Published in final edited form as:

Genes Chromosomes Cancer. 2018 September ; 57(9): 478–481. doi:10.1002/gcc.7.

Analysis of the Exome Aggregation Consortium (ExAC) database suggests that the *BAP1*-tumor predisposition syndrome is underreported in cancer patients

James B. Massengill^{1,¥}, Klarke M. Sample^{1,4,¥}, Robert Pilarski², Joseph McElroy³, Frederick H. Davidorf¹, Colleen M. Cebulla¹, and Mohamed H. Abdel-Rahman^{1,2,¥,*}

¹Department of Ophthalmology, The Ohio State University, Columbus, OH

²Division of Human Genetics, Department of Internal Medicine, The Ohio State University, Columbus, OH

³Center for Biostatistics, Department of Biomedical Informatics, The Ohio State University, Columbus, OH

Abstract

The *BAP1*-tumor predisposition syndrome (*BAP1*-TPDS) has been recently identified to predispose patients to a variety of cancers and preneoplastic lesions. About 130 unrelated probands have been identified worldwide; however, the impact of the syndrome is suspected to be much larger given the diversity of the cancer phenotype. To evaluate the frequency of germline *BAP1* mutations in the general and cancer populations, we analyzed the Exome Aggregation Consortium (ExAC), a database that contains 53105 exomes of unrelated individuals unaffected by cancer (general population) and exomes of 7601 unrelated individuals affected by cancer provided by the Cancer Genome Atlas (TCGA, cancer subjects). *BAP1* null variants were seen at much higher frequency in the cancer subjects (0.0526%) compared to the general population (0.00188%) with a relative risk of 27.93 and ($P=0.0011$, [95% CI: 3.122–249.883], Fisher's exact test). We also studied a reported *BAP1* null variant, c.1203T>G, p.T401* (rs200156887), observed commonly in the general population. Sequencing and restriction fragment polymorphism of the RT-4 cell line that contains this variant revealed that it is in fact a 3bp deletion/insertion, c.1201_1203delinsGAG, a likely benign missense alteration p.Y401E explaining the relative high frequency of this variant in the general population. In conclusion, germline null mutations in *BAP1* have a significantly higher frequency in cancer patients than the general population. Given the low frequency of reported families with *BAP1*-TPDS, our results suggest that the syndrome is underreported especially in patients with cancer.

Correspondence: Mohamed H. Abdel-Rahman, M.D., Ph.D., 400 W 12th Ave, Room 202, Columbus, OH, 43210, Phone: 614 292 1396, Fax: 614 247 7153, Mohamed.abdel-rahman@osumc.edu.

¥Authors contributed equally

⁴Current address: Guizhou Provincial People's Hospital, Guiyang, China

INTRODUCTION

BAP1-tumor predisposition syndrome *BAP1*-TPDS (OMIM # 614327) predisposes patients to a variety of cancers and preneoplastic lesions¹⁻⁵. Four main cancers (uveal melanoma, malignant mesothelioma, renal cell carcinoma and cutaneous melanoma) are currently confirmed to be part of the syndrome. However, other cancers could also be part of the clinical phenotype¹⁻⁵. The syndrome is caused by heterozygous germline mutation in *BAP1* and is inherited in an autosomal dominant pattern. About 130 unrelated probands have been identified worldwide; however, given the diversity of the cancer phenotype the syndrome could be underreported¹. In the following study we assessed the frequency of germline null variants in cancer and general population using the Exome Aggregation Consortium (ExAC)⁶, a publically available database. Our results suggest that *BAP1*-TPDS is likely underreported in cancer patients.

MATERIALS AND METHODS

Variant frequency assessment

We conducted an analysis of the *BAP1* mutations contained within the ExAC, which contains 53105 exomes of unrelated individuals unaffected by cancer and 7601 unrelated individuals affected by cancer provided by The Cancer Genome Atlas (TCGA). These publically available datasets were presented by the Broad Institute and MIT in two files: ExAC with TCGA and ExAC without TCGA. The distribution of variants in the TCGA was obtained by subtraction of the two datasets. ExAC data were filtered on PASS quality status. Computational assessment of the pathogenesis of missense variants was carried out using SIFT⁷ and PolyPhen2⁸ software.

Assessment of the p.Y401* truncating mutation

The RT4 cell line was obtained from the American Type Culture Collection (ATCC, Manassas, VA) and was cultured according to the supplier's recommended protocols. DNA was extracted using QIAamp DNA Micro Kit (Qiagen, Hilden, Germany). PCR amplification of the region surrounding the two variants was performed using the forward primer 5'-ATCCAGTGTAAG-TGGGTGGCAGC-3' and the reverse primer 5'-CCAAACTCCGCAGGTGCTCAACA-3'. PCR was carried out on a BioRad C1000 thermal cycler with the following PCR program: 95° for 15 min, 35 cycles of 94° for 30 sec, 60° for 30 sec, and 72° for 45 sec and a final elongation at 72° for 10 min. The product was then analyzed via restriction fragment polymorphism (RFLP) using the BfaI enzyme (New England BioLabs, Ipswich, MA) using the manufacturer's suggested protocol and visualized on a 2.5% agarose gel.

RESULTS AND DISCUSSION

Five different *BAP1* null variants, as defined by the ACMG guidelines⁹, were observed (Table 1). Null variants were seen at much higher frequency in the TCGA subjects (0.0526%) compared to the general population (0.00188%) with a relative risk of 27.93 ($P=0.0011$, [95% CI: 3.122–249.883], Fisher's exact test) (Table 1). Given the estimated carrier frequency in cancer patients of 1 in 1900, we should expect that about 7367 subjects,

among the 14 million patients diagnosed with cancer in the US, have germline null variants in *BAP1*. However, less than 50 families have been reported in the US and only 130 worldwide¹. This suggests that the *BAP1*-TPDS is significantly underreported. Recently we identified a high frequency of germline mutations in *BAP1* that were misreported as somatic mutations through tumor-only testing¹⁰. As such we recommend that somatic *BAP1* variants reported through tumor-only testing should be reflex tested in the germline. We also recommend education of practitioners in the oncology field about the syndrome and suggested management guidelines¹¹. *BAP1* has been added to cancer gene panels offered by several clinical laboratories. Our data strongly support such inclusion.

Thirty-one missense variants predicted as potentially deleterious by SIFT and/or PolyPhen 2 were identified, including 26 rare variants observed at frequency of <0.0001. Rare *BAP1* missense variants were seen at slightly higher frequency in cancer patients compared to the general population; however the difference was not statistically significant ($P>0.05$; relative risk = 1.747 [95%CI: 0.805–3.792])(Figure 1A and Table 1). The distribution of these missense variants in the different functional domains of BAP1 was similar in the cancer patients and general population(Figure 1). In addition, three in-frame deletions were observed in the general population but not in the cancer patients. Two of these variants, p.Gln441del and p.Glu406del, were located outside the deubiquitinase domain while one (p.Cys103del) was located in the deubiquitinase domain. All three in-frame deletions lead to a single amino acid loss. According to the established clinical guidelines, these missense and inframe deletion variants should be classified as variants of uncertain significance^{9,12}. Further laboratory functional studies will be needed for their optimal clinical classification.

In addition to the five null variants identified (Tables 1), a *BAP1* nonsense variant, c.1203T>G, p.T401* (rs200156887) was surprisingly observed with a heterozygote population frequency of 0.0358% (19/53105) in the general population and at a lower frequency 0.0132 % (1/7601) in the TCGA subjects. Given the high frequency of this variant in the general population we further investigated it. Using the Catalogue of Somatic Mutation in Cancer (COSMIC) cell line database we identified the RT4 cell line with the heterozygous c.1203T>G, p.Y401* variant. The cell line also has the *BAP1* c.1201T>G, p.Y401D (rs376563004) variant. It should be noted that the two variants, rs200156887 and rs376563004, have the same allele frequency and distribution in various populations, suggesting that they occur together. RFLP analysis showed that the two variants are in cis to each other(Figure 1B). As such, we assert that these two *BAP1* variants are erroneously reported in the ExAC and they are in fact a single 3-bp deletion/insertion variant c.1201_1203delinsGAG, leading to a likely benign missense alteration p.Y401E. It should be noted that this variant is reported in ClinVar database in one patient. The relatively high frequency of this variant in the general population and the predicted benign impact by multiple computational tools suggests that this variant is benign.

Limitations of studying the ExAC data include the inability of carrying out traditional case/control matching. Some of the subjects in the general population with variants could be younger than the expected age of cancer onset and could develop cancer later in their life. Hence *BAP1* variants could be overreported in the general population. However, the 1/53015 carrier frequency of *BAP1* pathogenic variants observed highlights the rarity *BAP1*

mutations in general population. Another limitation of studying ExAC data is that clinical information including age of the subject, type of cancer and status of biallelic inactivation of *BAP1* in the tumor are not available. Further studies of the TCGA data, including functional assessment of *BAP1* in tumors from subjects with germline alteration, will be needed. Although this analysis suggests that overall germline *BAP1* alterations is underreported in cancer patients, clinical testing for *BAP1* has been introduced only in the last 2–3 years, so underreporting would be expected. The low numbers of pathogenic mutations observed and the very wide confidence interval between cancer and non-cancer subjects highlights the preliminary nature of our findings which needs independent validation.

In conclusion, germline null mutations in *BAP1* appear to have a significantly higher frequency in cancer patients than general population. Given the low frequency of reported families with *BAP1*-TPDS, our results suggest that germline pathogenic mutations in *BAP1* are underreported especially in patients with cancer.

Acknowledgments

Funding

The Patti Blow Research Fund in Ophthalmology, the Ohio Lions Eye Research Foundation, the R21CA191943 Grant from the National Cancer Institute (PI: Abdel-Rahman, MH), and the National Eye Institute grant K08EY022672 (CMC).

The authors would like to thank the Exome Aggregation Consortium and the groups that provided exome variant data for comparison. A full list of contributing groups can be found at <http://exac.broadinstitute.org/about>.

References

1. Rai K, Pilarski R, Cebulla CM, Abdel-Rahman MH. Comprehensive review of *BAP1* tumor predisposition syndrome with report of two new cases. *Clin Genet*. 2016; 89(3):285–294. [PubMed: 26096145]
2. Popova T, Hebert L, Jacquemin V, et al. Germline *BAP1* mutations predispose to renal cell carcinomas. *American journal of human genetics*. 2013; 92(6):974–980. [PubMed: 23684012]
3. Wiesner T, Obenaus AC, Murali R, et al. Germline mutations in *BAP1* predispose to melanocytic tumors. *Nature genetics*. 2011; 43(10):1018–1021. [PubMed: 21874003]
4. Testa JR, Cheung M, Pei J, et al. Germline *BAP1* mutations predispose to malignant mesothelioma. *Nature genetics*. 2011; 43(10):1022–1025. [PubMed: 21874000]
5. Abdel-Rahman MH, Pilarski R, Cebulla CM, et al. Germline *BAP1* mutation predisposes to uveal melanoma, lung adenocarcinoma, meningioma, and other cancers. *J Med Genet*. 2011; 48(12):856–859. [PubMed: 21941004]
6. Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*. 2016; 536(7616):285–291. [PubMed: 27535533]
7. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc*. 2009; 4(7):1073–1081. [PubMed: 19561590]
8. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr Protoc Hum Genet*. 2013; Chapter 7(Unit7):20.
9. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015; 17(5):405–423. [PubMed: 25741868]

10. Abdel-Rahman MH, Rai K, Pilarski R, Davidorf FH, Cebulla CM. Germline BAP1 mutations misreported as somatic based on tumor-only testing. *Fam Cancer*. 2016; 15(2):327–330. [PubMed: 26748926]
11. Pilarski R, Rai K, Cebulla C, Abdel-Rahman M. BAP1 Tumor Predisposition Syndrome. In: Adam MP, Ardinger HH, Pagon RA. , et al., editors *GeneReviews*((R)). Seattle (WA): 1993.
12. Li MM, Datto M, Duncavage EJ, et al. Standards and Guidelines for the Interpretation and Reporting of Sequence Variants in Cancer: A Joint Consensus Recommendation of the Association for Molecular Pathology, American Society of Clinical Oncology, and College of American Pathologists. *J Mol Diagn*. 2017; 19(1):4–23. [PubMed: 27993330]

Figure 1. Predicted pathogenic variants reported in ExAC and assessment of the, p.Y401* mutation

A) Distribution of null and missense variants predicted as pathogenic with at least one online prediction tool (SIFT and/or PolyPhen2). Missense variants showed rather similar distribution in TCGA and general population. Several missense variants, including one commonly reported, are seen in different functional domains of BAP1.

B) Sequencing and restriction fragment polymorphism of the RT4 cell line show that the p.Y401*(rs200156887) and p.Y401D (rs376563004) variants are in cis. Within our 429bp PCR product, *BfaI* enzyme has two naturally occurring cut sites, generating three bands at 222bp, 120bp, and 87bp. If variant p. Y401D, is in trans with p. Y401*, a third cut site in the 222bp band will be generated, creating two extra bands 141bp & 81bp in the heterozygote. If the two variants are in cis, then the banding will be identical to a wild type sample with only 3 bands.

Table 1

Summary germline BAP1 variants reported in the ExAC database

Mutation Type	ExAC All			ExAC without TCGA			TCGA	
	% (n)	Carrier Frequency	% (n)	Carrier Frequency	% (n)	Carrier Frequency	% (n)	Carrier Frequency
Null Variants [§]	0.0082 (5)	1/12141	0.00188 (1)	1/53015	0.0526 (4)			1/1900
Inframe deletions	0.00494 (3)	1/20235	0.00565 (3)	1/17701	0		0	0
Rare (<0.0001) missense variants ^{§§}	0.0659 (40)	1/1528	0.0602 (32)	1/1660	0.11 (8)			1/950
Null variants	Data set			Frequency	Reported in ClinVar	ACMG classification		
c.783+2T>C	TCGA			0.0132 (1)	Yes	Pathogenic		
c.375+1G>A	TCGA			0.0132 (1)	No	Likely pathogenic		
c.2050C>T, p.Gln684*	TCGA			0.0132 (1)	Yes	Pathogenic		
c.781C>T, p.Gln261*	TCGA			0.0132 (1)	No	Likely pathogenic		
c.1224_1224delA, A410HfsX2	ExAC without TCGA			0.0188 (1)	No	Likely pathogenic		

[§]Null variants include nonsense, frameshift, canonical ±1 or 2 splice sites

^{§§}Only missense variant predicted pathogenic by SIFT and/or PolyPhen are included

Notes:

- ExAC all data are for 60,706 individuals, this includes 7601 individuals from the TCGA and 53105 individuals from non-TCGA.
- An additional splicing variant is listed (3:52437816 C/T, c.119+1G>A) for a short transcript [ENST00000469613](#) and not the full length transcript [ENST00000460680](#). For the full length transcript it leads to a missense variant p.A449T predicted as benign.
- Variant c.1203T>G, p.Y401* (rs200156887) a misreported 3bp deletion insertion c.1201_1203delinsGAG, p.Y401E is excluded.