in BRZ2 ( $MAF \geq 0.05$ ), of these, 34 SNPs were also found to be associated with age-of-onset in a consistent manner ( $p < 0.05$ , Mann–Whitney  $U$  test). In the adult ACC cohort, we were able to explore potential associations with differential age-of-onset for 23



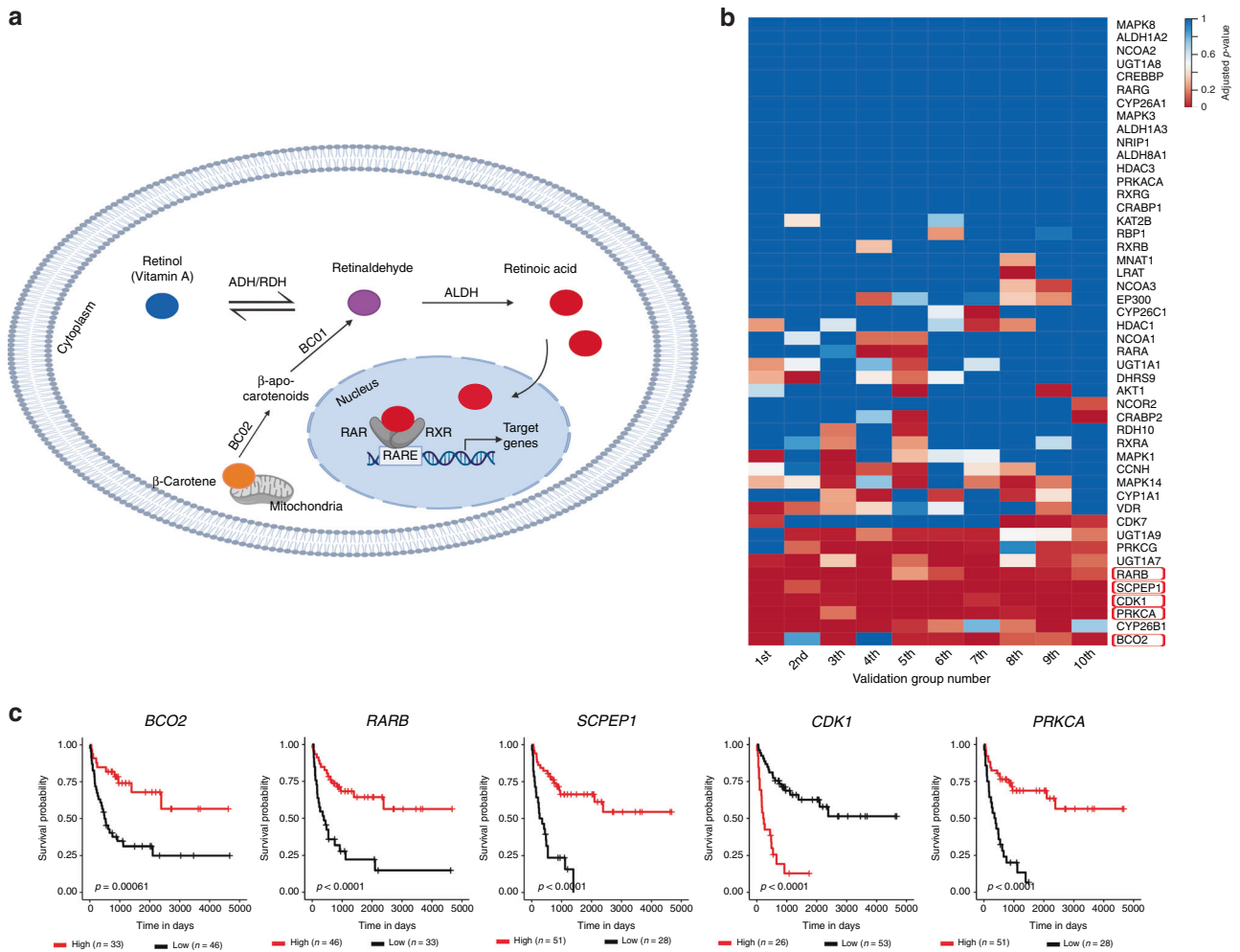
**Fig. 3 GWAS approach to identify the genetic modifiers in ACC age-of-onset.** **a** A flow chart representing the focused analytical approach taken to identify potential GWAS SNPs, which associate with differential age of ACC onset. **b–d** Manhattan plot representing GWAS SNPs tested in BRZ1 (discovery) cohort with p-value represented in  $-\log_{10}(p)$  for each SNP ( $n = 61,609$ ) (**b**). Manhattan plot representing the significant cancer GWAS SNPs ( $p < 0.05$ ) in BRZ1 and their validation in the BRZ2 cohort ( $n = 2,644$ , out of 2,878 significant in BRZ1) (**c**). Manhattan plot representing the significant cancer GWAS SNPs in both BRZ1, BRZ2 and validated in the adult ACC cohort ( $n = 23$ , out of 34 significant in BRZ1 and BRZ2) (**d**). The significant SNP in the three cohorts is marked in red and pointed with an arrow.

of the 34 SNPs (8 SNPs directly and 15 SNPs through proxies) Fig. 3a–d. Interestingly, only the rs971074 SNP was significant after multiple hypothesis correction (adj.  $p = 0.0044$ , Kruskal–Wallis test), thereby underlining the importance of the GWAS SNP's association with cancer and not other diseases and traits.

Expression levels of RA pathway genes in ACC associate with tumour progression

ADH7 is involved in the production of RA from retinol (also known as vitamin A), which can induce differentiation and inhibit cellular growth (Fig. 4a). This pathway is therapeutically malleable and RA isomers are utilised in both cancer prevention and treatment strategies.<sup>34,35</sup> Our data thus far suggest that the RA pathway could be important in ACC. To further test this hypothesis, we explored the association of RA pathway gene expression with ACC progression and survival. We focused our analyses on 50 RA pathway genes, whereby 30 genes are known to be involved in RA signalling and 20 genes in RA metabolism. In the adult ACC cohort, we processed RNA-seq data derived from tumours. We were able to detect significant RNA levels of 48 of the 50 genes.

Next, we randomly assigned the patients into Discovery and Validation groups at a 50:50 ratio. We repeated these analyses ten times. For each Discovery group, we first separated the patients into two groups based on the expression levels of a given RA pathway gene using a Rank statistic (high and low expressing groups). We then explored potential differences in PFI between these two groups using a Log-rank test. In the patients of the Discovery cohorts, we identified an average of 28 transcripts (ranging from 20 to 33) that were associated with PFI with a raw  $p < 0.05$ . Next, we sought validation of the associations identified in the Discovery groups, in the paired Validation groups. In the Validation groups, the raw-p-values for each association were adjusted for multiple hypothesis testing and significance was considered with an adjusted  $p < 0.05$  (Bonferroni adjustment, Supplementary Table 5). Interestingly, expression levels of 5 different genes consistently associated with PFI in minimally 6 different Discovery/Validation cohort pairs (Fig. 4b). One of the genes (*BCO2*) is involved in RA metabolism, while the other 4 genes (*RARB*, *SCPEP1*, *CDK1* and *PRKCA*) are involved in RA signalling (Fig. 4c).



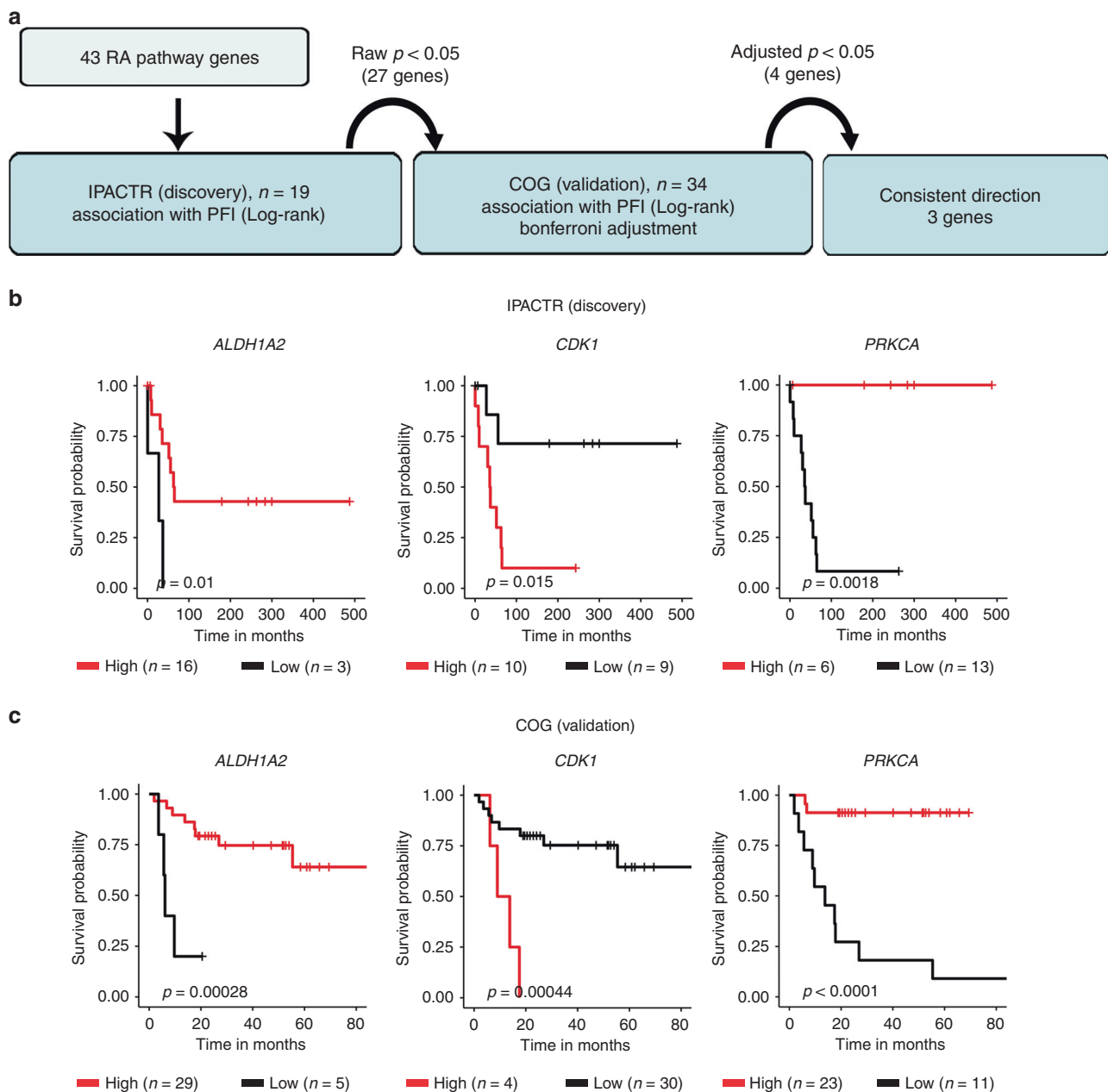
**Fig. 4 Expression levels of RA pathway genes in ACC associate with progression in adult patients.** **a** A depiction of the RA signalling pathway. Retinol (Vitamin A) is reversibly oxidised to retinaldehyde by the alcohol dehydrogenase (ADH) or retinol dehydrogenase (RDH) enzymes. Retinaldehyde is irreversibly oxidised to retinoic acid (RA), the active form of retinol, by the action of aldehyde dehydrogenases (ALDH) family members, also known as retinaldehyde dehydrogenases (RALDH). Retinaldehyde can also be produced from  $\beta$ -carotene. Beta-carotene oxygenase 2 (BCO2) asymmetrically cleaves  $\beta$ -carotene to generate  $\beta$ -apo-10'-carotenal, which is then converted to retinaldehyde, by BCO1. RA is the ligand for the nuclear receptors, retinoic acid receptor (RAR) and retinoid X receptor (RXR). RA regulates the transcription of its target genes following its binding to the RAR/RXR-retinoic acid responsive element (RARE) complex. **b** A heatmap representing the level of significance of the association of the transcript levels of RA pathway genes with progression-free interval of ACC patients. The results for the transcripts for all 48 genes are presented for the 10 Validation groups. *P*-values represent the Bonferroni adjusted *p*-values and considered significant if  $p < 0.05$ . Genes were considered hits if they were found to be significantly associated with ACC in six or more of the Validation groups and are boxed in red. **c** Kaplan–Meier survival curves for the significant hits. The plots were derived using all adult ACC patients ( $n = 79$ ). Statistical analyses were performed using the Log-rank test and the *p*-values are indicated on the graphs.

These data implicate a role for RA pathway genes in ACC progression. To further test this hypothesis, we next explored similar associations in paediatric ACC. We utilised two independent paediatric ACC cohorts with microarray gene expression data associated with outcome information (IPACTR and COG).<sup>30</sup> We investigated the expression data from histologically confirmed ACC paediatric patients (19 from the IPACTR and 34 from the COG).<sup>30</sup> We analysed these cohorts in a similar manner as with the adult ACC cohort, and thus randomly assigned the IPACTR as the Discovery cohort and the COG as the Validation cohort, and explored potential associations of gene expression for the 50 RA pathway genes with PFS using a Log-rank test. Specifically, as the cohorts utilised different arrays, we restricted our analyses to shared probes to maximise our ability to draw comparisons. Out of the 50 RA pathway genes, we were able to identify 79 shared probes for 43 RA pathway genes. Using these data and the above-described analytical pipeline, we determined that the expression

levels of 4 different transcripts from three different genes associated with PFS in both cohorts: Discovery cohort: raw  $p < 0.05$ ; Validation cohort: adjusted  $p < 0.05$ . One of the genes (*ALDH1A2*) is involved in RA metabolism, while the other two genes (*CDK1*, and *PRKCA*) are involved in RA signalling ( $p < 0.05$ , Bonferroni adjustment, Fig. 5a–c, Supplementary Table 6).

## DISCUSSION

In this report, we explored whether SNPs could potentially modify ACC cancer risk and thus give us a better understanding of ACC development and possible indications of novel treatments. Due to the fact that this is a rare cancer, we undertook a focused analytical approach, whereby we restricted our analyses to SNPs previously identified in GWAS studies to associate with disease or other human traits. We utilised two independent cohorts of ACC patients and clearly demonstrated that a SNP (rs971074 in *ADH7*)



**Fig. 5 Expression levels of RA pathway genes in ACC associate with progression in paediatric patients.** **a** A flow chart representing the analytical approach taken to identify the RA pathway genes that are associated with paediatric ACC progression. 17 genes showed an association with PFS in the Discovery cohort (Log-rank test, raw  $p < 0.05$ ), among these 17 genes, 3 genes showed a significant association with PFS in the Validation cohort (Log-rank test, Bonferroni adjusted  $p < 0.05$ ) in consistent manner. **b, c** Kaplan–Meier survival curves showing the RA pathway genes that are associated with PFS in paediatric ACC patients. Statistical analysis was performed using the Log-rank test and  $p$ -values are indicated on the graphs.

in a druggable pathway significantly associated with the age of ACC onset, suggesting an increased risk associated with ACC age-of-onset. We demonstrated that the minor allele of rs971074 in *ADH7* associates with a significantly earlier age of paediatric ACC in R337H carriers (CT) and adult ACC (TT), relative to patients homozygous for the major allele (CC) in a manner that is seemingly independent of both inherited and somatic *TP53* mutations. It is important to note that although the associations with the age-of-onset for the different genotypes were significant in both the paediatric and adult ACC cohorts, there are some differences. For example, in the adult cohort, the heterozygous patients (CT) demonstrated the latest average age-of-onset, while

in the paediatric, the heterozygotes were the youngest. It will be important to validate the associations in a second adult cohort, but this could further reflect the known genetic differences in the development of paediatric and adult ACC.<sup>36,37</sup> For instance, amplification of chromosome 9q occurs in 90% of paediatric ACC tumours but not in adult tumours.<sup>36</sup> Moreover, although upregulation of the Wnt/ $\beta$ -catenin signalling is common in paediatric and adult ACC, mutations in *ZNRF3* are only observed in adult tumours.<sup>36,37</sup>

We went on to demonstrate that amongst all SNPs identified to associate with human diseases or traits, only this SNP significantly and reproducibly associated with ACC age-of-onset. Intriguingly,



this SNP was shown to associate in GWAS with differential risk for the development of cancer, namely with upper aerodigestive tract cancer risk.<sup>32,33</sup> The rs971074 SNP is a synonymous SNP found in exon 6 of *ADH7*. *ADH7* is important in retinoic acid (RA) signalling as it is crucial for the production of RA from retinol (also known as vitamin A).<sup>38</sup> RA and its isomers such as all-*trans*-RA (ATRA), 9-*cis* RA, and 11-*cis* RA<sup>39</sup> are involved in the regulation of different genes that are important for cell growth, differentiation, and apoptosis by activating the nuclear retinoid receptors, RA receptors (RARs) and retinoid X receptors (RXRs).<sup>40</sup> The importance of *ADH7* and RA in the adrenal gland has also been clearly shown in mammals. In mice, class IV ADH (*Adh4*, recently re-named *Adh7*, <http://www.informatics.jax.org/searchtool/Search.do?query=adh&submit=Quick%0D%0ASearch>) and RA are expressed in the embryonic adrenal glands and adult adrenal cortex.<sup>41,42</sup> In fact, class IV ADH null mice are normal and fertile, with no growth defects when maintained on a normal diet. However, they show 100% lethality when placed on a vitamin A-deficient diet.<sup>38,43</sup> These data suggest that *ADH7* is important in producing RA during development when vitamin A levels are low. Moreover, RA signalling may be involved in the regulation of immune cells. For example, reduction of primary murine B cell adhesion was observed after removing retinoids or by treatment with corticosteroids.<sup>44</sup> These effects could suggest a putative RA-mediated activation of the immune system against ACC, a role that could be more important in paediatric and adult patients with and without Cushing syndrome.<sup>45</sup>

Although the effects of this SNP on RA signalling and ACC development remain to be further explored in mechanistic studies, we have been able to provide further evidence for the importance of the RA pathway genes in ACC. Specifically, using a two-step validation strategy to identify transcripts that significantly and reproducibly associate with ACC progression, we clearly demonstrated that transcripts for multiple RA pathway genes associate with differential progression in both adult (five genes) and paediatric (three genes) ACC. In both groups, we identified genes clearly involved in RA metabolism in a manner that could suggest that lower RA production associates with faster ACC progression. Specifically, in the adult ACC cohort, patients with low levels of *beta-carotene oxygenase 2* (*BCO2*) transcript associated with shorter PFI. *BCO2* is a mitochondrial enzyme involved in the production of retinoids from provitamin A carotenoids, which is a key source of retinoids in our diet.<sup>46</sup> Specifically, *BCO2* asymmetrically cleaves  $\beta$ -carotene to generate  $\beta$ -apo-10'-carotenal, which is then converted to retinoids by *BCO1*.<sup>47,48</sup> In the paediatric ACC cohorts, shorter PFS was seen in patients with low transcript levels of *Aldehyde Dehydrogenase 1 Family Member A2* (*ALDH1A2*). *ALDH1A2* is an enzyme involved in the rate-limiting step of RA production from retinaldehyde.<sup>49,50</sup> Consistent with the hypothesis that less RA signalling results in faster ACC progression is the association of low mRNA levels for the *RA receptor beta* (*RARB*) with faster progression of adult ACC patients. *RARB* is a member of the RAR family that binds the retinoid ATRA to further activate down-stream RA signalling.<sup>51</sup>

This is not the first time that lower expression levels of the RA pathway genes have been implicated in ACC.<sup>52</sup> In a meta-analysis of gene expression data from 164 adult adrenal tumours (97 adenomas and 67 carcinomas), significant reductions of three RA pathway transcripts were noted in malignant ACC relative to benign adrenal tumours. Specifically, *ALDH1A1*, *ALDH1A3*, and *RXRA* mRNAs were much lower in ACCs relative to adrenocortical adenoma.<sup>52</sup> The authors, therefore, hypothesised that RA production and signalling might be reduced in ACC and, therefore, its re-activation could be a therapeutic option,<sup>52</sup> as retinoids are already used in cancer treatment for acute promyelocytic leukaemia, and in cancer prevention for precancerous lesions in leukoplakia, actinic

keratosis, and cervical dysplasia.<sup>53</sup> Support of this came from experiments in both cell culture and murine xenografts, whereby treatment of ACC with 9-*cis* RA resulted in decreased cell viability, proliferation, and steroid hormone production.<sup>54,55</sup> Our results offer human genetic evidence for the importance of this therapeutically relevant pathway in ACC onset and extends its relevance to progression of both paediatric and adult ACC. It is intriguing to note that a recent study of metastatic ACC tumours has identified an enrichment of driver mutations in RA pathway genes, possibly lending further support to the inactivation of this pathway being a crucial step in ACC progression.<sup>56</sup> In addition, alterations in the retinoic acid pathway have also been linked recently to benign adrenocortical tumours secreting aldosterone.<sup>57</sup>

In conclusion, our results offer human genetic evidence for the importance of the druggable RA pathway in ACC onset and extend its relevance to progression of both paediatric and adult ACC. These findings will impact the design of future studies using RA in pre-clinical studies and guide preventive actions following genetic counselling and testing for children carriers of R337H germline mutation.

## ACKNOWLEDGEMENTS

We thank the High-Throughput Genomics Group at the Wellcome Trust Centre for Human Genetics (funded by Wellcome Trust grant reference 090532/Z/09/Z) for the generation of genotyping data. We thank Isaac Kitchen-Smith for reviewing the paper.

## AUTHOR CONTRIBUTIONS

Conception and design: M.S., M.W., B.C.F., G.B. Collection and assembly of data: M.S., M. W., E.B., P.Z., I.Z.S.P., H.K., M.M.P., D.M., E.M.P., E.L., B.C.F., G.B. Data analysis and interpretation: M.S., M.W., G.L.P., H.P., M.D.G., H.C., M.S.P., E.L., G.Z., D.M., E.M.P., B.C.F., G.B. Paper writing: M.S., B.C.F., G.B. Critical revising and final approval of paper: all authors.

## ADDITIONAL INFORMATION

**Ethics approval and consent to participate** The study was done in accordance with the principles of the Declaration of Helsinki. For the Brazilian paediatric ACC cohort, approval was given by the Ethics Committee from Pequeno Príncipe Hospital (CAA 0023.0.208.000-05 from 2005, CAAE 0612.0.015.000-08, 2009, and CAAE 50622315.0.0000.0097, 2015), and authorisation was obtained by signing a consent form. Approval from the NIH Data Access Committee was obtained (request number 34594-5) for the adult ACC cohort. For the Canadian cohort, approval was given by the Research Ethics Board of The Hospital for Sick Children (number 0019910602).

**Consent to publish** Consent for publication was approved by the Ethics Committee.

**Data availability** The websites used in the analysis for this study are publicly available websites: <http://www.cbioportal.org/>, <https://www.ebi.ac.uk/gwas>, <http://raggr.usc.edu> <http://hgdownload.soe.ucsc.edu/goldenPath/hg19/database/>, <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=%20GPL6801>, <http://software.broadinstitute.org/gsea/msigdb>, <https://www.ncbi.nlm.nih.gov/gds>, <http://firebrowse.org/>. The data generated from the analysis during this study are included in this published article and its supplementary information files. Genotyping data for the Cancer Genome Atlas participants requires approval from the NIH.

**Competing interests** The authors declare no competing interests.

**Funding information** This work was funded by the Ludwig Institute for Cancer Research, Conselho Nacional de Desenvolvimento e Pesquisa (CNPq), Comissão de Pessoal de Ensino Superior (CAPES) and the CNRS *EXPOGEN-CANCER*. International Associated Laboratory (LIA).

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41416-020-0764-3>.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

1. Wasserman, J. D., Zambetti, G. P. & Malkin, D. Towards an understanding of the role of p53 in adrenocortical carcinogenesis. *Mol. Cell Endocrinol.* **351**, 101–110 (2012).
2. Ribeiro, R. C. & Figueiredo, B. Childhood adrenocortical tumours. *Eur. J. Cancer* **40**, 1117–1126 (2004).
3. Pinto, E. M., Billerbeck, A. E., Villares, M. C., Domenice, S., Mendonca, B. B. & Latronico, A. C. Founder effect for the highly prevalent R337H mutation of tumor suppressor p53 in Brazilian patients with adrenocortical tumors. *Arq. Bras. Endocrinol. Metab.* **48**, 647–650 (2004).
4. Garritano, S., Gemignani, F., Palmero, E. I., Olivier, M., Martel-Planche, G., Le Calvez-Kelm, F. et al. Detailed haplotype analysis at the TP53 locus in p.R337H mutation carriers in the population of Southern Brazil: evidence for a founder effect. *Hum. Mutat.* **31**, 143–150 (2010).
5. Custodio, G., Parise, G. A., Kiesel Filho, N., Komechen, H., Sabbaga, C. C., Rosati, R. et al. Impact of neonatal screening and surveillance for the TP53 R337H mutation on early detection of childhood adrenocortical tumors. *J. Clin. Oncol.* **31**, 2619–2626 (2013).
6. Caminha I. P. Prevalence of TP53 germline mutation p.R337H in the metropolitan region of Campinas and surrounding cities. Campinas-SP. PhD thesis. (Universidade Estadual de Campinas (UNICAMP), Campinas-SP, Brasil, 2015)
7. Costa, T. E. J., Gerber, V. K. Q., Ibanez, H. C., Melanda, V. S., Parise, I. Z. S., Watanabe, F. M. et al. Penetrance of the TP53 R337H mutation and pediatric adrenocortical carcinoma incidence associated with environmental influences in a 12-year observational cohort in Southern Brazil. *Cancers (Basel)* **11**, E1804 (2019). pii.
8. Figueiredo, B. C., Sandrini, R., Zambetti, G. P., Pereira, R. M., Cheng, C., Liu, W. et al. Penetrance of adrenocortical tumours associated with the germline TP53 R337H mutation. *J. Med Genet* **43**, 91–96 (2006).
9. Ribeiro, R. C., Sandrini, F., Figueiredo, B., Zambetti, G. P., Michalkiewicz, E., Lafferty, A. R. et al. An inherited p53 mutation that contributes in a tissue-specific manner to pediatric adrenal cortical carcinoma. *Proc. Natl Acad. Sci. USA* **98**, 9330–9335 (2001).
10. Latronico, A. C., Pinto, E. M., Domenice, S., Fragoso, M. C., Martin, R. M., Zerbini, M. C. et al. An inherited mutation outside the highly conserved DNA-binding domain of the p53 tumor suppressor protein in children and adults with sporadic adrenocortical tumors. *J. Clin. Endocrinol. Metab.* **86**, 4970–4973 (2001).
11. Ribeiro, R. C., Pinto, E. M., Zambetti, G. P. & Rodriguez-Galindo, C. The International pediatric adrenocortical tumor registry initiative: contributions to clinical, biological, and treatment advances in pediatric adrenocortical tumors. *Mol. Cell Endocrinol.* **351**, 37–43 (2012).
12. Parise, I. Z. S., Parise, G. A., Noronha, L., Surakhy, M., Woiski, T. D., Silva, D. B. et al. The Prognostic Role of CD8(+) T lymphocytes in childhood adrenocortical carcinomas compared to Ki-67, PD-1, PD-L1, and the Weiss Score. *Cancers (Basel)* **11**, E1730 (2019).
13. Berruti, A., Fassnacht, M., Baudin, E., Hammer, G., Haak, H., Leboulleux, S. et al. Adjuvant therapy in patients with adrenocortical carcinoma: a position of an international panel. *J. Clin. Oncol.* **28**, e401–e402 (2010).
14. Fassnacht, M., Terzolo, M., Alolio, B., Baudin, E., Haak, H., Berruti, A. et al. Combination chemotherapy in advanced adrenocortical carcinoma. *N. Engl. J. Med* **366**, 2189–2197 (2012).
15. Finan, C., Gaulton, A., Kruger, F. A., Lumbers, R. T., Shah, T., Engmann, J. et al. The druggable genome and support for target identification and validation in drug development. *Sci. Transl. Med* **9**, eaag1166 (2017).
16. Nelson, M. R., Tipney, H., Painter, J. L., Shen, J., Nicoletti, P., Shen, Y. et al. The support of human genetic evidence for approved drug indications. *Nat. Genet.* **47**, 856–860 (2015).
17. Sanseau, P., Agarwal, P., Barnes, M. R., Pastinen, T., Richards, J. B., Cardon, L. R. et al. Use of genome-wide association studies for drug repositioning. *Nat. Biotechnol.* **30**, 317–320 (2012).
18. Stracquadanio, G., Wang, X., Wallace, M. D., Grawenda, A. M., Zhang, P., Hewitt, J. et al. The importance of p53 pathway genetics in inherited and somatic cancer genomes. *Nat. Rev. Cancer* **16**, 251–265 (2016).
19. Spencer, C. C., Su, Z., Donnelly, P. & Marchini, J. Designing genome-wide association studies: sample size, power, imputation, and the choice of genotyping chip. *PLoS Genet.* **5**, e1000477 (2009).
20. Visscher, P. M., Wray, N. R., Zhang, Q., Sklar, P., McCarthy, M. I., Brown, M. A. et al. 10 Years of GWAS discovery: biology, function, and translation. *Am. J. Hum. Genet* **101**, 5–22 (2017).
21. Villani, A., Shore, A., Wasserman, J. D., Stephens, D., Kim, R. H., Druker, H. et al. Biochemical and imaging surveillance in germline TP53 mutation carriers with Li-Fraumeni syndrome: 11 year follow-up of a prospective observational study. *Lancet Oncol.* **17**, 1295–1305 (2016).
22. Bond, G. L., Hu, W., Bond, E. E., Robins, H., Lutzker, S. G., Arva, N. C. et al. A single nucleotide polymorphism in the MDM2 promoter attenuates the p53 tumor suppressor pathway and accelerates tumor formation in humans. *Cell* **119**, 591–602 (2004).
23. Bougeard, G., Baert-Desurmont, S., Tournier, I., Vasseur, S., Martin, C., Brugieres, L. et al. Impact of the MDM2 SNP309 and p53 Arg72Pro polymorphism on age of tumour onset in Li-Fraumeni syndrome. *J. Med. Genet.* **43**, 531–533 (2006).
24. Ruijs, M. W., Schmidt, M. K., Nevanlinna, H., Tommiska, J., Aittomaki, K., Pruntel, R. et al. The single-nucleotide polymorphism 309 in the MDM2 gene contributes to the Li-Fraumeni syndrome and related phenotypes. *Eur. J. Hum. Genet* **15**, 110–114 (2007).
25. Bond, G. L., Hirshfield, K. M., Kirchhoff, T., Alexe, G., Bond, E. E., Robins, H. et al. MDM2 SNP309 accelerates tumor formation in a gender-specific and hormone-dependent manner. *Cancer Res.* **66**, 5104–5110 (2006).
26. Zhang, X., Pageon, L. & Post, S. M. Impact of the Mdm2(SNP309-G) allele on a murine model of colorectal cancer. *Oncogene* **34**, 4412–4420 (2015).
27. Zheng, S., Cherniack, A. D., Dewal, N., Moffitt, R. A., Danilova, L., Murray, B. A. et al. Comprehensive pan-genomic characterization of adrenocortical carcinoma. *Cancer Cell.* **29**, 723–736 (2016).
28. Liu, J., Lichtenberg, T., Hoadley, K. A., Poisson, L. M., Lazar, A. J., Cherniack, A. D. et al. An integrated TCGA pan-cancer clinical data resource to drive high-quality survival outcome analytics. *Cell* **173**, 400–16 e11 (2018).
29. Kang, G., Liu, W., Cheng, C., Wilson, C. L., Neale, G., Yang, J. J. et al. Evaluation of a two-step iterative resampling procedure for internal validation of genome-wide association studies. *J. Hum. Genet.* **60**, 729–738 (2015).
30. Pinto, E. M., Rodriguez-Galindo, C., Choi, J. K., Pounds, S., Liu, Z., Neale, G. et al. Prognostic significance of major histocompatibility complex class II expression in pediatric adrenocortical tumors: A St. Jude and Children’s Oncology Group Study. *Clin. Cancer Res* **22**, 6247–6255 (2016).
31. Birch, J. M., Hartley, A. L., Tricker, K. J., Prosser, J., Condie, A., Kelsey, A. M. et al. Prevalence and diversity of constitutional mutations in the p53 gene among 21 Li-Fraumeni families. *Cancer Res.* **54**, 1298–1304 (1994).
32. McKay, J. D., Truong, T., Gaborieau, V., Chabrier, A., Chuang, S. C., Byrnes, G. et al. A genome-wide association study of upper aerodigestive tract cancers conducted within the INHANCE consortium. *PLoS Genet.* **7**, e1001333 (2011).
33. Hashibe, M., McKay, J. D., Curado, M. P., Oliveira, J. C., Koifman, S., Koifman, R. et al. Multiple ADH genes are associated with upper aerodigestive cancers. *Nat. Genet.* **40**, 707–709 (2008).
34. Moreb, J. S., Ucar-Bilyeu, D. A. & Khan, A. Use of retinoic acid/aldehyde dehydrogenase pathway as potential targeted therapy against cancer stem cells. *Cancer Chemother. Pharm.* **79**, 295–301 (2017).
35. Uray, I. P., Dmitrovsky, E. & Brown, P. H. Retinoids and rexinoids in cancer prevention: from laboratory to clinic. *Semin. Oncol.* **43**, 49–64 (2016).
36. Pinto, E. M., Chen, X., Easton, J., Finkelstein, D., Liu, Z., Pounds, S. et al. Genomic landscape of paediatric adrenocortical tumours. *Nat. Commun.* **6**, 6302 (2015).
37. Lalli, E. & Figueiredo, B. C. Pediatric adrenocortical tumors: what they can tell us on adrenal development and comparison with adult adrenal tumors. *Front Endocrinol.* **6**, 23 (2015).
38. Deltour, L., Foglio, M. H. & Duester, G. Impaired retinol utilization in Adh4 alcohol dehydrogenase mutant mice. *Dev. Genet.* **25**, 1–10 (1999).
39. Cunningham, T. J. & Duester, G. Mechanisms of retinoic acid signalling and its roles in organ and limb development. *Nat. Rev. Mol. Cell Biol.* **16**, 110–123 (2015).
40. Rochette-Egly, C. & Germain, P. Dynamic and combinatorial control of gene expression by nuclear retinoic acid receptors (RARs). *Nucl. Recept. Signal* **7**, e005 (2009).
41. Ang, H. L., Deltour, L., Zgombic-Knight, M., Wagner, M. A. & Duester, G. Expression patterns of class I and class IV alcohol dehydrogenase genes in developing epithelia suggest a role for alcohol dehydrogenase in local retinoic acid synthesis. *Alcohol Clin. Exp. Res.* **20**, 1050–1064 (1996).
42. Haselbeck, R. J., Ang, H. L., Deltour, L. & Duester, G. Retinoic acid and alcohol/retinol dehydrogenase in the mouse adrenal gland: a potential endocrine source of retinoic acid during development. *Endocrinology* **138**, 3035–3041 (1997).
43. Deltour, L., Foglio, M. H. & Duester, G. Metabolic deficiencies in alcohol dehydrogenase Adh1, Adh3, and Adh4 null mutant mice. Overlapping roles of Adh1 and Adh4 in ethanol clearance and metabolism of retinol to retinoic acid. *J. Biol. Chem.* **274**, 16796–16801 (1999).
44. Whelan, J. T., Chen, J., Miller, J., Morrow, R. L., Lingo, J. D., Merrell, K. et al. 9-cis-retinoic acid promotes cell adhesion through integrin dependent and independent mechanisms across immune lineages. *J. Nutr. Biochem* **24**, 832–841 (2013).
45. Paez-Pereda, M., Kovalovsky, D., Hopfner, U., Theodoropoulou, M., Pagotto, U., Uhl, E. et al. Retinoic acid prevents experimental Cushing syndrome. *J. Clin. Invest* **108**, 1123–1131 (2001).
46. Grune, T., Lietz, G., Palou, A., Ross, A. C., Stahl, W., Tang, G. et al. Beta-carotene is an important vitamin A source for humans. *J. Nutr.* **140**, 2268S–2285S (2010).
47. Amengual, J., Widjaja-Adhi, M. A., Rodriguez-Santiago, S., Hessel, S., Golczak, M., Palczewski, K. et al. Two carotenoid oxygenases contribute to mammalian provitamin A metabolism. *J. Biol. Chem.* **288**, 34081–34096 (2013).

48. Wu, L., Guo, X., Wang, W., Medeiros, D. M., Clarke, S. L., Lucas, E. A. et al. Molecular aspects of beta, beta-carotene-9', 10'-oxygenase 2 in carotenoid metabolism and diseases. *Exp. Biol. Med.* **241**, 1879–1887 (2016).
49. Zhao, D., McCaffery, P., Ivins, K. J., Neve, R. L., Hogan, P., Chin, W. W. et al. Molecular identification of a major retinoic-acid-synthesizing enzyme, a retinaldehyde-specific dehydrogenase. *Eur. J. Biochem.* **240**, 15–22 (1996).
50. Niederreither, K., Subbarayan, V., Dolle, P. & Chambon, P. Embryonic retinoic acid synthesis is essential for early mouse post-implantation development. *Nat. Genet.* **21**, 444–448 (1999).
51. Duester, G. Retinoic acid synthesis and signaling during early organogenesis. *Cell* **134**, 921–931 (2008).
52. Szabo, P. M., Tamasi, V., Molnar, V., Andrasfalvy, M., Tombol, Z., Farkas, R. et al. Meta-analysis of adrenocortical tumour genomics data: novel pathogenic pathways revealed. *Oncogene* **29**, 3163–3172 (2010).
53. Bushue, N. & Wan, Y. J. Retinoid pathway and cancer therapeutics. *Adv. Drug Deliv. Rev.* **62**, 1285–1298 (2010).
54. Nagy, Z., Baghy, K., Hunyadi-Gulyas, E., Micsik, T., Nyiro, G., Racz, G. et al. Evaluation of 9-cis retinoic acid and mitotane as antitumoral agents in an adrenocortical xenograft model. *Am. J. Cancer Res.* **5**, 3645–3658 (2015).
55. Szabo, D. R., Baghy, K., Szabo, P. M., Zsippai, A., Marczell, I., Nagy, Z. et al. Antitumoral effects of 9-cis retinoic acid in adrenocortical cancer. *Cell Mol. Life Sci.* **71**, 917–932 (2014).
56. Gara, S. K., Lack, J., Zhang, L., Harris, E., Cam, M. & Kebebew, E. Metastatic adrenocortical carcinoma displays higher mutation rate and tumor heterogeneity than primary tumors. *Nat. Commun.* **9**, 4172 (2018).
57. El Zein, R. M., Soria, A. H., Golib Dzib, J. F., Rickard, A. J., Fernandes-Rosa, F. L., Samson-Couterie, B. et al. Retinoic acid receptor alpha as a novel contributor to adrenal cortex structure and function through interactions with Wnt and Vegfa signalling. *Sci. Rep.* **9**, 14677 (2019).



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020