Original Article

Assessing the Spectrum of Germline Variation in Fanconi Anemia Genes Among Patients With Head and Neck Carcinoma Before Age 50

Settara C. Chandrasekharappa, PhD[1]; Steven B. Chinn, MD, MPH[2]; Frank X. Donovan, MS[1]; Naweed I. Chowdhury, MD[3]; Aparna Kamat, MSc[1]; Adebowale A. Adeyemo, MD[4]; James W. Thomas, PhD[5]; Meghana Vemulapalli, MS[5]; Caroline S. Hussey, BA[6]; Holly H. Reid, MD[7]; James C. Mullikin, PhD[5]; Qingyi Wei, MD, PhD[8]; and Erich M. Sturgis, MD, MPH [9,10]

BACKGROUND: Patients with Fanconi anemia (FA) have an increased risk for head and neck squamous cell carcinoma (HNSCC). The authors sought to determine the prevalence of undiagnosed FA and FA carriers among patients with HNSCC as well as an age cutoff for FA genetic screening. **METHODS:** Germline DNA samples from 417 patients with HNSCC aged <50 years were screened for sequence variants by targeted next-generation sequencing of the entire length of 16 FA genes. **RESULTS:** The sequence revealed 194 FA gene variants in 185 patients (44%). The variant spectrum was comprised of 183 nonsynonymous point mutations, 9 indels, 1 large deletion, and 1 synonymous variant that was predicted to effect splicing. One hundred eight patients (26%) had at least 1 rare variant that was predicted to be damaging, and 57 (14%) had at least 1 rare variant that was predicted to be damaging and had been previously reported. Fifteen patients carried 2 rare variants or an X-linked variant in an FA gene. Overall, an age cutoff for FA screening was not identified among young patients with HNSCC, because there were no significant differences in mutation rates when patients were stratified by age, tumor site, ethnicity, smoking status, or human papillomavirus status. However, an increased burden, or mutation load, of FA gene variants was observed in carriers of the genes FA complementation group D2 (*FANCD2*), *FANCE*, and *FANCL* in the HNSCC patient cohort relative to the 1000 Genomes population. **CONCLUSIONS:** FA germline functional variants offer a novel area of study in HNSCC tumorigenesis. *FANCE* and *FANCL*, which are components of the core complex, are known to be responsible for the recruitment and ubiquitination, respectively, of *FANCD2*, a critical step in the FA DNA repair pathway. In the current cohort, the increased mutation load of *FANCD2, FANCE*, and *FANCL* variants among younger patients with HNSCC indicates the importance of the FA pathway in HNSCC. ***Cancer* 2017;123:3943-54.** *© 2017 American Cancer Society.*

KEYWORDS: Fanconi anemia, germline variations, head and neck cancers, recessive inherited disorders, squamous cell carcinoma.

INTRODUCTION

Fanconi anemia (FA) is a rare, predominantly recessive, inherited disorder with an incidence of 1 in 130,000 births and an estimated carrier rate of 0.6%.[1,2] Genetically, FA is a heterogeneous disease with 21 causative genes known to date, including 5 new genes that were added within the past two years.[3,4] Phenotypically, FA is associated with congenital defects (short stature, renal defects, cafe-au-lait spots, microphthalmia, hearing difficulties, and abnormal thumb or radii) and progressive bone marrow failure. In addition to congenital anomalies and the inevitable bone marrow failure, patients with FA are at increased risk of acute myelogenous leukemia and head and neck squamous cell carcinoma (HNSCC).[5,6] However, 30% of patients fail to display these FA congenital defects,[7,8] and up to 25% present with solid or hematologic malignancies as the first sign of the condition.[9] In many such patients, an FA diagnosis is prompted after severe toxicities are encountered upon initiating chemotherapy or radiation therapy to treat the malignancy.[10] The pathogenesis is based

Corresponding authors: Settara C. Chandrasekharappa, PhD, Cancer Genetics and Comparative Genomics Branch, Building 50, Room 5232, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD 20892; chandra@mail.nih.gov; or Erich M. Sturgis, MD, Department of Head and Neck Surgery, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard, Unit 1445, Houston, TX 77030; Fax: (713) 794-4662; esturgis@mdanderson.org

[1]Cancer Genetics and Comparative Genomics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland; [2]Department of Otolaryngology-Head and Neck Surgery, The University of Michigan, Ann Arbor, Michigan; [3]Department of Otolaryngology-Head and Neck Surgery, University of Kansas Medical Center, Kansas City, Kansas; [4]Center for Research on Genomics and Global Health, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland; [5]Intramural Sequencing Center, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland; [6]The University of Texas Health Science Center School of Medicine, Houston, Texas; [7]Department of Dermatology, The University of Texas Health Science Center School of Medicine, Houston, Texas; [8]Department of Medicine, Duke University School of Medicine, Durham, North Carolina; [9]Department of Head and Neck Surgery, The University of Texas MD Anderson Cancer Center, Houston, Texas; [10]Department of Epidemiology, The University of Texas MD Anderson Cancer Center, Houston, Texas.

The second and third authors contributed equally to this work.

We thank Stephanie P. Deming for assistance with editing the article.

Additional supporting information may be found in the online version of this article.

DOI: 10.1002/cncr.30802, **Received:** December 28, 2016; **Revised:** March 27, 2017; **Accepted:** April 24, 2017, **Published online** July 5, 2017 in Wiley Online Library (wileyonlinelibrary.com)

on the finding that the normal function of FA genes is related to DNA repair and genome stability; biallelic mutations of these genes confer an annual risk of cancer from 0.7% to 2%,[5,9,11] and some, including FA complementation group D1 (*FANCD1*) (breast cancer 2 [*BRCA2*]) and *FANCN* (partner and localizer of BRCA2 [*PALB2*]), are closely associated with solid organ malignancies[12-14]

HNSCC is traditionally related to tobacco and alcohol consumption and has recently been associated with human papillomavirus, and exposed individuals have a 10-fold to 15-fold increased risk of HNSCC compared with unexposed individuals.[15,16] However, FA confers a 500-fold to 700-fold increased risk of HNSCC compared with the normal population.[9,17,18] The median age of HNSCC onset in patients with FA is 33 years compared with 60 years in traditional patients with HNSCC; thus, patients with FA develop HNSCC at a significantly younger age than those with sporadic HNSCC.[19]

Younger patients with HNSCC have reduced DNA repair capacity, but the prevalence of FA mutations in patients with HNSCC remains unknown.[20] Because 30% of patients with FA mutations do not display the congenital stigmata of FA, we sought to determine: 1) the prevalence of undiagnosed FA among patients with HNSCC aged <50 years, 2) an age cutoff for FA screening among younger patients with HNSCC, and 3) the prevalence of FA carriers (heterozygote germline mutations) among younger patients with HNSCC.

MATERIALS AND METHODS

DNA Extraction

DNA was isolated from blood using the Puregene kit and DNeasy blood and tissue DNA extraction kit (Qiagen, Hilden, Germany) and was subjected to phenol/chloroform extraction and ethanol precipitation.

Samples for Sequencing

In total, 647 patients aged < 50 years with HNSCC were enrolled in a prospective molecular epidemiologic study of newly diagnosed HNSCC that included completion of a prospective, standardized, epidemiologic questionnaire and blood draw. Patients who had cancer (including known FA) before an HNSCC diagnosis were excluded. The Institutional Review Board of the University of Texas MD Anderson Cancer Center approved this study, and all patients provided informed consent.

Of the 617 patients initially enrolled for the study (Supporting Fig. 1; see online supporting information), DNA was available from 468 patients. DNA from 417 of 468 patients was of sufficient quality and quantity required for the targeted capturing and sequencing approach. Thus, DNA samples from 417 patients were sequenced. There was no difference between the sequenced group (417 patients) and the nonsequenced group (230 patients) in terms of age, sex, ethnicity, smoking, and alcohol drinking (Supporting Table 1; see online supporting information). There was a higher proportion of laryngeal/hypopharyngeal disease among those who were excluded and a higher proportion of oral cavity disease in those who were included.

Targeted Next-Generation Sequencing

The genomic regions representing the entire length of 16 FA genes were targeted for capturing and sequencing (Supporting Table 2; see online supporting information). The exceptions were the exon 1 regions of *FANCA, FANCB*, and *FANCE*, which were not covered by the sequencing approach. Excessive repeat sequences prevented the successful design of probes for exon 1 of *FANCA*. Probes were designed for the *FANCB* and *FANCE* exon 1 regions but did not yield product. The targeted TruSeq (Illumina Inc., San Diego, Calif) capturing design, capture, and sequencing were done as previously described.[21]

Variant Calling and Filtering

Novoalign (http://www.novocraft.com/products/novoalign/, Accessed May 24, 2017) was used for sequence alignment. The Most Probably Genotype (MPG) genotype caller (https://research.nhgri.nih.gov/software/bam2mpg/, Accessed May 24, 2017; National Human Genome Research Institute Genome Technology Branch) was used to call variants, and SnpEff[22] was used to filter the high-quality functional variants based on the following criteria: quality score ≥ 20, read depth ≥ 10, nonsense, missense, indel, and splicing (± 2 base pairs). The extracted functional variants were subsequently annotated and filtered using a population-specific maximum frequency, where applicable, of 0.5% in each of the 1000 Genomes (2504 individuals), the National Heart, Lung, and Blood Institute-Exon Sequencing Project 6500 (NHLBI-ESP6500) (6503 individuals), and the Exome Aggregation Consortium (ExAC) non-The Cancer Genome Atlas (60,706 individuals) variant databases. The allele frequency threshold of 0.5% is derived from the FA carrier frequency of 1:181.[2] All coordinates are in accord with Human Genome Build 19.

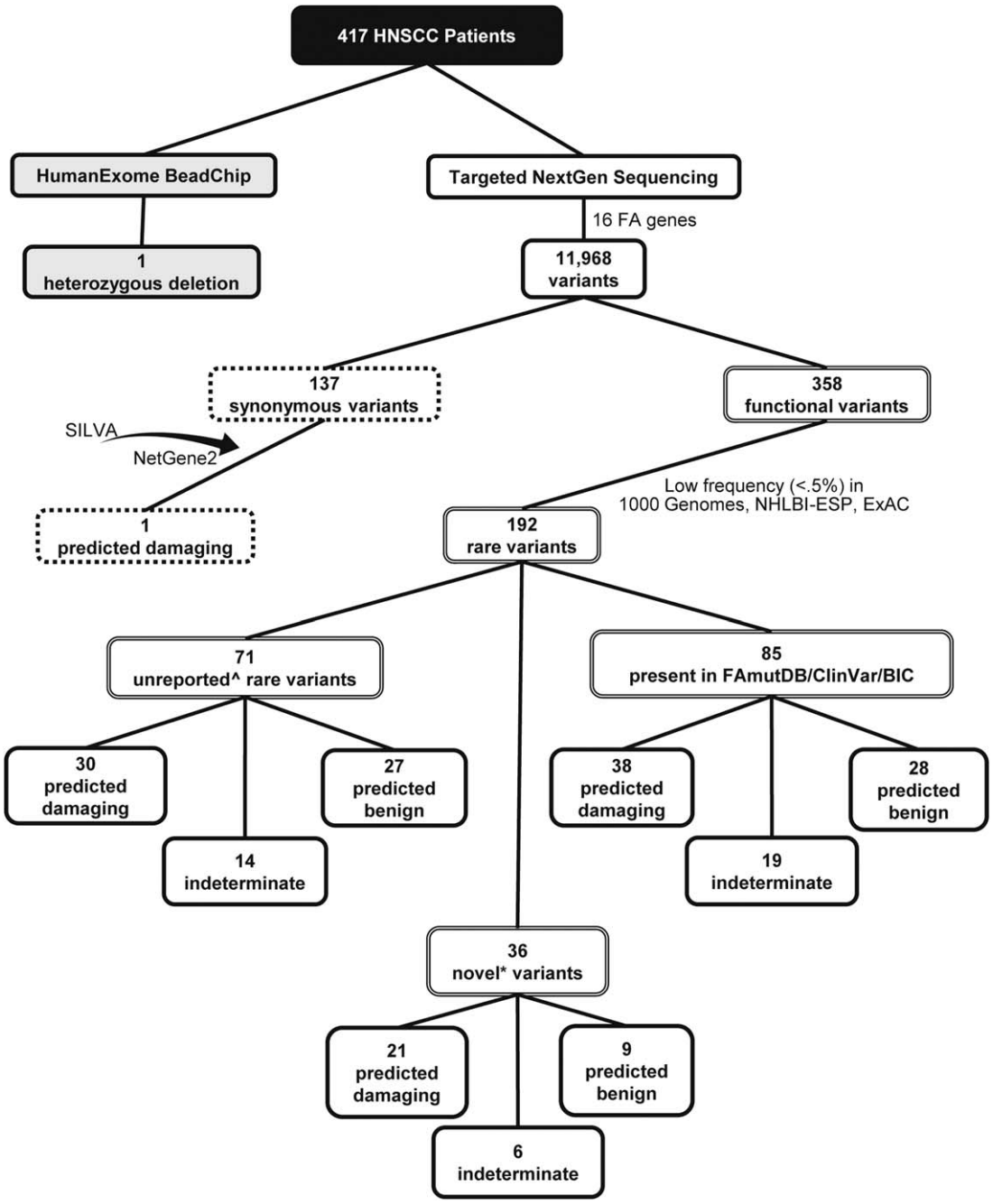

Figure 1. Schematics illustrate Fanconi anemia (FA) gene variant analysis in patients with head and neck squamous cell carcinoma (HNSCC). Data from a single nucleotide polymorphism array were analyzed for copy number variations. Single nucleotide variant discovery was performed from targeted sequencing of 16 FA genes. Synonymous variants were analyzed for splicing effects. Indels and nonsynonymous and splicing variants were screened for quality and filtered using population-specific frequencies, where applicable. The resulting rare variants were screened for presence in the FA Mutation Database (FAmutDB) and ClinVar and were analyzed by functional prediction algorithms to determine potential pathogenicity (^unreported variants were present at a frequency below 0.5% in the public databases but were not present in the FAmutDB, ClinVar, or Breast Cancer Information Core [BIC] databases; *novel variants were not present at any frequency in the public databases used for filtering).

Predicted Damaging Variants

To predict the functional consequence of a variant, we compared the results from 5 prediction algorithms: Scale-Invariant Feature Transform (SIFT) (http://sift.jcvi.org, Accessed May 24, 2017), Polymorphism Phenotyping v2 (PolyPhen-2) (http://genetics.bwh.harvard.

edu/pph2, Accessed May 24, 2017), MutationTaster (http://www.mutationtaster.org/, Accessed May 24, 2017), Combined Annotation-Dependent Depletion (CADD) (http://cadd.gs.washington.edu, Accessed May 24, 2017), and Genomic Evolutionary Rate Profiling 2 (GERP++) (http://mendel.stanford.edu/SidowLab/downloads/gerp/). We used a minimum threshold of 20 for Phred-scaled CADD scores, representing the 1% most damaging variants in the genome, and a minimum threshold of 2 for GERP++ rejected substitution scores. GERP++ was included to highlight constrained sites. Annovar[23] was used to annotate the variants with the results from all 5 prediction algorithms. For consideration as a "damaging" variant, at least 4 of the 5 algorithms had to meet their specified threshold. "Benign" variants did not meet the threshold in at least 4 of the 5 algorithms. Variants lacking consensus between at least 4 algorithms were labeled as "indeterminate" (Fig. 1, Supporting Table 3; see online supporting information).

Previously Reported Variants

The following databases were used to identify previously reported variants: The Leiden Open Variation Database (LOVD) for FA (http://www.rockefeller.edu/fanconi/, Accessed May 24, 2017) the ClinVar National Center for Biotechnology Information database (http://www.ncbi.nlm.nih.gov/clinvar/, Accessed May 24, 2017), and Breast Cancer Information Core (BIC).

Synonymous Variants

Synonymous variants with quality scores $\geq$20 and a read depth $\geq$10 were analyzed using SILVA v1.1.1.[24] Variants that were determined to be "potentially pathogenic" by SILVA were further analyzed using NetGene2[25] to predict splicing effects.

Single Nucleotide Polymorphism Array Analysis

The patient DNA samples were run on the Illumina HumanExome BeadChip, which contains approximately 250,000 single nucleotide polymorphisms. The data were processed using GenomeStudio (Illumina, Inc.), and copy number variants were detected using cnvPartition v3.2 (Illumina, Inc.) and Nexus v7.5 (BioDiscovery, Inc., El Segundo, Calif).

Statistical Analysis of Mutation Load

The Mann-Whitney-Wilcoxon nonparametric statistical test was used to evaluate the burden or mutation load of FA gene variants in the HNSCC patient cohort compared with the 1000 Genomes data set. Because the HNSCC cohort was comprised predominantly of patients of Caucasian ethnicity (Table 1), the statistical test was performed using data from the 356 Caucasian patients with HNSCC and the 503 EUR 1000 Genomes individuals (of European ancestry) to create a more homogenous comparison group. We obtained the nonsynonymous and indel variant alleles from each set and implemented the test in R (R Foundation for Statistical Computing, Vienna, Austria).[26]

Low-frequency Variants

LoFreq[27] was implemented to call low-frequency variants occurring between 5% and 40% with a genotype quality >500.

RESULTS

Of the 417 patients who had DNA available for sequencing, 88 (21%) were aged < 40 years (a traditionally accepted definition of "young" for a patient with HNSCC), 108 (26%) were ages 40 to 44 years, and 221 (53%) were ages 45 to 49 years. Tumor site was the oral cavity in 149 patients (36%), the oropharynx in 230 (55%), and the larynx in 38 (9%). The cohort was comprised of 4 different ethnic populations, 356 (85%) are Caucasian, 40 (10%) are Hispanic, 14 (3%) are Asian, and 7 (2%) are African American (Table 1).

No patient in the cohort had a known diagnosis of FA, and there were very few with any potential signs of an FA phenotype. Classic phenotypes of FA were then evaluated in the cohort. Seventeen patients (4%) had a first-degree relative with a hematologic malignancy, 25 (6%) had short stature (<5th percentile), and 8 (2%) had macrocytic anemia and/or leukopenia. Among the patients who had received chemotherapy, there were no grade IV toxic effects.

Germline DNA targeted capturing and next-generation sequencing of 16 FA genes revealed 11,968 initial variants. The targeted region for capturing and sequencing included the entire gene. The postsequence coverage of high-quality sequence, particularly for the entire coding region of 56,120 base pairs, was 100% for all genes except for a total of 287 base pairs from exon 1 of *FANCA, FANCB* and *FANCE*, with an approximate depth of coverage of 240 reads at each base (Supporting Table 2; see online supporting information).[21]

Among the initial 11,968 called variants, there were 137 synonymous variants and 358 functional variants (nonsense, missense, indel, or splicing) (Fig. 1). By using a population-specific variant frequency threshold of $\leq$0.5% in each of the 1000 Genomes, NHLBI-ESP6500, and ExAC databases, the subset of functional variants was

TABLE 1. Distribution of 417 Patients With Head and Neck Squamous Cell Carcinoma by Cancer Site, Age Group, and Ethnicity

	Cancer Site and Age Group									
	Oral Cavity, n = 149			Oropharynx, n = 230			Larynx, n = 38			
Ethnicity	<40 y	40-44 y	45-49 y	<40 y	40-44 y	45-49 y	<40 y	40-44 y	45-49 y	Total
Caucasian	40	32	45	23	54	129	3	9	21	356
African American	0	0	2	0	1	1	1	0	2	7
Hispanic	8	4	8	7	5	7	0	0	1	40
Asian	3	4	3	1	0	2	0	0	1	14
Total	51	40	58	31	60	139	4	9	25	417

further reduced to 192 (183 single nucleotide variants [SNVs] and 9 indels). Functional prediction algorithms (SIFT, PolyPhen-2, MutationTaster, CADD, and GERP++) analyzed the 183 rare SNVs to identify which are likely to induce deleterious functional consequences and which are likely benign. Assignment of a definitive prediction to a particular variant depended on at least 4 of the 5 algorithms reaching a consensus. If a consensus was not met between at least 4 algorithms, then the prediction was classified as indeterminate and labeled as such. Eighty rare SNVs were predicted to be damaging, and 64 were predicted to be benign, whereas 39 were indeterminate. The 9 indels were presumed to be deleterious by the nature of the variant. Splicing prediction algorithms (SILVA and NetGene2) analyzed the 137 synonymous variants and identified 1 variant as likely pathogenic by creating a new donor site.

The 193 resulting variants (184 SNVs and 9 indels) were compared with variants that had been previously reported to the LOVD FA disease database, BIC, and/or listed in ClinVar. Eighty-five of 193 variants (44%) had been previously reported in at least 1 of the 3 databases, but only 5 of 85 were listed specifically as pathogenic, whereas the rest were listed as either benign or of uncertain significance. Variants in *BRCA2*(36) and *PALB2*(9) comprised 52% (45 of 85) of the reported variants. In addition to the 85 reported variants, our cohort carried 38 completely novel variants and 71 variants that were unreported but were present at a frequency below 0.5% in the public databases.

In addition to high-throughput sequencing of the 16 FA genes, genotype data were collected by single nucleotide polymorphism array (HumanExome BeadChip; Illumina, Inc.) for all 417 patients with HNSCC. Copy number analysis revealed an approximately 154-kb heterozygous deletion of *SLX4*, spanning the entire gene, in 1 patient with HNSCC. This patient also carried an *SLX4* missense mutation that was predicted to be damaging (Table 2).

We also implemented LoFreq to call low-frequency variants (between 5% and 40%) that had quality scores >500 and a minor allele frequency <0.5% in public databases. It is known that revertant mosaicism in FA may result in the loss of a variant or a variant being present to a lower extent; and, in such patients, cancer may precede the diagnosis of FA.[28] However, our search for low-frequency variants did not yield any potential candidates.

Of all 417 patients in the cohort, 185 (44%) carried at least 1 rare variant (Table 3). These 185 patients carried 194 variations (192 rare variants, 1 large deletion, and 1 synonymous variant that was predicted to affect splicing). A rare variant was observed in 42% to 55% of patients from Caucasian, Hispanic, and African American ancestry, whereas 11 of 14 Asian patients (79%) carried a variant. Rare variants were identified in similar proportions of patients irrespective of tumor site (larynx, 42%; oral cavity, 44%; or oropharynx, 45%) or age group (<40 years, 40%; 40-44 years, 49%; or 45-49 years, 44%) (Tables 3). The proportion of patients carrying rare predicted damaging variants also was similar irrespective of tumor site (range, 25%-28%) or age group (range, 24%-28%) (Tables 4 and 5).

Fifteen patients (4%) either had 2 rare FA variants in the same FA gene or had an X-linked variant of FA (Table 2). Among these 15 patients, the median age was 45 years (only 2 were younger than 40 years), 10 patients (67%) had oropharyngeal primary tumors, 8 (47%) were never smokers, 2 (13%) had a first-degree relative with a hematologic malignancy, 5 (33%) had short stature (<5th percentile), and 3 (20%) had macrocytic anemia and/or leukopenia. Relative to the entire cohort, there were no differences in age, sex, smoking status, tumor site, or human papillomavirus status between patients who did or did not carry either 2 rare variants in the same FA gene or an X-linked FA variant. Of the 15 patients who carried 2 rare variants or an X-linked variant, 6 carried variants that had been previously documented either in the FA mutation database, ClinVar, or BIC.

TABLE 2. Patients Carrying 2 Variants in 1 Fanconi Anemia (FA) Gene or an X-Linked Variant of FA

									Mutation 1				Mutation 2			
Patient	Ethnicity	Sex	Age, y	Cancer Site	FA Phenotypes	Smoking	HPV	Gene	cDNA	Protein	In FA Database[a]	Functional Prediction[b]	cDNA	Protein	In FA Database[a]	Functional Prediction[b]
A5421[c]	C	Man	46	OP	None	Never	Positive	*BRCA2*	c.1792A>G	p.T598A	Yes	Benign	c.1804G>A	p.G602R	Yes	Benign
A5809	C	Woman	45	OP	None	MD	Positive	*BRCA2*	c.8573A>G	p.Q2858R	Yes	Damaging	c.1151C>T	p.S384F	Yes	Damaging
A1105	C	Man	44	OP	None	Never	MD	*ERCC4*	c.1336G>T	p.A446S	—	Indeterminate	c.1347C>A	p.V449V[d]	—	Damaging
A4798	C	Man	42	OP	None	Current	Positive	*FANCA*	c.1046C>T	p.A349V	—	Indeterminate	c.2390C>T	p.A797V	—	Benign
A4675	C	Man	43	OP	An	Never	Positive	*FANCB*[e]	c.30C>A	p.N10K	—	Benign	—	—	—	—
A2766	H	Man	31	OC	None	Never	MD	*FANCI*	c.868G>A	p.V290M	—	Benign	c.1114G>A	p.V372I	—	Indeterminate
A3494	H	Man	47	OP	An	Never	MD	*FANCI*	c.3493delG	p.D1165fs	Yes	Damaging	c.3946G>A	p.G1316R	—	Benign
A4164	C	Man	49	OC	None	Never	MD	*FANCI*	c.1461T>A	p.Y487X	Yes	Damaging	c.362T>C	p.L121P	—	Damaging
A4741	C	Woman	33	OC	None	Former	MD	*FANCM*	c.5117A>C	p.N1706T	—	Benign	c.3827C>T	p.S1276L	—	Benign
A2217	C	Man	46	OP	L and An	Current	MD	*SLX4*	c.3739G>A	p.E1247K	—	Damaging	c.833G>A	p.R278Q	—	Benign
A2281	C	Man	43	OP	None	Former	MD	*SLX4*	c.2182G>A	p.A728T	—	Damaging	c.5281C>T	p.R1761C	—	Damaging
A3094	C	Woman	43	OC	Ht	Current	Positive	*SLX4*	c.4264C>G	p.P1422A	—	Indeterminate	c.2364G>C	p.Q788H	—	Benign
A4325	C	Woman	49	OP	None	Never	MD	*SLX4*	c.3368C>A	p.S1123Y	Yes	Damaging	Large deletion[f]	—	—	Damaging
A5423	H	Man	47	OP	None	Former	Positive	*SLX4*	c.4261A>T	p.I1421F	Yes	Indeterminate	c.2290C>G	p.P764A	—	Benign
A2674	AA	Woman	49	Lar	Ht	Current	MD	*ERCC4*	c.109C>T	p.R37C	—	Damaging	c.109C>T	p.R37C	—	Damaging

Abbreviations: AA, African American; An, anemia; *BRCA2*, breast cancer 2; C, Caucasian; cDNA, complementary DNA; *ERCC4*, excision repair 4, endonuclease catalytic subunit; *FANCA* through *FANCM*, Fanconi anemia complementation groups A through M, respectively; H, Hispanic; Ht, short stature; HPV, human papillomavirus; L, leukopenia; Lar, larynx; MD, missing data; OP, oropharynx; *SLX4*, SLX4 structure-specific endonuclease subunit.

[a] The variation was previously reported in the Leiden Open Variation Database for Fanconi anemia (http://www.rockefeller.edu/fanconi/, Accessed May 24, 2017) and/or ClinVar (http://www.ncbi.nlm.nih.gov/clinvar/, Accessed May 24, 2017).

[b] Functional prediction consensus results are from the algorithms Scale-Invariant Feature Transform (SIFT), Polymorphism Phenotyping (PolyPhen), MutationTaster, Combined Annotation-Dependent Depletion (CADD), and Genomic Evolutionary Rate Profiling 2 (GERP++); synonymous variants were analyzed by SILVA and NetGene2; nonsense mutations, indels, and the large deletion were presumed to be damaging by nature of the variant.

[c] This patient carried a third mutation in *BRCA2* that also was predicted to be benign (c.125A>G; p.Y42C).

[d] This was a synonymous variant that was predicted to effect splicing.

[e] *FANCB* is X-linked.

[f] This *SLX4* deletion removes the entire gene (deletion coordinates: chr16_3586230-3740926).

TABLE 3. Distribution of 185 Patients With Head and Neck Squamous Cell Carcinoma Who Carried a Rare Fanconi Anemia Gene Variant by Cancer Site, Age Group, and Ethnicity

	Cancer Site and Age Group										
	Oral Cavity, n = 66			Oropharynx, n = 103			Larynx, n = 16				
Ethnicity	<40 y	40-44 y	45-49 y	<40 y	40-44 y	45-49 y	<40 y	40-44 y	45-49 y	Total No.	% of entire cohort
Caucasian	16	11	22	4	29	55	2	4	6	149	42
African American	0	0	0	0	0	0	1	0	2	3	43
Hispanic	5	3	2	4	3	5	0	0	0	22	55
Asian	1	3	3	1	0	2	0	0	1	11	79
Total	22	17	27	9	32	62	3	4	9	185	
% of entire cohort		44			45			42			

TABLE 4. Distribution of 108 Patients With Head and Neck Squamous Cell Carcinoma Who Carried a Rare Predicted Damaging Fanconi Anemia Gene Variant by Cancer Site, Age Group, and Ethnicity

	Cancer Site and Age Group										
	Oral Cavity, n = 41			Oropharynx, n = 57			Larynx, n = 10				
Ethnicity	<40 y	40-44 y	45-49 y	<40 y	40-44 y	45-49 y	<40 y	40-44 y	45-49 y	Total No.	% of entire cohort
Caucasian	10	7	13	1	15	32	2	1	4	85	24
African American	0	0	0	0	0	0	0	0	2	2	29
Hispanic	4	3	0	3	3	2	0	0	0	15	38
Asian	1	2	1	0	0	1	0	0	1	6	43
Total	15	12	14	4	18	35	2	1	7	108	
% of entire cohort		28			25			26			

TABLE 5. Percentage of Patients With Head and Neck Squamous Cell Carcinoma Carrying Rare Variants and Rare Predicted Damaging Variants in Fanconi Anemia Genes Segregated by Age Group

Age Group, y	Total No. of Patients	No. Carrying a Rare Variant (%)	No. Carrying a Predicted Damaging Rare Variant (%)
<40	86	34 (40)	21 (24)
40-44	109	53 (49)	31 (28)
45-49	222	98 (44)	56 (25)
Total	417	185 (44)	108 (26)

Thirty-nine patients carried a rare variant in 2 or more different FA genes (Supporting Table 4; see online supporting information). When comparing the age at presentation between patients with and without multiple variants, there was no difference in age, sex, smoking, alcohol use, or human papillomavirus status.

The 194 germline variations were comprised of 176 missense variants, 9 indels, 6 nonsense variants (including 1 stop loss), 2 splicing variants (including 1 synonymous), and 1 large deletion (Fig. 2, Supporting Table 3; see online supporting information). These variants amounted to 255 occurrences throughout the HNSCC cohort (Table 6). *BRCA2, FANCP, FANCM, FANCA*, and *FANCI* were the most common genes to carry rare variants (21% [54 of 255 variants], 14% [36 of 255 variants], 11% [27 of 255 variants], 9% [22 of 255 variants], and 8% [20 of 255 variants], respectively). Twenty-six percent of patients (108 of 417) carried 91 FA rare variants that were predicted to be damaging: *BRCA2* had the highest proportion of occurrences at 10% (25 of 255 variants); and *SLX4, FANCI, FANCM, FANCQ* each accounted for 5% (12-14 of 255 variants). Fourteen percent of patients (57 of 417) had rare variants that were predicted to be damaging and had previously been reported. *BRCA2*, with 10% (25 of 255 variants), and *SLX4* and *FANCQ*, with 3% each (7-8 of 255 variants), are the top 3 most prevalent carriers. Detailed characterization of each of the 194 rare variants is presented in Supporting Table 3 (see online supporting information).

Table 7 presents the incidence of rare variants by FA gene among the HNSCC cohort, segregated by tumor

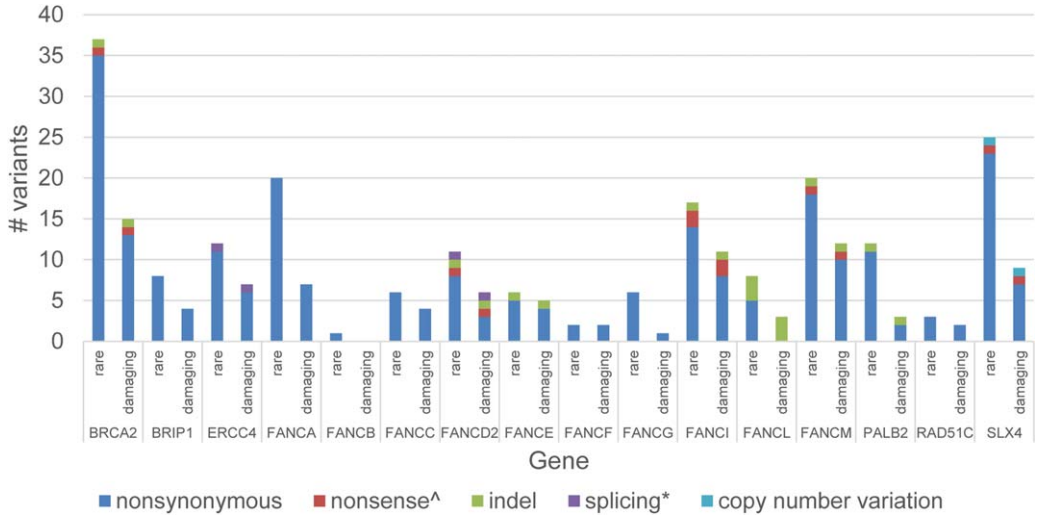

Figure 2. The 194 rare variants observed in 16 Fanconi anemia (FA) genes from 185 of 417 patients with head and neck squamous cell carcinoma are illustrated by mutation type. The plot illustrates the number and type of rare variants observed in each FA gene along with the number of rare variants that were predicted to be damaging (^includes 1 stop loss; *includes 1 synonymous variant). *BRCA2* indicates breast cancer 2; *BRIP1*, BRCA-interacting protein C-terminal helicase 1; *ERCC4*, excision repair 4, endonuclease catalytic subunit; *FANCA* through *FANCM*, Fanconi anemia complementation group genes A through M, respectively; *PALB2*, partner and localizer of BRCA2 (*FANCN*); *RAD51C*, RAD51 paralog C; *SLX4*, SLX4 structure-specific endonuclease subunit.

TABLE 6. Variant Counts for Each Fanconi Anemia Gene and the Number of Occurrences in the Current Head and Neck Squamous Cell Carcinoma Cohort Characterized as Rare, Predicted Damaging, and Predicted Damaging and Reported

	Rare Variants, n = 185[a]		Predicted Damaging Variants, n = 108[b]		Predicted Damaging and Reported Variants, n = 57[c]	
Gene	No. of Variants	No. of Occurrences	No. of Variants	No. of Occurrences	No. of Variants	No. of Occurrences
FANCA	20	22	7	9	2	2
FANCB	1	1	0	0	0	0
FANCC	6	8	4	5	3	4
FANCD1/BRCA2	37	54	15	25	15	25
FANCD2	11	11	6	6	2	2
FANCE	6	6	5	5	1	1
FANCF	2	2	2	2	0	0
FANCG	6	6	1	1	0	0
FANCI	17	20	11	14	2	2
FANCJ/BRIP1	8	13	4	5	3	4
FANCL	8	13	3	8	0	0
FANCM	20	27	12	13	0	0
FANCN/PALB2	12	15	3	3	2	2
FANCO/RAD51C	3	4	2	3	2	3
FANCP/SLX4	25	36	9	14	3	8
FANCQ/ERCC4	12	17	7	12	3	7
Totals	194	255	91	125	38	60

Abbreviations: *BRCA2*, breast cancer 2; *BRIP1*, BRCA interacting protein C-terminal helicase 1; *ERCC4*, excision repair 4, endonuclease catalytic subunit (Fanconi anemia complementation group Q [*FANCQ*]); *FANCA* through *FANCM*, Fanconi anemia complementation groups A through M; *PALB2*, partner and localizer of BRCA2 (*FANCN*); *RAD51C*, RAD51 paralog C (*FANCO*); *SLX4*, SLX4 structure-specific endonuclease subunit (*FANCP*).

[a] Rare indicates that the variant exists at a frequency <0.5% in the 1000 Genomes; National Heart, Lung, and Blood Institute-Exon Sequencing Project 6500; or Exome Aggregation Consortium database.

[b] Predicted damaging variants are determined according to prediction algorithm (criteria): Scale-Invariant Feature Transform (SIFT) (Damaging), Polymorphism Phenotyping v2 (PolyPhen-2) (Damaging or Probably damaging), MutationTaster (Disease causing or Disease causing automatic), Combined Annotation-Dependent Depletion (CADD) (Phred-scaled score > 20), or Genomic Evolutionary Rate Profiling 2 (GERP++) (rejected substitution score > 2). If 4 of 5 of the required criteria were met, then the variant was considered to be "predicted damaging."

[c] Predicted damaging and reported indicates that the criteria for predicted damaging variants were met and the variant was present in either the FA mutation database and/or ClinVar (a National Center for Biotechnology Information database).

TABLE 7. Distribution of Rare Fanconi Anemia Gene Variants in Patients With Head and Neck Squamous Cell Carcinoma Segregated by Cancer Site, Age of Onset, and Ethnicity

	Cancer Site, Age Group, and Ethnicity																														
	Oral Cavity, n = 149											Oropharynx, n = 230											Larynx, n = 38								
	<40 y, n = 51			40-44 y, n = 40			45-49 y, n = 58					<40 y, n = 31			40-44 y, n = 60			45-49 y, n = 139					<40 y, n = 4		40-44 y, n = 9	45-49 y, n = 25					
Variable	A	H	C	A	H	C	A	H	C	AA	Total	A	H	C	H	C	AA	A	H	C	AA	Total	AA	C	C	A	AA	C	H	Total	All
Patient totals																															
No. of patients	3	8	40	4	4	32	3	8	45	2		1	7	23	5	54	1	2	7	129	1		1	3	9	1	2	21	1		
No. with rare variants	1	5	16	3	3	11	3	2	22	0	66	1	4	4	3	29	0	2	5	55	0	103	1	2	4	1	2	6	0	16	185
No. of FA gene rare variants																															
BRCA2/FANCD1		2	6	2		2	1		6		19		2	1		10		1		18		32			1	1		1		3	54
FANCJ/BRIP1		1	1			2	1		1		6					2				3		5				1	1			2	13
FANCQ/ERCC4			1						2		3	1				4				3		8		1	1		2	2		6	17
FANCA		1	1	1		2			2		7				1	4		1		8		14			1					1	22
FANCB																1						1									1
FANCC		1				1			2		4				1	2						3			1					1	8
FANCD2		1				1	1		2		5				1					4		5			1					1	11
FANCE			2								2		1			1				1		3					1			1	6
FANCF				1							1									1		1									2
FANCG	1	1					1				3					2				1		3									6
FANCI		2	1		1				3		7			1	1	1			3	5		11						2		2	20
FANCL	1	1	1			2					5		1	1	1	2			1	2		8									13
FANCM		1	6			2		1	6		16					2				7		9				1	1			2	27
FANCN/PALB2					2			1	1		4					1			1	5		7	1					3		4	15
FANCO/RAD51C	1		1						1		3									1		1									4
FANCP/SLX4	1	1	2			4	1		3		12			1		6			2	12		21		1	1			1		3	36
Variant totals																															
Ethnicity	4	12	22	4	3	16	5	2	29	0		1	4	4	5	38	0	2	7	71	0		1	2	6	3	5	9	0		255
Age group	38			23			36					9			43			80					3		6		17				
Cancer site	97											132											26								
Total occurrences	255																														

Abbreviations: A, Asian; AA, African American; *BRCA2*, breast cancer 2; *BRIP1*, BRCA interacting protein C-terminal helicase 1; C, Caucasian; *ERCC4*, excision repair 4, endonuclease catalytic subunit; *FANCA* through *FANCM*, Fanconi anemia complementation group genes A through M; H, Hispanic; *PALB2*, partner and localizer of BRCA2 (*FANCN*); *RAD51C*, RAD51 paralog C (*FANCO*); *SLX4*, SLX4 structure-specific endonuclease subunit (*FANCP*).

TABLE 8. Analysis of Fanconi Anemia Gene Mutation Burden in Patients With Head and Neck Squamous Cell Carcinoma

FA Gene	Adjusted *P*	HNSCC Direction
FANCA	0.93744	—
FANCB	2.8672	—
FANCC	15.3552	—
FANCD1 [BRCA2]	3.52E-15	↓
FANCD2	3.52E-15	↑
FANCE	0.0081088	↑
FANCF	1.55936	—
FANCG	3.52E-15	↓
FANCI	11.1664	—
FANCJ [BRIP1]	2.5776	—
FANCL	3.52E-15	↑
FANCM	8.1728	—
FANCN [PALB2]	6.2352	—
FANCO [RAD51C]	14.776	—
FANCP [SLX4]	1.7248	—
FANCQ [ERCC4 or *XPF]*	3.52E-15	↓

Abbreviations: *BRCA2*, breast cancer 2; *BRIP1*, BRCA interacting protein C-terminal helicase 1; *ERCC4*, excision repair 4, endonuclease catalytic subunit (Fanconi anemia complementation group Q [*FANCQ*]); FA, Fanconi anemia; *FANCA* through *FANCM*, Fanconi anemia complementation group genes A through M; HNSCC, head and neck squamous cell carcinoma; *PALB2*, partner and localizer of BRCA2 (*FANCN*); *RAD51C*, RAD51 paralog C (*FANCO*); *SLX4*, SLX4 structure-specific endonuclease subunit (*FANCP*); *XPF*, DNA repair endonuclease XPF.

site, ethnicity, and age group. There were no significant differences in mutation rates when patients were stratified by age, tumor site, ethnicity, smoking status, or human papillomavirus status.

Finally, we searched for an increased burden or mutation load of FA gene variants in our cohort compared with the 1000 Genomes data set using Mann-Whitney-Wilcoxon nonparametric tests. Because our cohort is comprised predominantly of patients with Caucasian ethnicity, the statistical comparison was between the 356 Caucasian patients with HNSCC and the 503 individuals of European ancestry in the 1000 Genomes data set. This comparison revealed that *FANCD2, FANCE*, and *FANCL* had a significantly increased mutation burden in our cohort (Table 8). At the same time, the mutation burden for *BRCA2, FANCG*, and *FANCQ* was significantly reduced.

DISCUSSION

Whereas HNSCC is highly associated with FA, the prevalence of FA germline variants in younger HNSCC populations has not been explored. Although the primary purpose of this study was to define an age cutoff for genetic FA screening among patients with HNSCC, there was no correlation of younger age with FA germline variants. In a previously reported tumor genomic analysis of HNSCC in low-risk patients (nonsmokers aged < 45 years) and traditional high-risk patients (smokers aged > 45 years), age was not a marker of genome instability: rates of gene-specific mutations and copy number alterations were similar in oral tongue cancers among low-risk patients and oral tongue cancers among traditional high-risk patients.[29] Similarly, we did not observe a difference in the prevalence of FA germline mutations between patients ages <40, 40-44, and 45-50 years or between smokers and nonsmokers.

Our patient cohort consisted of 4 different ethnicities, so it was important to use an approach that involved population-specific frequencies to ensure that low frequency variants were properly characterized. One hundred eighty-five patients in our cohort carried 194 total variations, including 183 rare nonsynonymous variants, 1 synonymous variant that was predicted to affect splicing, 9 indels, and 1 large deletion (Fig. 1). One hundred eight patients carried at least 1 variant that was predicted to be damaging. We identified 38 novel FA germline variants, including 1 synonymous variant and a large deletion, that were not present in control populations or had not previously been documented in disease-associated databases, of which 23 (61%) were predicted to be damaging (Fig. 1, Supporting Table 3; see online supporting information). There were 15 patients who carried 2 rare alleles in the same FA gene, which included 3 who were known to have anemia and/or leukopenia and 2 who presented with short stature. FA germline variants and genome instability may play a broader role in HNSCC susceptibility regardless of age or traditional risk factors.

Mutations in the FA core complex (FANCA, FANCB, FANCC, FANCE, FANCF, FANCG, FANCL, FANCM, and the recently identified FANCT) affect ubiquitination of the FANCD2/FANCI complex, a critical step in the FA pathway that repairs DNA damage and maintains genomic stability.[5,30] FANCL is a member of the FA core complex with E3 ligase enzymatic activity for FANCD2 monoubiquitination.[31] FANCE recruits FANCD2 to the FA core complex for ubiquitination and subsequent DNA repair.[32] It is interesting to note that FANCD2, FANCL, and FANCE homologs form a subset of the FA genes present in *Ciona intestinalis*, which is believed to be the closest invertebrate relative of vertebrates[33,34]; thus, these homologs appear to be an evolutionarily conserved part of the FA pathway. It is likely that mutations affecting *FANCD2, FANCE*, or *FANCL* modify cancer susceptibility through altered DNA repair and genomic instability. In the current study, we observed an increased burden, or mutation load, of variants in

FANCD2, FANCL, and *FANCE* in patients with HNSCC compared with population-level estimates (Table 8). Although it might be expected that patients with FA with mutations in *FANCD2, FANCL*, and *FANCE* may be predisposed to developing HNSCC, it is not practical to evaluate this from the clinical experience of patients with FA, because those with *FANCL, FANCE*, and *FANCD2* mutations represent only 0.4%, 1%, and 4% of all patients with FA,[4] respectively, and may often die from other causes, not surviving long enough to develop HNSCC. Recently, missense mutations in the FANCD2/FANCI complex have been associated with colorectal carcinoma.[35] The FA pathway, and FANCD2 in particular, is implicated in facilitating replication through common fragile sites,[36] which is a critical process in maintaining genomic stability. It is noteworthy that a study reporting results from sequencing 190 patients who had esophageal squamous cell carcinoma (ESCC) for germline variants in 12 FA genes identified 1 heterozygous indel variant each in *FANCD2, FANCE*, and *FANCL* in 3 patients, each with a strong family history of ESCC.[37] ESCC and squamous cell carcinoma of the anogenital tract, in addition to HNSCC, form the cancer spectrum displayed in patients with FA.[17] Germline mutations in *FANCD2, FANCE*, and *FANCL* may provide a novel area of study for HNSCC susceptibility and tumorigenesis.

In conclusion, our analysis of FA gene germline variations in 417 patients with HNSCC aged <50 years identified 15 patients who carried 2 variants in an FA gene, of whom 5 presented with an FA-associated phenotype, but we did not identify a specific age cutoff for FA screening. In our cohort, 185 of 417 patients (44%) were identified as heterozygous carriers of a rare FA gene variant, and variants in 60% of these patients (108 of 185) were predicted to be damaging. In addition, the increased mutation burden of *FANCE, FANCL*, and *FANCD2*, which are key players in DNA repair pathway activation, may indicate the importance of the FA pathway in HNSCC tumorigenesis.

FUNDING SUPPORT

This work was supported in part by the Fanconi Anemia Research Fund, The University of Texas MD Anderson Cancer Center Christopher and Susan Damico Chair in Viral Associated Malignancies, the National Institutes of Health through The MD Anderson Cancer Center Support Grant (CA016672), and the Intramural Research Program of the National Human Genome Research Institute, National Institutes of Health.

CONFLICT OF INTEREST DISCLOSURES

The authors made no disclosures.

AUTHOR CONTRIBUTIONS

Settara C. Chandrasekharappa: Conceptualization, methodology, investigation, resources, writing–original draft, writing–review and editing, visualization, supervision, project administration, and funding acquisition. **Steven B. Chinn**: Validation, formal analysis, investigation, data curation, writing–original draft, writing–review and editing, and visualization. **Frank X. Donovan**: Methodology, software, validation, formal analysis, investigation, resources, data curation, writing–original draft, writing–review and editing, and visualization. **Naweed I. Chowdhury**: Data curation and funding acquisition. **Aparna Kamat**: Formal analysis and investigation. **Adebowale A. Adeyemo**: Methodology and formal analysis. **James W. Thomas**: Methodology, software, validation, and formal analysis. **Meghana Vemulapalli**: Software and formal analysis. **Caroline S. Hussey**: Investigation and data curation. **Holly H. Reid**: Investigation and data curation. **James C. Mullikin**: Methodology and software. **Qingyi Wei**: Resources and funding acquisition. **Erich M. Sturgis**: Conceptualization, investigation, resources, data curation, writing–original draft, writing–review and editing, visualization, supervision, project administration, and funding acquisition.

REFERENCES

1. Rochowski A, Rosenberg PS, Alonzo TA, Gerbing RB, Lange BJ, Alter BP. Estimation of the prevalence of Fanconi anemia among patients with de novo acute myelogenous leukemia who have poor recovery from chemotherapy. *Leukemia Res*. 2012;36:29-31.
2. Rosenberg PS, Tamary H, Alter BP. How high are carrier frequencies of rare recessive syndromes? Contemporary estimates for Fanconi anemia in the United States and Israel. *Am J Med Genet A*. 2011; 155A:1877-1883.
3. Mamrak NE, Shimamura A, Howlett NG. Recent discoveries in the molecular pathogenesis of the inherited bone marrow failure syndrome Fanconi anemia [published online ahead of print October 13, 2016]. *Blood Rev*. doi: 10.1016/j.blre.2016.10.002.
4. Wang AT, Smogorzewska A. SnapShot: Fanconi anemia and associated proteins. *Cell*. 2015;160:354-354.e1.
5. Kottemann MC, Smogorzewska A. Fanconi anaemia and the repair of Watson and Crick DNA crosslinks. *Nature*. 2013;493:356-363.
6. Kutler DI, Singh B, Satagopan J, et al. A 20-year perspective on the International Fanconi Anemia Registry (IFAR). *Blood*. 2003;101: 1249-1256.
7. Schneider M, Chandler K, Tischkowitz M, Meyer S. Fanconi anaemia: genetics, #molecular |biology, and cancer—implications for clinical management in children and adults. *Clin Genet*. 2015;88:13-24.
8. Giampietro PF, Verlander PC, Davis JG, Auerbach AD. Diagnosis of Fanconi anemia in patients without congenital malformations: an international Fanconi Anemia Registry Study. *Am J Med Genet*. 1997;68:58-61.
9. Alter BP, Greene MH, Velazquez I, Rosenberg PS. Cancer in Fanconi anemia [letter]. *Blood*. 2003;101:2072.
10. Tan IB, Cutcutache I, Zang ZJ, et al. Fanconi's anemia in adulthood: chemoradiation-induced bone marrow failure and a novel FANCA mutation identified by targeted deep sequencing. *J Clin Oncol*. 2011;29:e591-e594.
11. Joenje H, Patel KJ. The emerging genetic and molecular basis of Fanconi anaemia. *Nat Rev Genet*. 2001;2:446-457.
12. Friedenson B. BRCA1 and BRCA2 pathways and the risk of cancers other than breast or ovarian [serial online]. *MedGenMed*. 2005;7:60.
13. Foulkes WD, Shuen AY. In brief: BRCA1 and BRCA2. *J Pathol*. 2013;230:347-349.

14. Tischkowitz M, Xia B. PALB2/FANCN: recombining cancer and Fanconi anemia. *Cancer Res*. 2010;70:7353-7359.
15. Sturgis EM, Cinciripini PM. Trends in head and neck cancer incidence in relation to smoking prevalence: an emerging epidemic of human papillomavirus-associated cancers? *Cancer*. 2007;110:1429-1435.
16. Neville BW, Day TA. Oral cancer and precancerous lesions. *CA Cancer J Clin*. 2002;52:195-215.
17. Rosenberg PS, Greene MH, Alter BP. Cancer incidence in persons with Fanconi anemia. *Blood*. 2003;101:822-826.
18. Rosenberg PS, Alter BP, Ebell W. Cancer risks in Fanconi anemia: findings from the German Fanconi Anemia Registry. *Haematologica*. 2008;93:511-517.
19. Velleuer E, Dietrich R. Fanconi anemia: young patients at high risk for squamous cell carcinoma [serial online]. *Mol Cell Pediatr*. 2014; 1:9.
20. Wang LE, Hu Z, Sturgis EM, et al. Reduced DNA repair capacity for removing tobacco carcinogen-induced DNA adducts contributes to risk of head and neck cancer but not tumor characteristics. *Clin Cancer Res*. 2010;16:764-774.
21. Chandrasekharappa SC, Lach FP, Kimble DC, et al. Massively parallel sequencing, aCGH, and RNA-Seq technologies provide a comprehensive molecular diagnosis of Fanconi anemia. *Blood*. 2013;121: e138-e148.
22. Cingolani P, Platts A, Wang LL, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w^{1118} ; iso-2; iso-3. *Fly (Austin)*. 2012;6:80-92.
23. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data [serial online]. *Nucleic Acids Res*. 2010;38:e164.
24. Buske OJ, Manickaraj A, Mital S, Ray PN, Brudno M. Identification of deleterious synonymous variants in human genomes. *Bioinformatics*. 2013;29:1843-1850.
25. Brunak S, Engelbrecht J, Knudsen S. Prediction of human mRNA donor and acceptor sites from the DNA sequence. *J Mol Biol*. 1991; 220:49-65.
26. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2015.
27. Wilm A, Aw PP, Bertrand D, et al. LoFreq: a sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. *Nucleic Acids Res*. 2012;40:11189-11201.
28. Alter BP, Joenje H, Oostra AB, Pals G. Fanconi anemia: adult head and neck cancer and hematopoietic mosaicism. *Arch Otolaryngol Head Neck Surg*. 2005;131:635-639.
29. Stransky N, Egloff AM, Tward AD, et al. The mutational landscape of head and neck squamous cell carcinoma. *Science*. 2011;333:1157-1160.
30. Rickman KA, Lach FP, Abhyankar A, et al. Deficiency of UBE2T, the E2 ubiquitin ligase necessary for FANCD2 and FANCI ubiquitination, causes FA-T subtype of Fanconi anemia. *Cell Rep*. 2015;12:35-41.
31. Hodson C, Walden H. Towards a molecular understanding of the Fanconi anemia core complex [serial online]. *Anemia* 2012:926787, 2012.
32. Polito D, Cukras S, Wang X, et al. The carboxyl terminus of FANCE recruits FANCD2 to the Fanconi anemia (FA) E3 ligase complex to promote the FA DNA repair pathway. *J Biol Chem*. 2014;289:7003-7710.
33. Stanley EC, Azzinaro PA, Vierra DA, Howlett NG, Irvine SQ. The simple chordate Ciona intestinalis has a reduced complement of genes associated with Fanconi anemia. *Evol Bioinform Online*. 2016; 12:133-148.
34. Delsuc F, Brinkmann H, Chourrout D, Philippe H. Tunicates and not cephalochordates are the closest living relatives of vertebrates. *Nature*. 2006;439:965-968.
35. Segui N, Mina LB, Lazaro C, et al. Germline mutations in FAN1 cause hereditary colorectal cancer by impairing DNA repair. *Gastroenterology*. 2015;149:563-566.
36. Madireddy A, Kosiyatrakul ST, Boisvert RA, et al. FANCD2 facilitates replication through common fragile sites. *Mol Cell*. 2016;64: 388-404.
37. Akbari MR, Malekzadeh R, Lepage P, Roquis D, Sadjadi AR, Aghcheli K, et al. Mutations in Fanconi anemia genes and the risk of esophageal cancer. *Hum Genet*. 2011;129:573-582.