AI-guided identification of risk variants for adrenocortical tumours in *TP53* p.R337H carrier children: a genetic association study

Bonald C. Figueiredo,[a,b,*] Florent Dupont,[c] Geoffrey Portelli,[c] Tatiana E. J. Costa,[d] Gislaine Custódio,[e] Mariana M. Paraizo,[a] Heloisa Komechen,[a] Hadrien Gascuel,[c] Maxime Bottau,[c] Elodie Callea,[c] Ana Paula Percicote,[f] Leila G. Telles,[f] Mehdi Jendoubi,[c] and Enzo Lalli[g,h,i,**]

[a]Instituto de Pesquisa Pelé Pequeno Príncipe, Curitiba, PR, Brazil
[b]Faculdades Pequeno Príncipe, Curitiba, PR, Brazil
[c]Thales Services Numériques, Valbonne - Sophia Antipolis, France
[d]Hospital Infantil Joana de Gusmão, Florianópolis, SC, Brazil
[e]Laboratório de Análises Clínicas, Hospital de Clínicas, Universidade Federal do Paraná, Curitiba, PR, Brazil
[f]Hospital Pequeno Príncipe, Curitiba, PR, Brazil
[g]Institut de Pharmacologie Moléculaire et Cellulaire CNRS, Valbonne - Sophia Antipolis, France
[h]Inserm, Valbonne - Sophia Antipolis, France
[i]Université Côte d'Azur, Valbonne - Sophia Antipolis, France

The Lancet Regional Health - Americas 2024;38: 100863

Published Online xxx
https://doi.org/10.1016/j.lana.2024.100863

Summary

Background Adrenocortical tumours (ACT) in children are part of the Li-Fraumeni cancer spectrum and are frequently associated with a germline *TP53* pathogenic variant. *TP53* p.R337H is highly prevalent in the south and southeast of Brazil and predisposes to ACT with low penetrance. Thus, we aimed to investigate whether genetic variants exist which are associated with an increased risk of developing ACT in *TP53* p.R337H carrier children.

Methods A genetic association study was conducted in trios of children (14 girls, 7 boys) from southern Brazil carriers of *TP53* p.R337H with (n = 18) or without (n = 3) ACT and their parents, one of whom also carries this pathogenic variant (discovery cohort). Results were confirmed in a validation cohort of *TP53* p.R337H carriers with (n = 90; 68 girls, 22 boys) or without ACT (n = 302; 165 women, 137 men).

Findings We analysed genomic data from whole exome sequencing of blood DNA from the trios. Using deep learning algorithms, according to a model where the affected child inherits from the non-carrier parent variant(s) increasing the risk of developing ACT, we found a significantly enriched representation of non-coding variants in genes involved in the cyclic AMP (cAMP) pathway known to be involved in adrenocortical tumorigenesis. One among those variants (rs2278986 in the *SCARB1* gene) was confirmed to be significantly enriched in the validation cohort of *TP53* p.R337H carriers with ACT compared to carriers without ACT (OR 1.858; 95% CI 1.146, 3.042, p = 0.01).

Interpretation Profiling of the variant rs2278986 is a candidate for future confirmation and possible use as a tool for ACT risk stratification in *TP53* p.R337H carriers.

Funding Centre National de la Recherche Scientifique (CNRS), Behring Foundation, Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Keywords: Adrenal gland; Paediatrics; Oncology; Adrenocortical tumours; Li-Fraumeni spectrum; *TP53* p.R337H

Introduction

Pathogenic variants in the *TP53* tumour suppressor gene are common in many types of cancer. When people inherit certain harmful modifications in this gene, especially when the gene's function is severely affected, they can develop multiple and repeated tumours starting from a young age. This condition is known as Li-Fraumeni syndrome (LFS). These pathogenic variants

*Corresponding author. Instituto de Pesquisa Pelé Pequeno Príncipe, Curitiba, PR, Brazil.
**Corresponding author. Institut de Pharmacologie Moléculaire et Cellulaire CNRS, Valbonne - Sophia Antipolis, France.
E-mail addresses: bonald.figueiredo@professor.fpp.edu.br (B.C. Figueiredo), ninino@ipmc.cnrs.fr (E. Lalli).

Research in context

Evidence before this study

To gather existing evidence related to genetic factors that modify the risk of developing cancer in carriers of the *TP53* p.R337H pathogenic variant, we conducted a systematic search of the literature using the PubMed and Web of Science databases for papers published after 2001 (the time of identification of *TP53* p.R337H in the Brazilian population) and before July 30, 2024, using the search terms (*TP53* OR p53) AND (p.R337H OR R337H) AND cancer risk, without language limitations. The search identified 50 studies. This search aimed to identify studies that discussed the prevalence, genetic implications, and associated risks of the *TP53* p.R337H pathogenic variant and allowed us to understand the current knowledge landscape and reveal knowledge gaps that our study could address.

Added value of this study

Our study significantly advances the current understanding of the cancer risk associated with the *TP53* p.R337H pathogenic variant by identifying specific non-coding genetic changes that increase the risk of developing adrenocortical tumours (ACT) in carriers of this variant. Unlike most previous studies that primarily focused on the broad cancer risk associated with *TP53* mutations, our study identified the risk for a specific type of tumour linked to the presence of particular genetic variants (in *GNAS*, *ADCY7*, and *SCARB1*) associated with genes involved in the cAMP signalling pathway. The identification of the rs2278986 variant in *SCARB1*, which shows a significantly higher frequency in *TP53* p.R337H carrier children with ACT, adds an important element to understanding the genetic susceptibility to this disorder and provides a potential target for genetic screening and risk stratification in the clinical setting.

Implications of all the available evidence

The combined evidence from our study and the existing literature has several important implications. Clinically, the identification of the rs2278986 variant in *SCARB1* as a significant risk factor for ACT in *TP53* p.R337H carriers can inform personalised screening and early intervention strategies. From a public health perspective, incorporating rs2278986 genotyping into existing neonatal screening programs, such as the one ongoing in the State of Paraná, Brazil, could enhance early detection and prevention efforts for ACT. For future research, our findings underscore the importance of investigating non-coding genetic variants and their role in cancer susceptibility. Further studies are needed to validate these findings in larger and more diverse populations, as well as to explore the underlying biological mechanisms. Additionally, our methodology can be applied to other cancers associated with germline pathogenic variants, potentially leading to broader applications in cancer risk assessment and management.

are rare because they reduce reproductive success. However, some variants in the *TP53* gene only partially affect its function, leading to a lower probability of developing cancer in their carriers.[1] One specific *TP53* variant, c.1010G > A (p.R337H), is very common in the south and southeast of Brazil. The first study describing this variant found it to be linked to a specific type of childhood cancer, adrenocortical tumours (ACT), which is much more common in this area than elsewhere in the world.[2] Further research linked this *TP53* variant to LFS and to a similar condition called Li-Fraumeni-like syndrome (LFL), which shows some but not all features of LFS.[3,4] In families where *TP53* p.R337H carriers are present, those carrying the variant were reported to be 5.7 times more likely to develop cancer than non-carriers. In children and adolescents with this pathogenic variant, the most common cancers were ACT and a specific type of brain cancer (choroid plexus carcinoma). In adults, TP53 p.R337H carriers were more likely to get breast and stomach cancers compared to non-carriers.[5]

Nevertheless, the likelihood of developing cancer in *TP53* p.R337H carriers is much lower than in people who carry more severe *TP53* pathogenic variants. Recent studies show that fewer than 5% of *TP53* p.R337H carriers develop ACT during their childhood.[6] Similarly, mice carrying the homologous pathogenic variant in *Tp53* developed tumours with long latency and incomplete penetrance.[7,8] Unlike other *TP53* pathogenic variants, p.R337H does not seem to affect reproductive health, as most carriers do not develop cancer until later in life and are healthy during their reproductive years. In addition to neonatal screening for *TP53* p.R337H carrier status,[9] it is crucial to identify the genetic and environmental factors that influence ACT penetrance in carriers of this pathogenic variant. This represents an important public health concern for a population of approximately 100 million people residing across four major states of south and southeast Brazil (São Paulo, Paraná, Santa Catarina, Rio Grande do Sul) and in bordering countries (Paraguay and possibly Argentina).[10] Furthermore, epidemiological data about other Brazilian states are scarce at present and *TP53* p.R337H contribution to the cancer burden in other regions is currently poorly known.[11] Our study aimed to identify additional genetic factors that increase the risk of ACT in *TP53* p.R337H carriers by analysing genomic data of trios composed of one carrier child, one cancer-free carrier parent, and one non-carrier parent.

Methods

Study design and participants

We investigated genomic data from whole exome sequencing (WES) of blood DNA from 21 trios including carrier children with (n = 18; 13 girls; 5 boys) or without (n = 3; 1 girl, 2 boys) ACT and their parents, one of them being a carrier in each trio and who never had ACT (discovery cohort) (Supplementary Table S1). Written informed consent was obtained from all participants of the study or from their guardians. The study was approved by the Pequeno Príncipe Hospital Ethics Committee (CAAE: 50622315.0.0000.0097, 2015).

Given the low penetrance of the *TP53* p.R337H pathogenic variant,[6] we reasoned that genetic variants influencing ACT development would be distributed according to two alternative models (Fig. 1a and b):

- In one model (***protecting variant model; PVM***) the carrier parent would harbour a variant that protects them from developing ACT. This variant would be lost in the affected child (Fig. 1a);
- In the second model (***damaging variant model; DVM***) the affected child would inherit a variant which increases the risk of developing ACT from the non-carrier parent. The carrier parent would not harbour this variant (Fig. 1b).

To decrease the potential impact of environmental factors on ACT development and to maximize our chances of identifying genetic variants increasing the risk of developing ACT, we aimed to recruit a majority of early onset ACT cases (<3 years of age) in our discovery cohort.

Whole exome sequencing of blood DNA from trios

Features of trios composed of parents and their children are reported in Supplementary Table S1. Genomic DNA extracted from blood (ReliaPrep gDNA tissue Miniprep System, Promega) was captured using Agilent's solution enrichment method with their biotinylated oligonucleotide probe library (Human All Exon V5–50 Mb, Agilent) followed by high-throughput 75 bp paired-end sequencing on an Illumina HiSeq 2000 instrument.

Sequence capture, enrichment, and elution were performed according to the manufacturer's protocol and recommendations (SureSelect, Agilent) without modifications. Briefly, 3 μg of each genomic DNA sample was fragmented by sonication and purified to obtain 150–200 bp fragments. Oligonucleotide adapters for paired-end sequencing were ligated onto the repaired fragments, with adenine added to the ends, then purified and enriched by 4–6 PCR cycles. 500 ng of those purified libraries were then hybridized onto the SureSelect capture oligonucleotide probe library for 24 h. After hybridization, washing and elution, the eluted fraction was amplified by 10–12 PCR cycles, purified and quantified by quantitative PCR to obtain sufficient template DNA for further downstream processing. Each eluted DNA sample was then sequenced on an Illumina HiSeq 2000 instrument to obtain 75 bp paired-end sequences. Image analysis and base determination were performed using the Illumina RTA version 1.14 pipeline with default parameters.

AI-guided data analysis

For variant calling from hg38 aligned WES bam files, we developed a data processing pipeline and integrated genomic tools for cohort data analysis (variant extraction, annotation, pathway identification). The following tools were adapted and integrated into the pipeline (Fig. 1c–e):

1. DeepTrio[12]: This variant calling tool developed by Google uses CNN (convolutional neural networks) for highly accurate variant calling and takes as input cohorts specifically composed of trios (parents and child). This method runs HaplotypeCaller to call single sample variants, followed by GenomicsDBImport and GenotypeGVCFs to consolidate and jointly genotype the cohort, and finally VariantRecalibrator and ApplyVQSR for variant quality score recalibration (VQSR). This tool can leverage trio data weighting sequencing errors by modelling the potential error rate based on the context and other family members. It therefore offers a real added value, in the specific context of this study, compared to other variant calling tools such as DeepVariant (Google) or Clara Parabricks (NVIDIA) (Fig. 1c);
2. GLnexus[13]: This tool is used to cluster data by family. It allows the manipulation of large.bcf files obtained after the variant calling step. Using this method, about 13 million variants were called in the trio data after standard quality thresholding (Fig. 1c);
3. Bcftools[14]: This tool allows the optimized conversion of.bcf files to.vcf format. This format is required for the application of the subsequent analyses (Fig. 1c);
4. SNPEff[15]: This variant annotation tool adds functional information to the variants extracted and sorted at the previous steps.
5. KEGGSeq[16]: This tool allows assigning one or more pathways (set of genes leading to a given phenotypic expression) to each identified and annotated variant. The objective is to be able to exclude from the analysis variants concerning genes involved in pathways that are not related to the targeted disease (for example, in the case of ACT, genes coding for eye colour were not considered relevant).

From the two-model hypothesis (PVM and DVM), we devised a binary rule for each variant and each child based on the existence of at least one allele of the variant among the parents. Importantly, we fixed no threshold

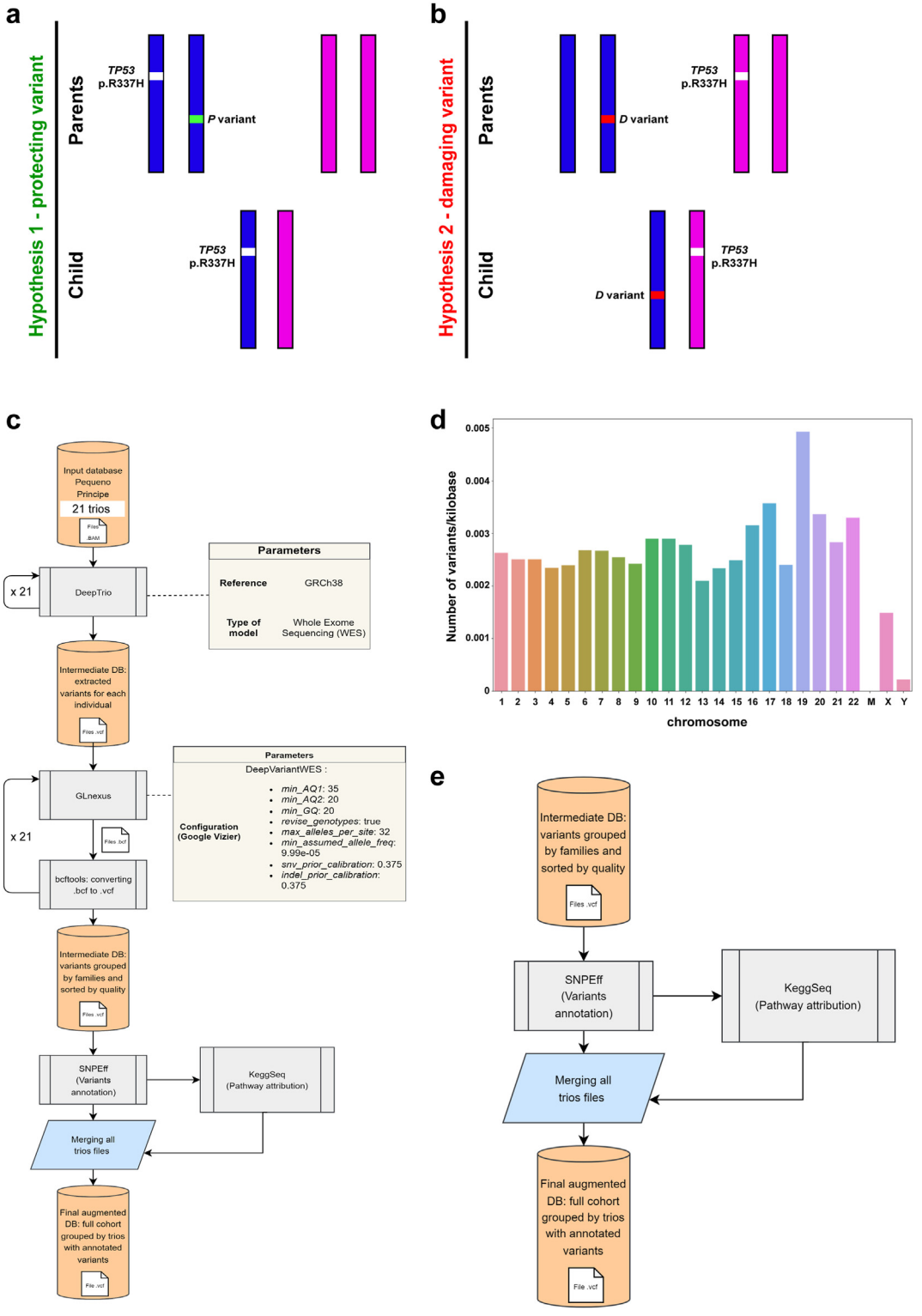

Fig. 1: **Variant calling and annotation in trio WES data. (a)** Hypothesis 1: a protecting variant (P) is present in the germline of the parent carrying the *TP53* p.R337H pathogenic variant and absent in the other parent. The affected child does not carry the P variant. **(b)** Hypothesis 2: a damaging variant (D) is present in the germline of the parent not carrying *TP53* p.R337H and is absent in the carrier parent. The child carries D. Each bar represents the haploid genome; **(c)** Variant calling pipeline using DeepTrio. **(d)** Number of variants found per kilobase for each chromosome in the trios. **(e)** Variant annotation pipeline.

to the frequency of the variants we considered for further analysis, either in the general population [gnomAD (https://gnomad.broadinstitute.org/), dbSNP (https://www.ncbi.nlm.nih.gov/snp/)] or in Brazilian population-specific databases (ABraOM,[17] SELAdb[18]). Also, we considered both coding and non-coding variants, without restrictions linked to the predicted functional impact of the variant on its gene product function. According to the PVM, 35% of total variants (4,638,646) were carried by parents harbouring the *TP53* p.R337H pathogenic variant and were not transmitted to their affected children. According to the DVM, 13.5% of total variants (1,771,398) were present in children with ACT and were inherited from their parents who were not carriers of *TP53* p.R337H. De novo variants (present in children and absent in their parents) were integrated into the analysis. Variants found present in at least 50% (9/18) of ACT cases according to the PVM or the DVM and absent in subjects without ACT were further analysed for Gene Ontology (GO) classification by expert analysis using Metascape.[19]

SNP genotyping

SNPs were genotyped in an independent validation cohort of *TP53* p.R337H carriers 1) with ACT ($n = 90$; 68 girls, 22 boys), 2) cancer-free ($n = 302$; 165 women, 137 men), or 3) with other cancers ($n = 21$; all women). All subjects are part of a project led by the Pequeno Principe Institute carrying out a prolonged follow-up after ACT development in any person in the family. In addition to *TP53* p.R337H, participants enrolled in the project also agreed to investigate associated germline and somatic variants. The parents were invited to participate in the study and authorised their children by signing a consent form (in addition, children older than 8 years signed an assent form). They are all part of a large cohort including 598 *TP53* p.R337H carrier families identified until the present time since the discovery of the germline TP53 p.R337H in Brazil.[2]

Taqman assays: C___7611069_10 for rs919196, C_102147632_10 for rs78051832 and C__15966787_10 for rs2278986 (Thermo Fisher) were performed according to the manufacturer's method. Subjects bearing at least one alternative allele were considered carriers of that allele, independently from their hetero- or homozygote status.

Statistical analysis

Odds ratio, confidence intervals and 1-sided p-values for differences in variant frequencies in the validation cohort groups (*TP53* p.R337H carriers with or without ACT) were calculated by the Fisher's exact test in GraphPad Prism v.10.2.3. The frequencies of the genotyped variants and their 95% confidence intervals in *TP53* p.R337H subjects with and without ACT of the validation cohort were calculated using Sample Size Calculators (https://sample-size.net/confidence-interval-proportion/).

Role of the funding source

The funders had no role in the study design, data collection, data analysis, interpretation or writing of the report.

Results

Variants associated with ACT in trios according to the PVM and the DVM models

We ranked variants in order of frequency according to the PVM-DVM hypothesis and submitted their associated genes to GO classification using Metascape,[19] considering recurrent variants present in at least 50% (9/18) of the trios including children with ACT in a pattern consistent with either the PVM or the DVM (Fig. 2a and b and Supplementary Table S2) and absent in children without ACT. Comparison of enriched GO terms in DisGeNET[20] highlighted a significant enrichment for variants associated with genes involved in adrenocortical adenoma pathogenesis in the DVM group, but not among PVM variants (Fig. 2c). The variants evidenced by our analysis are listed in Table 1.

All three variants are associated with genes involved in cyclic AMP (cAMP) signalling, a pivotal signal transduction pathway in adrenocortical physiology and tumorigenesis.[21] The first is a variant found in intron 6 of the *GNAS* gene, encoding the stimulatory G-protein alpha subunit, a key element linking receptor–ligand interactions with the activation of adenylyl cyclase and a variety of cellular responses. The rs919196 SNP lies inside a haplotype block with high-degree genomic constraint in the 3' portion of the gene where disease-causing variants and strong eQTLs for *GNAS* are localized[22] (Supplementary Figure S1). rs78051832 is an intronic SNP and eQTL for the *ADCY7* gene, encoding a member of the adenylyl cyclase class-4/guanylyl cyclase enzyme family, playing a fundamental role in the biosynthesis of cAMP. rs2278986 is an intronic SNP for *SCARB1*, encoding a prominent member of the scavenger receptor family (SR-BI), a membrane receptor for high-density lipoprotein (HDL) cholesterol, the major route for cholesterol delivery to the steroidogenic pathway. This SNP is situated in a region of high-degree genomic constraint (Supplementary Figure S2).

Risk variants frequency in a validation cohort composed of TP53 p.R337H carriers without cancer/with ACT/with other cancers

To confirm the association of each one of those variants with ACT in children carriers of *TP53* p.R337H, we compared their distribution in an independent cohort of carriers with or without ACT (Supplementary Table S3) by SNP genotyping analysis using TaqMan assays. The age and sex distribution of this validation cohort of *TP53* p.R337H carrier children with ACT were similar to the children with ACT from the trios, while participants without cancer were significantly older, reaching ages beyond the time limit of paediatric ACT diagnosis

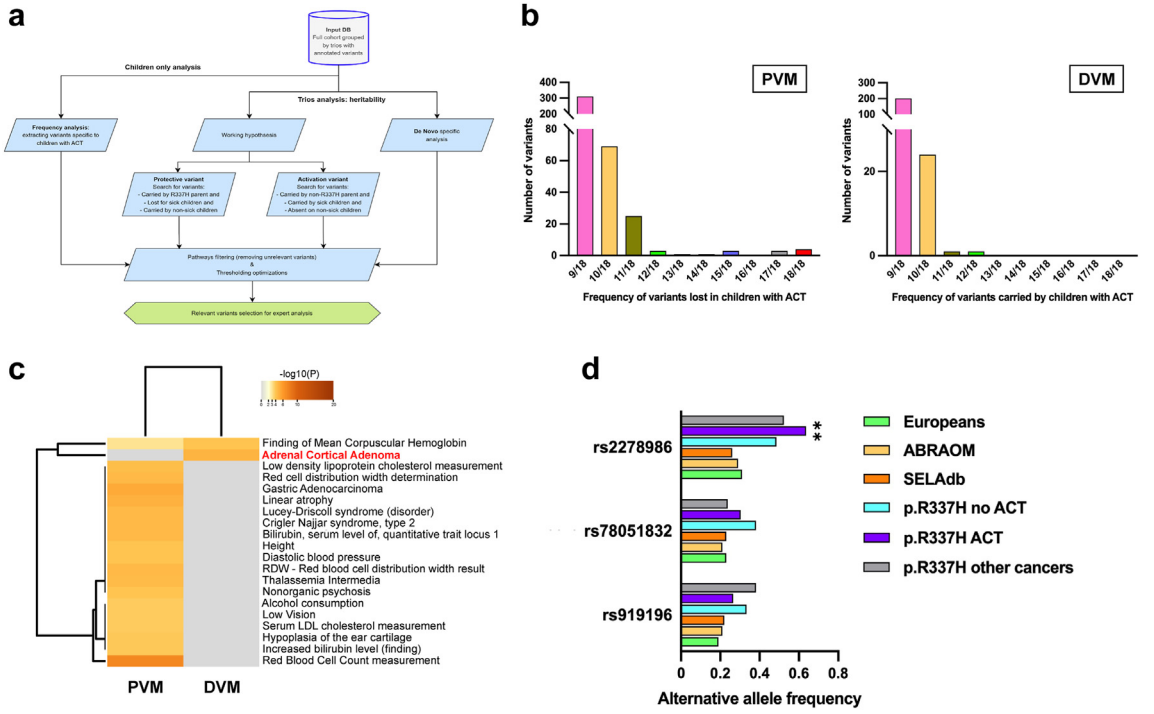

Fig. 2: **Variants associated with ACT development in carriers of the *TP53* p.R337H pathogenic variant**. **(a)** Pipeline for variant classification, filtering and analysis. **(b)** Frequency of variants according to the PVM and DVM, respectively. **(c)** Differentially enriched GO terms in the PVM and DVM. Log10 probability scale is shown. **(d)** Alternative allele frequency of three SNPs (rs919196, rs78051832 and rs2278986) associated with ACT development in Europeans, ABRAOM-SELAdb databases and in our validation cohort of tumor-negative and -positive children carrier of *TP53* p.R337H without (n = 302) and with ACT (n = 90) and of adult carriers with other cancers (n = 21). **p < 0.01, Fisher's exact test.

Variant	SNP	Associated gene	Localization	Fequency in trios
chr20: 58,909,030 T > C	rs919196	*GNAS*	intron	12/18 (63%)
chr16: 50,309,400 C > T	rs78051832	*ADCY7*	intron	9/18 (50%)
chr12: 124,814,823 A > G	rs2278986	*SCARB1*	intron	9/18 (50%)

Table 1: **Variants in genes involved in adrenocortical adenoma pathogenesis according to the DVM in our trio cohort.**

(Supplementary Figure S3). Those participants can then be considered definitively free of risk of developing paediatric ACT. The alternative allele frequencies and their 95% confidence limits in the validation cohort are reported in Supplementary Table S4. In the validation cohort, the alternative allele frequency for rs2278986 (0.636) was significantly higher in participants with ACT compared with those who never had ACT (0.485) (OR 1.858, 95% CI 1.146, 3.042, p = 0.01), while no significant difference was present in the frequency of alternative alleles for rs919196 (0.280 in *TP53* p.R337H carriers with ACT vs. 0.333 in carriers without ACT; OR 0.779, 95% CI 0.455, 1.329, p = 0.22) and rs78051832 (0.372 in *TP53* p.R337H carriers with ACT vs. 0.381 in carriers without ACT; OR 0.961, 95% CI 0.581, 1.581, p = 0.49). Also, in *TP53* p.R337H carriers with cancers other than ACT there was no significant increase in alternative allele frequency for rs2278986 (0.523) (OR 1.16, 95% CI 0.501, 2.832, p = 0.45), rs919196 (0.381 in *TP53* p.R337H carriers with other cancers; OR 0.72, 95% CI 0.414, 1.237, p = 0.14) and rs78051832 (0.238 in *TP53* p.R337H carriers with other cancers; OR 0.50, 95% CI 0.199, 1.423, p = 0.13) compared with cancer-free *TP53* p.R337H carriers (Fig. 2d). No advantage in predictive value for ACT development was obtained by combining data from the different variants to obtain a polygenic risk score. The cancer ratio data (number of cancers divided by the number of carriers) per family shown in Supplementary Table S3 suggest that *TP53* p.R337H has an attenuated LFS spectrum, with ratios ranging from 0.071 to 0.800 (mean 0.38) in the validation group, with similar findings in the group of other cancers. Moreover, we did not find any correlation between those ratios and the variant frequencies.

Discussion

The finding that not everyone bearing inherited low-penetrance *TP53* pathogenic variants gets cancer suggests that other factors influence cancer risk and timing.[23] Specifically, in the case of the Brazilian *TP53* p.R337H variant, a nonsense variant in the *XAF1* gene (p.E134X) increases the risk of developing sarcomas and multiple tumours.[24] Additionally, certain non-coding genetic modifiers are linked to the types and number of cancers in *TP53* p.R337H carriers,[25,26] while a synonymous SNP in the *ADH7* gene affects the age at which *TP53* p.R337H carrier children develop ACT.[27]

Our study shows that certain non-coding genetic changes in genes linked to adrenal gland tumours and involved in the cAMP pathway (*GNAS*, *ADCY7*, and *SCARB1*) are significantly more frequent in children with ACT carrying *TP53* p.R337H, according to a model where those variants were inherited from the parent negative for the *TP53* variant (DVM) and then supposed to be predisposed to neoplasia. One specific variant (rs2278986 in *SCARB1*) was confirmed to have a significantly higher frequency in *TP53* p.R337H carrier children with ACT compared to cancer-free carriers in an independent cohort. Conversely, the same variant was not enriched in *TP53* p.R337H carriers affected with other cancers. Since the variants identified are non-coding, we interpret our findings in the sense that they likely do not act *per se* but rather highlight regions with a high degree of genomic constraint (Supplementary Figures S1 and S2) that contain elements important for ACT tumorigenesis in the context of a germline low-penetrance *TP53* pathogenic variant. These data are consistent with previous studies that showed that genetic variation in genes involved in the cAMP pathway are important for paediatric ACT development, regardless of the *TP53* status of carriers.[21,28–31] Interestingly, among the genes associated with the identified variants, *ADCY7* is linked to paediatric autoimmune disorders[32] and may play a role in the immune response against tumours. Carriers of a coding variant in the *SCARB1* gene (p.P297S) have altered platelet function and reduced adrenal hormone production,[33] suggesting this gene is a candidate for primary adrenal insufficiency.[34] Remarkably, carriers of the alternative allele of rs2278986 express lower levels of the SR-BI protein compared with noncarriers.[35] *Scarb1* null mice have enlarged adrenal glands, a consequence of decreased glucocorticoid production and increased ACTH levels.[36] These findings lead us to the speculation that the lower SR-BI expression and subsequent adrenal gland hyperstimulation by ACTH in children carrying both the *TP53* p.R337H and the rs2278986 *SCARB1* variants might increase their risk of developing ACT. This hypothesis needs further investigation. Notably, a few children in the validation cohort were homozygous for the risk variant identified according to the DVM. We would like to emphasize that our strategy for variant identification in trios is primarily designed as a tool to uncover novel genetic elements that increase the risk of developing ACT, rather than purely focusing on the requirement for these variants to be transmitted from a non-carrier parent to an affected child. Even if these variants are present in the carrier parent, they may not invariably lead to ACT in every case. This variability can be attributed to the intricate genetic makeup of each individual. Various factors may influence tumour development, including a robust immune response, the activation of apoptosis, and the regulation of blood vessel development, among others. These factors may collectively contribute to a protective effect, preventing tumour formation despite the presence of potentially predisposing genetic variants. Another possibility is that those alleles became homozygotes in the child because of gene conversion.

Overall, our data, interpreted in the context of the previous literature in the field, suggest that different genetic variants in carriers of low-penetrance *TP53* pathogenic variants are associated with different cancer types, according to the specific biological features of different tissues. However, several limitations of this study need to be considered. First, despite ours being one of the largest ACT patient cohorts composed of *TP53* p.R337H carriers available, our total sample size could still be considered relatively small to identify and validate tumour risk variants. Also, the absence of pathological material from the investigated patients prevented us from analysing loss of heterozygosity (LOH) for *TP53* and rs2278986 in the tumour tissue. Further validation will be required to evaluate the association of the variants identified in our study on ACT development in other populations, in patients carrying another germline *TP53* pathogenic variants and in other tumours linked to LFS. In addition, demographic data, including race and ethnicity and other possible mediators of environmental exposure was not collected for the population enrolled in this study and could not be analysed. Further functional studies are required to dissect the mechanisms of how the ACT risk variants we identified modulate the tumorigenic process in the adrenal gland. Unfortunately, it will not be possible to produce a mouse model associating the p.R334H (homologous to the human p.R337H) pathogenic variant in *Tp53*[7,8] with the *Scarb1* variant since the human genomic region encompassing rs2278986 (chr12: 124,814,823 in hg38) is not conserved in the mouse genome.

In conclusion, we have shown in this study that it is possible to identify genetic variants associated with an increased risk of developing ACT by AI-guided analysis of WES data from a small number of trios. If validated, the implementation of rs2278986 profiling could be a useful tool for ACT risk stratification in carriers of the TP53 p.R337H pathogenic variant in the clinical setting. In the State of Paraná, in the south of Brazil, a newborn

screening program is currently ongoing to identify *TP53* p.R337H carriers for early detection of ACT and other tumours.[9] Associating rs2278986 genotyping into this screening program to detect individuals at higher risk of developing ACT could help validate its clinical usefulness in a large prospective study. In addition, our methods could also be applied to other paediatric and adult cancers associated with germline pathogenic variants in oncogenes or tumour suppressor genes.

Contributors

BCF and EL designed research, acquired funding, and coordinated the study. BCF, TEJC, APP, and LGT provided clinical samples. BCF, FD, GP, GC, MMP, HK, HG, MB, EC, MJ, and EL collected and analysed the data. BCF, FD, GP, and EL directly accessed and verified the data reported in this article. BCF and EL were responsible for the decision to submit the article. All authors contributed to writing the report and have seen and approved the final text. All authors had access to all the data in the study and had final responsibility for the decision to submit for publication.

Data sharing statement

Anonymized WES data from the trio study are available from the authors upon reasonable request after signing a data access agreement.

Declaration of interests

None of the authors has anything to declare.

Acknowledgements

We thank Liv Bastos, Noé Dumas, Julie Galliere, Rapahëlle Roty, and Yasser Tba for their contribution to the initial part of the project; Emilia Pinto, Raul Ribeiro, and Gerard Zambetti for discussions. This work was supported by the Behring Foundation (Big Data Grant 2023-76.591) and the CNPq 303592/2021-4 grant to BCF, and by the CNRS EXPOGEN-CANCER International Research Project and IIPACT-ICPACT International Research Network to BCF and EL.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.lana.2024.100863.

References

1 Kratz CP, Freycon C, Maxwell KN, et al. Analysis of the Li-Fraumeni spectrum based on an international germline TP53 variant data set: an international agency for research on cancer TP53 database analysis. *JAMA Oncol*. 2021;7:1800–1805.
2 Ribeiro RC, Sandrini F, Figueiredo B, et al. An inherited p53 mutation that contributes in a tissue-specific manner to pediatric adrenal cortical carcinoma. *Proc Natl Acad Sci U S A*. 2001;98:9330–9335.
3 Achatz MIW, Olivier M, Le Calvez F, et al. The *TP53* mutation, p.R337H, is associated with Li-Fraumeni and Li-Fraumeni-like syndromes in Brazilian families. *Cancer Lett*. 2007;245:96–102.
4 Palmero EI, Schüler-Faccini L, Caleffi M, et al. Detection of R337H, a germline *TP53* mutation predisposing to multiple cancers, in asymptomatic women participating in a breast cancer screening program in Southern Brazil. *Cancer Lett*. 2008;261:21–25.
5 Mastellaro MJ, Seidinger AL, Kang G, et al. Contribution of the TP53 p.R337H mutation to the cancer burden in southern Brazil: insights from the study of 55 families of children with adrenocortical tumors. *Cancer*. 2017;123:3150–3158.
6 Costa TEJ, Gerber VKQ, Ibañez HC, 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*. 2021;11:1804.
7 Park JH, Li J, Starost MF, et al. Mouse homolog of the human *TP53* R337H mutation reveals its role in tumorigenesis. *Cancer Res*. 2018;78:5375–5383.
8 Jeffers JR, Pinto EM, Rehg JE, et al. The common germline *TP53*-R337H mutation is hypomorphic and confers incomplete penetrance and late tumor onset in a mouse model. *Cancer Res*. 2021;81:2442–2456.
9 Custódio G, Parise GA, Kiesel Filho N, et al. Impact of neonatal screening and surveillance for the *TP53* p.R337H mutation on early detection of childhood adrenocortical tumors. *J Clin Oncol*. 2013;31:2619–2626.
10 Legal EF, Ascurra M, Custódio G, et al. Prevalence of an inherited cancer predisposition syndrome associated with the germ line *TP53* R337H mutation in Paraguay. *Cancer Epidemiol*. 2015;39:166–169.
11 Vieira IA, Andreis TF, Fernandes BV, et al. Prevalence of the Brazilian *TP53* founder c.1010G>A (p.Arg337His) in lung adenocarcinoma: is genotyping warranted in all Brazilian patients? *Front Genet*. 2021;12:606537.
12 Kolesnikov A, Goel S, Nattestad M, et al. DeepTrio: variant calling in families using deep learning. *bioRxiv*. 2021. https://doi.org/10.1101/2021.04.05.438434.
13 Poplin R, Chan PC, Alexander D, et al. A universal SNP and small-indel variant caller using deep neural networks. *Nat Biotechnol*. 2018;36:983–987.
14 Danecek P, McCarthy SA. BCFtools/csq: haplotype-aware variant consequences. *Bioinformatics*. 2017;33:2037–2039.
15 Cingolani P, Platts A, Wang le L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of drosophila melanogaster strain w1118; iso-2; iso-3. *Fly (Austin)*. 2012;6:80–92.
16 Korani W, Chu Y, Holbrook C, et al. Insight into genes regulating post-harvest aflatoxin contamination of tetraploid peanut from transcriptional profiling. *Genetics*. 2018;209:143–156.
17 Naslavsky MS, Scliar MO, Yamamoto G, et al. Whole-genome sequencing of 1,171 elderly admixed individuals from São Paulo, Brazil. *Nat Commun*. 2022;13:1004.
18 Lerario AM, Mohan DR, Montenegro LR, et al. SELAdb: a database of exonic variants in a Brazilian population referred to a quaternary medical center in São Paulo. *Clinics*. 2020;75:e1913.
19 Zhou Y, Zhou B, Pache L, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun*. 2019;10:1523.
20 Piñero J, Bravo À, Queralt-Rosinach N, et al. DisGeNET: a comprehensive platform integrating information on human disease-associated genes and variants. *Nucleic Acids Res*. 2017;45:D833–D839.
21 Kamilaris CDC, Hannah-Shmouni F, Stratakis CA. Adrenocortical tumorigenesis: lessons from genetics. *Best Pract Res Clin Endocrinol Metab*. 2020;4:101428.
22 Yang W, White B, Spicer EK, et al. Complex haplotype structure of the human GNAS gene identifies a recombination hotspot centred on a single nucleotide polymorphism widely used in association studies. *Pharmacogenetics*. 2004;14:741–747.
23 Majhi PD, Sharma A, Jerry DJ. Genetic modifiers of p53: opportunities for breast cancer therapies. *Oncotarget*. 2023;14:236–241.
24 Pinto EM, Figueiredo BC, Chen W, et al. *XAF1* as a modifier of p53 function and cancer susceptibility. *Sci Adv*. 2021;6:eaba3231.
25 Bandeira IC, Vieira IA, Andreis TF, et al. MIR605 rs2043556 is associated with the occurrence of multiple primary tumors in *TP53* p.(Arg337His) mutation carriers. *Cancer Genet*. 2020;240:54–58.
26 Vieira IA, Pezzi EH, Bandeira IC, et al. Functional *pri-miR-34b/c* rs4938723 and *KRAS* 3′UTR rs61764370 SNPs: novel phenotype modifiers in Li-Fraumeni syndrome? *Gene*. 2024;898:148069.
27 Surakhy M, Wallace M, Bond E, et al. A common polymorphism in the retinoic acid pathway modifies adrenocortical carcinoma age-dependent incidence. *Br J Cancer*. 2020;122:1231–1241.
28 Fragoso MCBV, Domenice S, Latronico AC, et al. Cushing's syndrome secondary to adrenocorticotropin- independent macronodular adrenocortical hyperplasia due to activating mutations of *GNAS1* gene. *J Clin Endocrinol Metab*. 2003;88:2147–2151.
29 Sidhu A, Debelenko L, Misra VK. Infantile adrenocortical tumor with an activating *GNAS1* mutation. *J Clin Endocrinol Metab*. 2013;98:E1115–E1118.
30 Pinto EM, Faucz FR, Paza LZ, et al. Germline variants in phosphodiesterase genes and genetic predisposition to pediatric adrenocortical tumors. *Cancers*. 2020;12:E506.
31 Rege J, Hoxie J, Liu C-J, et al. Targeted mutational analysis of cortisol-producing adenomas. *J Clin Endocrinol Metab*. 2022;107:e594–e603.

32 Li YR, Li J, Zhao SD, et al. Meta-analysis of shared genetic architecture across ten pediatric autoimmune diseases. *Nat Med*. 2015;21:1018–1027.

33 Vergeer M, Korporaal SJ, Franssen R, et al. Genetic variant of the scavenger receptor BI in humans. *N Engl J Med*. 2011;364:136–145.

34 Hoekstra M. Identification of scavenger receptor BI as a potential screening candidate for congenital primary adrenal insufficiency in humans. *Am J Physiol Endocrinol Metab*. 2020;319:E102–E104.

35 West M, Greason E, Kolmakova A, et al. Scavenger receptor class B type I protein as an independent predictor of high-density lipoprotein cholesterol levels in subjects with hyperalphalipoproteinemia. *J Clin Endocrinol Metab*. 2009;94:1451–1457.

36 Hoekstra M, Ye D, Hildebrand RB, et al. Scavenger receptor class B type I-mediated uptake of serum cholesterol is essential for optimal adrenal glucocorticoid production. *J Lipid Res*. 2009;50:1039–1046.