CLINICAL GENETICS An International Journal of Genetics, Molecular and Personalized Medicine

Clin Genet 2016
Printed in Singapore. All rights reserved

© 2016 John Wiley & Sons A/S.
Published by John Wiley & Sons Ltd
CLINICAL GENETICS
doi: 10.1111/cge.12748

Original Article

Hereditary truncating mutations of DNA repair and other genes in *BRCA1/BRCA2/PALB2*-negatively tested breast cancer patients

F. Lhota[a,†], P. Zemankova[a,†], P. Kleiblova[a,b], J. Soukupova[a], M. Vocka[c], V. Stranecky[d], M. Janatova[a], H. Hartmannova[d], K. Hodanova[d], S. Kmoch[d] and Z. Kleibl[a]

[a]Institute of Biochemistry and Experimental Oncology, First Faculty of Medicine, Charles University in Prague, Prague, Czech Republic, [b]Institute of Biology and Medical Genetics, First Faculty of Medicine, Charles University in Prague and General University Hospital in Prague, Prague, Czech Republic, [c]Department of Oncology, First Faculty of Medicine, Charles University in Prague and General University Hospital in Prague, Prague, Czech Republic, and [d]Institute of Inherited Metabolic Disorders, First Faculty of Medicine, Charles University in Prague and General University Hospital in Prague, Prague, Czech Republic

[†]These authors contributed equally to this work.

Lhota F., Zemankova P., Kleiblova P., Soukupova J., Vocka M., Stranecky V., Janatova M., Hartmannova H., Hodanova K., Kmoch S., Kleibl Z. Hereditary truncating mutations of DNA repair and other genes in *BRCA1/BRCA2/PALB2*-negatively tested breast cancer patients.
Clin Genet 2016. © John Wiley & Sons A/S. Published by John Wiley & Sons Ltd, 2016

Hereditary breast cancer comprises a minor but clinically meaningful breast cancer (BC) subgroup. Mutations in the major BC-susceptibility genes are important prognostic and predictive markers; however, their carriers represent only 25% of high-risk BC patients. To further characterize variants influencing BC risk, we performed SOLiD sequencing of 581 genes in 325 BC patients (negatively tested in previous *BRCA1/BRCA2/PALB2* analyses). In 105 (32%) patients, we identified and confirmed 127 truncating variants (89 unique; nonsense, frameshift indels, and splice site), 19 patients harbored more than one truncation. Forty-six (36 unique) truncating variants in 25 DNA repair genes were found in 41 (12%) patients, including 16 variants in the Fanconi anemia (FA) genes. The most frequent variant in FA genes was c.1096_1099dupATTA in *FANCL* that also show a borderline association with increased BC risk in subsequent analysis of enlarged groups of BC patients and controls. Another 81 (53 unique) truncating variants were identified in 48 non-DNA repair genes in 74 patients (23%) including 16 patients carrying variants in genes coding proteins of estrogen metabolism/signaling. Our results highlight the importance of mutations in the FA genes' family, and indicate that estrogen metabolism genes may reveal a novel candidate genetic component for BC susceptibility.

Conflict of interest

All authors declare no conflict of interest.

Key words: DNA repair – FANCL – genetic predisposition testing – hereditary breast cancer – high-throughput nucleotide sequencing – sequencing panel – SOLiD next-generation sequencing

Corresponding author: Zdenek Kleibl, MD, PhD, Institute of Biochemistry and Experimental Oncology, First Faculty of Medicine, Charles University in Prague, U Nemocnice 5, 128 53 Prague 2, Czech Republic.
Tel.: +420 22 496 5745;
fax.: +420 22 496 5732;
e-mail: zdekleje@lf1.cuni.cz

Received 12 December 2015, revised and accepted for publication 22 January 2016

Breast cancer (BC; OMIM#114480) emerges as a leading cause of cancer death in female population worldwide. Hereditary breast cancer (HBC) accounts approximately for 5–10% of cases. Clinical importance of HBC results from the high lifetime risk of BC development, increased risk of other associated cancers, early disease onset, and 50% probability of the mutant allele's transmission to the offspring (1). Hence, the identification of germline mutations that confer BC susceptibility is an important task of clinical oncogenetics with considerable clinical utility, including tailored healthcare focused on early cancer identification, preventive surgical strategies decreasing cancer risk, and specific therapeutic strategies (2). The most frequently mutated genes in HBC patients are *BRCA1* and *BRCA2*; however, mutations in these genes account for less than 25% of cases in HBC patients. Since the identification of major BC-susceptibility genes, numerous other predisposition genes have been identified. Their characterization has been strongly accelerated with the availability of next-generation sequencing (NGS) technologies (3). Mutational analyses of recently identified BC-susceptibility genes indicate that frequencies of their mutations are substantially lower than that in *BRCA1* and *BRCA2*, besides being highly variable among populations worldwide (4). However, mutations in these newly established BC-susceptibility genes could collectively epitomize another 25% of ascertained genetic risk in HBC patients and thus their analyses are gradually introduced into the clinical analyses (5).

A striking characteristic of the majority of known BC-susceptibility genes is the contribution of their protein products in DNA damage repair (6). On the other hand, the existence of known BC-susceptibility genes that code for proteins not directly involved in these processes (e.g. *PTEN*, *CDH1*, or *NF1*) indicates that non-DNA repair genes could also contribute to BC susceptibility (7).

Our previous gene-by-gene mutational analyses revealed that the most frequent mutations in Czech HBC patients are found in the *BRCA1* gene (8–10). Less frequently, we identified pathogenic variants in *BRCA2* or *PALB2* (11). In this study, we aimed to describe the presence of potentially pathogenic hereditary variants in other known BC-susceptibility genes using the targeted NGS and to identify further variants that may contribute to BC susceptibility in high-risk Czech BC patients.

Materials and methods

Detailed methods are available in Supporting information methods.

Patients and samples

The 325 successfully sequenced patients' samples were selected from a sample collection of high-risk Czech BC patients that fulfilled testing criteria described previously (8, 9, 11), were negatively tested for the presence of mutations in *BRCA1*/*BRCA2*/*PALB2*, and gave their informed consent approved by local ethical committee. As controls, we analyzed 105 samples obtained from Czech non-cancer elder females selected according to their age (>50 years; median age 71 years; ranged 54–95 years) from non-cancer controls described previously (12). The genotyping of the c.1096_1099dupATTA variant in *FANCL* was performed on additional sample sets of 337 high-risk BC patients, 673 sporadic BC patients and 686 non-cancer controls (13, 14) using high-resolution melting analysis and confirmed by Sanger sequencing (Fig. S1, Supporting information). Clinical and histopathological characteristics of analyzed high-risk BC patients are available in Tables S1 and S2.

Sequencing gene panel

The targeted genes comprised two groups consisting of 141 DNA repair genes and 449 genes retrieved from Phenopedia database (15) using the disease term 'breast neoplasms' with at least two entries (assessed February 2012). Finally, 581 targeted genes (listed in Table S3) were sequenced successfully.

Library construction, sequence capture and sequencing

Fragmented DNA was subjected to ligation of SOLiD sequencing adaptors and polymerase chain reaction (PCR)-based incorporation of bar codes, as described previously (16). The target DNA enrichment was performed by a custom SeqCap EZ Choice Library (Roche), and SOLiD sequencing primers were introduced by PCR. The final libraries were amplified by an emulsion PCR and sequenced on a SOLiD 4 System (Thermo Fisher, Waltham, MA, USA).

Bioinformatics pipeline, variant filtration, and prioritization of missense variants

Sequencing reads were aligned to the human genome reference (GRCh37/hg19) using Novoalign (CS1.01.08). Picard was used for duplicate removal and SAMtools (0.1.8) for SAM-to-BAM conversion and calling of single nucleotide variants (SNVs) and small insertions and deletions (indels). Variant annotation was performed with ANNOVAR (17).

Variant filtration excluded off-target sequences and low confidence variants (sequence quality <150; sequencing coverage <10). We also excluded common variants with allelic frequencies in ESP6500 and 1000 Genomes databases >0.01. To reflect the population-specific variants and variants influencing cancer susceptibility, we excluded variants presented in no patient or in more than two controls.

To identify missense variants with a putative contribution to BC susceptibility, we performed a prioritization that considered five prediction algorithms (SIFT, PolyPhen-2, LRT, MutationTaster, and PhyloP) and two databases (ClinVar and HGMD) aggregating data about genotypes and corresponding clinical characteristics. Prioritized variants were considered those that were

called by each prediction software as deleterious (or unknown) or considered as disease-associated in ClinVar and HGMD databases.

Confirmation of truncating variants

All truncating variants were confirmed by conventional Sanger sequencing. The variants affecting a conservative splice site were analyzed from the blood-isolated patient's RNA, when available, using RT-PCR and sequencing as described previously (18). All primers are listed in Table S4.

Statistical analysis

The differences among analyzed groups and subgroups were calculated by the chi-square test or Fischer exact test if the expected number of events was lower than six.

Results

In the set of 325 patients' samples and 105 controls, we obtained 491,385 variants in exome and adjacent intronic sequences of 581 targeted genes. The mean sequencing coverage was 56.5 and 93% of the captured sequence was covered by >10 reads. Using the variant filtration, we identified 4540 rare variants in the final dataset representing 2647 unique variants of 496/581 targeted genes (85.4%). We found 144 truncating variants (either nonsense, frameshift indels, or splice site alterations), representing 89 unique variants, in 73/581 targeted genes (12.6%).

The set of 325 BC patients harbored 4053 rare variants (2647 unique) including 127 truncating variants (89 unique), 34 in-frame indels (22 unique), 2347 missense SNVs (1599 unique), and 1545 synonymous SNVs (937 unique). We primarily focused on the truncating variants that were identified in 105/325 (32.3%) BC patients (Fig. 1) and were all confirmed by Sanger sequencing. Nineteen patients carried more than one truncating variant (1 patient carried four, 1 patient carried three, and 17 patients carried two truncations), 86 patients carried one truncating variant. The group of truncating variants included 20 splicing variants (14 unique, each affecting one particular gene) flanking to intronic (±2 bp) sequences. Their impact on splicing was studied at the mRNA level (available from eight patients). Seven out of eight analyzed splicing variants showed frameshift (Figs S2 and S3). The prioritization analysis revealed 356 potentially pathogenic variants out of 1599 rare unique missense variants (22%).

Hereditary variants in DNA repair genes

In 25 DNA repair genes, we identified 46/127 truncations (36/89 unique) in 41 (12.6%) BC patients (Table 1). The most frequent alterations affected genes that code for DNA double-strand break (DDSB)/interstrand crosslink (ICL) repair proteins. These included 16 patients carrying nine unique truncating variants in five Fanconi anemia (FA) genes (Fig. 1).

Another 19 patients harbored 19 unique variants affecting other genes coding for proteins involved in the DDSB repair pathways, including homologous recombination (HR; *ATM*, *EXO1*, *WRN*, *BLM*, *DCLRE1C*, *FAM175B/ABRO1*, *HELQ, NBN*, *RAD18*, *RAD50*, *RAD51D*, *CHEK2*, and *RFC1*) but also non-homologous end joining (NHEJ; *XRCC4*) repair. Finally, eight truncating variants, each in one patient, were identified in the genes that code for proteins involved in other DNA repair processes including single-strand DNA repair (*ATR, ATRIP*), nucleotide excision repair (NER; *ERCC2, ERCC6*), mismatch repair (MMR; *MSH5*), and direct removal of alkylated guanine (*MGMT*). Two patients carried truncating variants in more than one gene involved in different DNA repair pathways.

Altogether, 106 unique prioritized missense variants in 59 DNA repair genes were identified in 133 patients (34 of these variants, in 56 patients, were found in 15 genes in which at least 1 truncating variant was also detected; Table S5). The most frequent potentially pathogenic missense variants were found in *ATM* (12 variants in 17 patients) and *CHEK2* (4 variants in 13 patients). Among prioritized variants, we also identified pathogenic missense variants in *BRCA1* (c.115T>C; p.C39R), *TP53* (c.733G>A; p.G245S), and *CDH1* (c.1018A>G; p.T340A) in three young BC patients.

Extended analysis of *FANCL* c.1096_1099dupATTA

The most frequent frameshift variant found in six BC patients and none NGS control was c.1096_1099dupATTA (p.T367Nfs*13) in *FANCL* (previously described in an FA patient belonging to the FA-L complementation group; OMIM#614083) (19). Because of the insufficient number of NGS controls, we first compared the frequency of this *FANCL* variant among our patients with data from the Exome Aggregation Consortium (ExAC) database (http://exac.broadinstitute.org; accessed May 2015) indicating an overrepresentation of this variant among our high-risk BC patients (Table 1). Therefore, we further analyzed another 337 high-risk BC Czech patients (329 females, 8 males; all *BRCA1*/*BRCA2*/*PALB2*-negative). Among these, we identified another three c.1096_1099dupATTA carriers with BC (all diagnosed before age of 38 years). Overall, the c.1096_1099dupATTA was identified in 9/662 high-risk BC individuals (1.3%).

To identify the carriers of c.1096_1099dupATTA in sporadic BC patients and other controls, we genotyped 673 unselected BC cases and 686 non-cancer controls (313 females and 373 males). This analysis revealed three carriers in each analyzed group, showing its frequency as 0.4% in both BC cases (3/693) and controls (3/791; including 105 NGS controls and 686 genotyped controls), respectively. Thus, the frequency of c.1096_1099dupATTA was significantly (Fisher exact test) overrepresented only among high-risk individuals ($p = 0.04$) but not in sporadic BC patients ($p = 0.9$). All 14 carriers among patients were females, while all three carriers in controls were males.

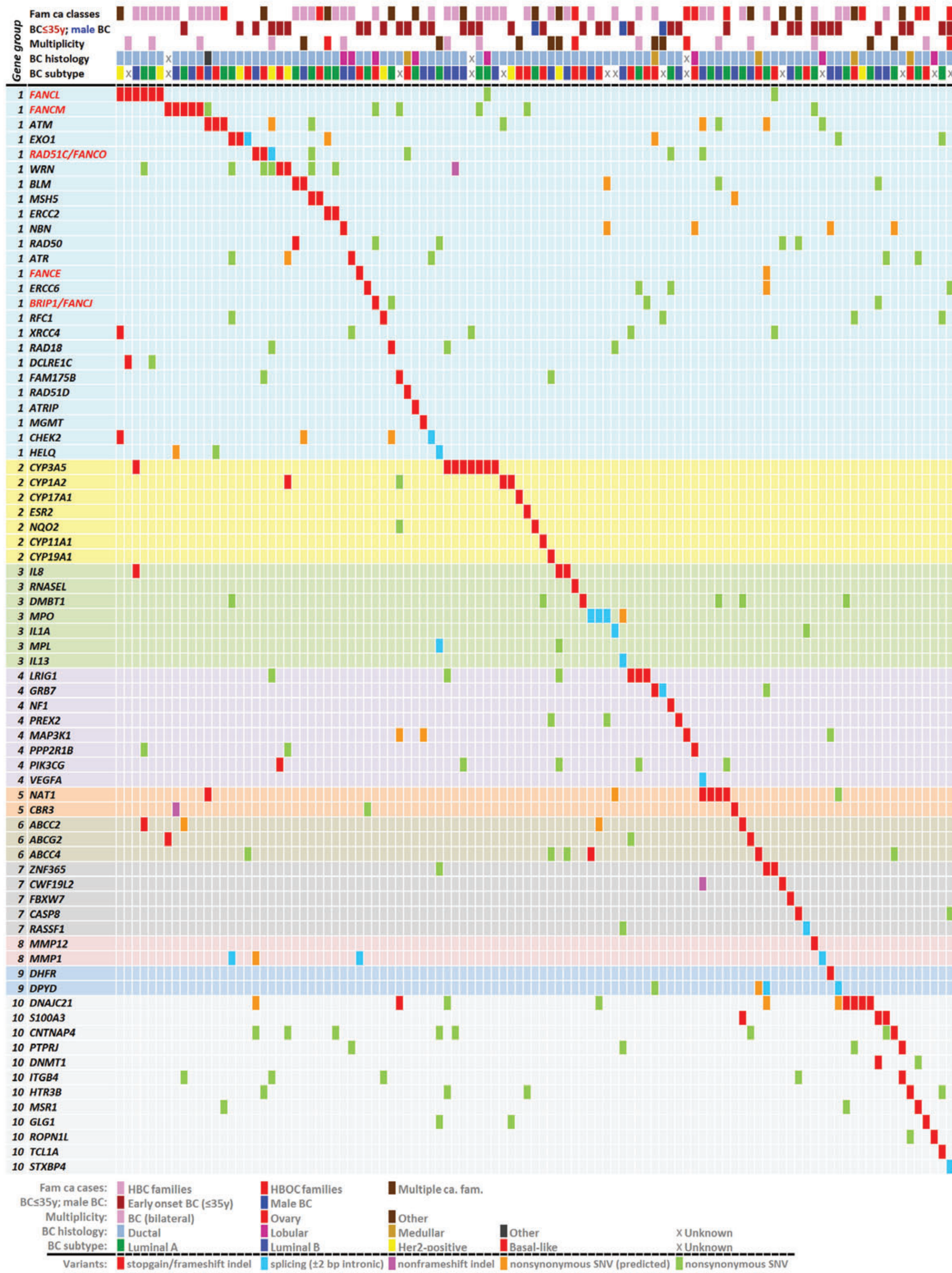

Fig. 1. Overview of variants in 73 genes (rows) affected by at least one truncating (nonsense, frameshift, or splicing) variant, that were identified in 105 BC patients (columns). Pathological characteristics of BC tumors (histology and subtypes) and selected clinical characteristics (BC in females at the age of <35 years or male BC, and the presence of familial cancer) are shown in five upper lines (color markings are displayed in Fig. 2; X denotes a missing information). The patients and genes are ordered according to the overall number of found variants, genes (with the Fanconia anemia gene members highlighted in red letters) are grouped by functional relationship of coded proteins (see note). Note: Genes in gene groups (1–10; number (N) patients with at least one truncating variant) were ascertained as follows: the genes coding proteins involved in DNA repair (1; $N = 41$); steroid hormones synthesis, turnover or signaling (2; $N = 16$); immune response (3; $N = 11$); membrane receptor signaling (4; $N = 11$); metabolism of xenobiotics (5; $N = 6$); membrane transport of molecules (6; $N = 6$); cell cycle/apoptosis regulation (7; $N = 6$); cell-to-cell communication (8; $N = 4$); nucleotide metabolism (9; $N = 3$); or other (unsorted) processes (10; $N = 16$). Color markings used for pathological and clinical characteristics (shown in legend) are identical to that presented in Fig. 2. Clinical and histopathological characteristics of truncating mutation carriers are shown in Table S6.

Table 1. List of 36 unique truncating variants (nonsense, frameshift indels, or splicing) that were found in 25 genes coding for proteins involved in DNA repair and DNA damage response identified in 41/325 BC patients (Pts) and 7/105 non-cancer controls (Ctrls)[a]

Gene	HGVS coding	HGVS protein[b]	Classification	Rs number	HGMD/ ClinVar	Pts (*N*)	Ctrls (*N*)	ExAC (mut/all)[b, c]
FANCL	c.1096_1099dupATTA	p.T367Nfs*13	Indel			6	0	232/65648*
FANCM	c.1972C>T	p.R658*	Nonsense		DM	1	1	7/66502*
	c.3979_3980delCA	p.Q1327Vfs*16	Indel			1	0	0/66498*
	c.5101C>T	p.Q1701*	Nonsense	rs147021911	DM?	2	0	95/66562
	c.5791C>T	p.R1931*	Nonsense	rs144567652	DM	1	0	63/66622
ATM	c.3850delA	p.T1284Qfs*9	Indel		DM	1	0	n.r.
	c.7327C>T	p.R2443*	Nonsense	rs121434220	DM/P	2	0	n.r.
EXO1	c.1522dupT	p.C508Lfs*7	Indel			1	0	n.r.
	c.2358delG	p.L787Yfs*37	Indel			1	0	n.r.
	c.2212-1G>C	p.V738_K743del	Splicing	rs4150000	DM	1	1	172/63478
CHEK2	c.277delT	p.W93Gfs*17	Indel			1	0	n.r.
	c.444+1G>A	p.R148Vfs*6	Splicing		DM	2	0	11/66720*
RAD51C	c.502A>T	p.R168*	Nonsense			2	0	n.r.
	c.905-2_1delAG	p.L301Gfs*42	Splicing		DM	1	0	n.r.
BLM	c.1642C>T	p.Q548*	Nonsense	rs200389141	DM	2	0	21/66322*
ERCC2	c.230_231delTG	p.V77Afs*4	Indel			1	0	n.r.
	c.1703_1704delTT	p.F568Yfs*2	Indel		DM	1	2	11/65444*
MSH5	c.541C>T	p.R181*	Nonsense	rs147515280		1	0	13/65882*
	c.1900C>T	p.R634*	Nonsense			1	0	n.r.
WRN	c.604A>T	p.K202*	Nonsense			1	0	n.r.
	c.4216C>T	p.R1406*	Nonsense	rs11574410		1	0	87/65788
ATR	c.5342T>A	p.L1781*	Nonsense			1	0	n.r.
ATRIP	c.827_828delAG	p.E276Gfs*2	Indel			1	0	n.r.
BRIP1	c.2392C>T	p.R798*	Nonsense	rs137852986	DM/P	1	0	16/65688*
DCLRE1C	c.1903dupA	p.S635Kfs*6	Indel			1	0	27/66734*
ERCC6	c.3693C>G	p.Y1231*	Nonsense			1	0	n.r.
FAM175B	c.1084delC	p.Q362Kfs*19	Indel			1	0	n.r.
FANCE	c.929dupC	p.V311Sfs*2	Indel		DM	1	1	n.r.
HELQ	c.2677-1G>A	p.Q348Pfs*17	Splicing	rs200992133		1	1	27/66528
MGMT	c.207_210dupACGT	p.S70Yfs*5	Indel			1	0	n.r.
NBN	c.657_661delACAAA	p.K219Nfs*16	Indel			1	1	21/65324
RAD18	c.1430_1431insGCGG	p.T478Rfs*6	Indel			1	0	n.r.
RAD50	c.1093C>T	p.R365*	Nonsense			1	0	n.r.
RAD51D	c.355_358deldelTGTA	p.C119Wfs*16	Indel			1	0	n.r.
RFC1	c.2191delA	p.R731Gfs*7	Indel			1	0	n.r.
XRCC4	c.25delC	p.H9Tfs*8	Indel			1	0	42/66632
Total variants						46	7	

Variants listed in HGMD or ClinVar databases: DM, disease-causing (pathological) mutations; DM?, likely disease-causing (likely pathological) mutation; P, pathogenic. ExAC, Exome Aggregation Consortium; n.r., variant not reported in ExAC.
[a]The enhanced version of the table (including missense variant predicted as pathogenic, numbers of reference transcripts, and frequencies in ExAC, ESP6500, and 1000 genomes databases) is available as Table S5.
[b]ExAC allelic frequency in European non-Finnish population (mutated alleles/wt alleles).
[c]Asterisk (*) indicates significant differences ($p < 0.05$) between allelic frequencies in European non-Finnish population (ExAC) and in analyzed population of patients (Fisher exact test).

Hereditary variants in non-DNA repair genes

The remaining 81/127 truncations (53/89 unique) in 48 non-DNA repair genes were identified in 74 (22.8%) BC patients (Table 2). We found variants in only non-DNA repair genes in 64 of them, while in 10 patients we also detected some truncating variants in DNA repair genes. To identify possible defects in pathways that may contribute to BC susceptibility, we sorted the affected genes into nine groups (Group 2–9 in Table 2 and Fig. 1) clustering functionally related proteins. Twelve genes (Group 10) comprised proteins with unrelated or unclear function. Sixteen carriers (5% of all patients) of eight different truncating variants have been identified in the 'Group 2' associating genes that code for proteins involved in steroid hormones metabolism or signaling.

Further, we detected 250 unique, prioritized, potentially pathogenic missense variants in 150 genes in 213 patients. The most frequent prioritized SNVs in non-DNA repair genes affected the *APC* gene (in eight patients).

Individual and disease characteristics in carriers of truncating variants

We found no significant differences in the characteristics of patients and tumors between the carriers of truncating

Table 2. List of 53 unique truncating variants (nonsense, frameshift indels, or splicing) in 48 non-DNA repair genes identified in 74/325 BC patients and 10/105 non-cancer controls[a]

Gene	Gr	HGVS coding	HGVS protein[b]	Classification	Rs number	HGMD/ClinVar	Pts (*N*)	Ctrs (*N*)	ExAC (mut/all)[b, c]
CYP3A5	2	c.92dupG	p.L32Tfs*3	Indel			7	1	732/66688
	2	c.246dupG	p.A83Gfs*40	Indel			1	0	21/66728
CYP1A2	2	c.816T>A	p.Y272*	Nonsense	rs140421378	FTV	3	0	16/65958*
CYP11A1	2	c.835delA	p.I279Yfs*10	Indel		DM	1	0	2/66694*
CYP17A1	2	c.1072C>T	p.R358*	Nonsense		DM	1	0	n.r.
CYP19A1	2	c.1058dupT	p.L353Ffs*10	Indel			1	0	1/60606*
ESR2	2	c.76G>T	p.E26*	Nonsense			1	0	1/66734*
NQO2	2	c.628C>T	p.Q210*	Nonsense			1	0	1/66386*
IL8	3	c.91G>T	p.E31*	Nonsense	rs188378669		3	2	104/66426
DMBT1	3	c.2227delC	p.Q743Rfs*4	Indel			1	0	n.r.
IL13	3	c.174+2delT	p.(?)	Splicing			1	0	n.r.
IL1A	3	c.319+2T>C	p.(?)	Splicing			1	0	n.r.
MPL	3	c.79+2T>A	p.(?)	Splicing	rs146249964	DM	1	0	114/66230
MPO	3	c.2031-2A>C	p.R677Wfs*73	Splicing	rs35897051	DM	3	1	470/66434
RNASEL	3	c.793G>T	p.E265*	Nonsense	rs74315364	DM/P	1	1	381/66212
LRIG1	4	c.3149_3150delCG	p.A1050Gfs*17	Indel			3	0	102/66704*
GRB7	4	c.862C>T	p.Q288*	Nonsense			1	0	2/48666*
	4	c.801+1G>C	p.(?)	Splicing			1	0	n.r.
MAP3K1	4	c.4151dupT	p.L1384Lfs*36	Indel			1	0	n.r.
NF1	4	c.5690delG	p.G1897Vfs*28	Indel			1	0	n.r.
PIK3CG	4	c.41_42delAG	p.E14Gfs*147	Indel			1	0	5/62474*
PPP2R1B	4	c.342_343delTG	p.V115Cfs*3	Indel			1	0	81/66082
PREX2	4	c.3210_3213delAACA	p.D1072Vfs*17	Indel			1	0	n.r.
VEGFA	4	c.1085+2T>C	p.(?)	Splicing	rs149528656		1	0	15/66648*
NAT1	5	c.559C>T	p.R187*	Nonsense	rs5030839	FP	5	1	252/66632
CBR3	5	c.533delA	p.D178Afs*46	Indel			1	0	102/66716
ABCC2	6	c.3196C>T	p.R1066*	Nonsense	rs72558199	DM/P	2	0	35/66738*
ABCC4	6	c.2468dupA	p.N823Kfs*12	Indel			1	1	26/66634
	6	c.1150C>T	p.R384*	Nonsense			1	0	n.r.
ABCG2	6	c.706C>T	p.R236*	Nonsense	rs140207606	FP	1	0	24/66634
	6	c.736C>T	p.R246*	Nonsense	rs200190472	FP/P	1	0	5/66692*
ZNF365	7	c.1065G>A	p.W355*	Nonsense	rs142406094		2	0	4/66740*
CASP8	7	c.106delG	p.E36Nfs*7	Indel			1	0	n.r.
CWF19L2	7	c.1605delA	p.K535Nfs*4	Indel			1	0	n.r.
FBXW7	7	c.310delC	p.H104Mfs*389	Indel			1	0	n.r.
RASSF1	7	c.888+1G>A	p.V258Gfs*7	Splicing			1	0	1/64856*
MMP1	8	c.105+2T>C	pQ35Vfs*11	Splicing	rs139018071	FTV	3	0	114/66230
MMP12	8	c.327C>T	p.W109*	Nonsense			1	0	0/65722*
DPYD	9	c.1905+1G>A	p.D581_N635del	Splicing	rs3918290	DM	2	1	389/66688
DHFR	9	c.95delT	p.F32Sfs*7	Indel			1	0	n.r.
DNAJC21	10	c.1503delA	p.K501Nfs*10	Indel			3	0	33/66560*
	10	c.1629delT	p.F543Lfs*4	Indel			2	0	5/11578*
S100A3	10	c.208delG	p.V70Wfs*83	Indel			3	1	291/66718
CNTNAP4	10	c.3913G>T	p.E1305*	Nonsense			1	0	3/56224*
DNMT1	10	c.1035dupC	p.K346Qfs*35	Indel			1	0	n.r.
GLG1	10	c.3520C>T	p.R1174*	Nonsense			1	0	n.r.
HTR3B	10	c.871C>T	p.Q291*	Nonsense			1	0	n.r.
ITGB4	10	c.665delG	p.G222Efs*60	Indel			1	0	n.r.
MSR1	10	c.569delT	p.L190Cfs*5	Indel			1	0	n.r.
PTPRJ	10	c.1191T>A	p.Y397*	Nonsense			1	0	n.r.
ROPN1L	10	c.135T>A	p.Y45*	Nonsense	rs41280363		1	1	126/66680
STXBP4	10	c.181-1G>A	p.K60Vfs*28	Splicing			1	0	3/66388*
TCL1A	10	c.253C>T	p.R85*	Nonsense			1	0	n.r.
Total variants							81	10	

Variants listed in HGMD or ClinVar databases: DM, disease-causing (pathological) mutations; DM?, likely disease-causing (likely pathological) mutation; P, pathogenic. ExAC, Exome Aggregation Consortium; n.r., variant not reported in ExAC.

[a]Genes are grouped (Gr 2–10) according to the functional relationship of coded proteins, as described in the legend of Fig. 1. The enhanced version of the table (including missense variant predicted as pathogenic, numbers of reference transcripts, and frequencies in ExAC, ESP6500, and 1000 genomes databases) is available as Table S5.

[b]ExAC allelic frequency in European non-Finnish population (mutated alleles/wt alleles).

[c]Asterisk (*) indicates Significant differences ($p < 0.05$) between allelic frequencies in European non-Finnish population (ExAC) and in analyzed population of patients (Fisher exact test).

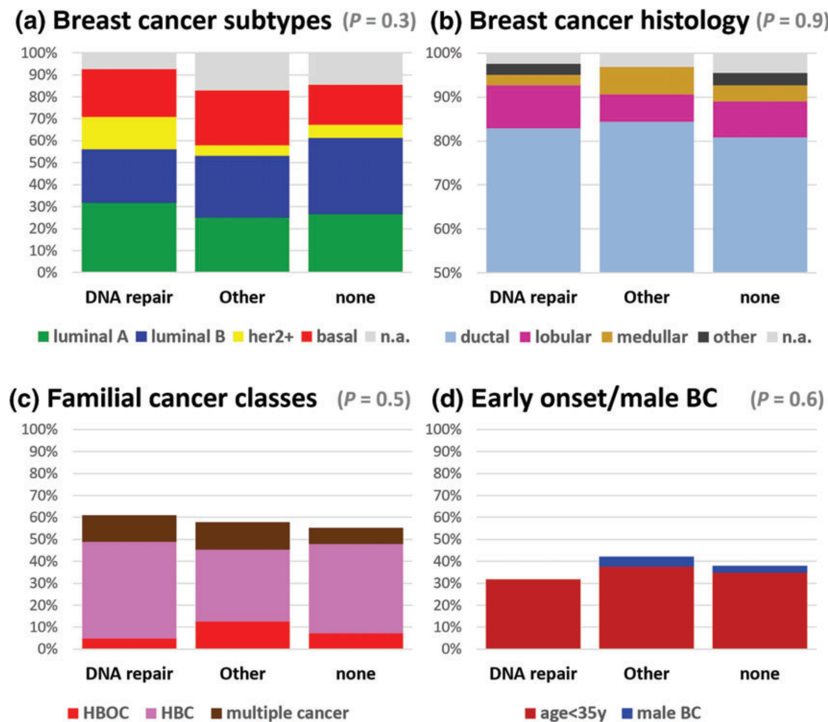

Fig. 2. Pathological characteristics of tumors and clinical characteristics of 325 analyzed BC patients grouped according to the presence of truncating variant in any DNA repair gene (DNA repair; 41 patients), variant in only other genes (other; 64 patients), and no truncating variant (none; 220 patients). The p-values (chi-square test) indicate insignificant differences in displayed characteristics among the analyzed subgroups.

variants in the DNA repair genes, carriers of truncating variants in other genes, and patients not carrying truncating variants (Fig. 2).

Discussion

Panel NGS represents a reliable approach for the analysis of cancer susceptibility in clinical settings but also in identification of candidate genes in high-risk individuals. In contrast to exome or even genome NGS, it allows the identification of the carriers of pathogenic variants in a cost-effective manner, with flexibility in the selection of gene targets, sensitivity, and manageable bioinformatics load for routine practice (20). Our analysis revealed the presence of truncating variants in nearly one third of analyzed patients and 30 patients (9%) carried truncating variants in some of 15 genes (*ATM*, *ATR*, *BLM*, *BRIP1*, *ERCC2*, *FANCE*, *FANCL*, *FANCM*, *CHEK2*, *NBN*, *NF1*, *RAD50*, *RAD51C*, *RAD51D*, *WRN*) analyzed by currently clinically used NGS panels (5). Out of 73 genes with truncating variants, in 51 genes we found only a single truncation. This indicates that rare variants could be identified in a substantial proportion of high-risk individuals; however, their clinical interpretation and differentiation from incidental findings not associating with BC susceptibility would be difficult.

Characterization of variants in DNA repair genes

The interesting result of our study is the high frequency of potentially pathogenic variants in five FA genes in 4.9% high-risk patients. FA genes code for DNA repair proteins contributing to genome stability maintenance by the ICL repair [reviewed in (21, 22)]. FA proteins form several protein–protein complexes (22). Hereditary bi-allelic mutations of FA genes are responsible for the development of FA characterized by congenital abnormalities, bone marrow failure, cellular hypersensitivity to DNA crosslinking agents and cancer susceptibility. The most frequent truncating variant was c.1096_1099dupATTA in *FANCL* that codes an ubiquitin ligase catalyzing the monoubiquitination of FANCI/FANCD2 (ID2) complex – the key step in FA pathway activation (23, 24). The c.1096_1099dupATTA variant was described by Ali et al. (19) in a patient that belonged to the FA-L complementation group. The mutated FANCL protein (p.T367Nfs*13) contains an aberrant chain of 12 amino acid residues that flanks to the PHD/RING finger domain catalyzing ubiquitin ligase activity. Ali et al. (19) performed its functional characterization revealing that the c.1096_1099dupATTA is a hypomorphic mutation resulting in the formation of altered protein with reduced binding to FA core complex and reduced FANCD2 monoubiquitination. Same variant was also identified by Akbari et al. (25) in a patient with familial esophageal squamous cell carcinoma. The results of our study, showing the overrepresentation of c.1096_1099dupATTA among high-risk BC patients, indicate that this variant may represent a novel BC-susceptibility allele. However, further studies, including segregation analyses providing information

about the association of c.1096_1099dupATTA with cancer phenotype in affected families and analyses of the variant in other populations, will be necessary to evaluate its potential clinical utility. As four out of six c.1096_1099dupATTA carriers identified by our panel NGS carried also other truncating variants (Fig. 1), we could not rule out the possibility that c.1096_1099dupATTA could act as a rather modifying variant. The recurrent mutations affecting *FANCL* and *FANCM* at their C-termini indicate that truncating variants of FA genes located in far C-terminal regions may impair the FA pathway under specific and so far uncomprehended circumstances. Such phenomenon has been proposed also for the nonsense c.9976A>T (p.K3326*) variant in *BRCA2/FANCD1* truncating the last 93 amino acids. In contrast to the majority of *BRCA2* pathogenic variants, the p.K3326* has been recognized as only a modest BC-susceptibility allele (OR = 1.26) increasing a risk of other cancers (26).

In *FANCM*, coding a helicase contributing to the formation of the FA anchor complex, we identified four truncating variants in five patients. Truncations in *FANCM* were recently associated with susceptibility for triple-negative BC (27). In three patients (none of them triple-negative), we identified previously characterized nonsense or exon skipping mutations that were shown to increase BC risk (27, 28). The remaining two *FANCM* variants included the rare nonsense mutation c.1972C>T (p.R658*; in a luminal BC patient whose mother and her sister suffered from bilateral BC) and the novel mutation c.3979_3980delCA (p.Q1327Vfs*16; in a BC patient with multiple breast and colorectal cancer (CRC) cases in the family). The association between CRC and germinal *FANCM* mutation has recently been identified in CRC tumor samples obtained from two c.5791C>T (p.R1931*) carriers (29).

We have also identified three *RAD51C/FANCO* mutation carriers (0.9% of patients). The *RAD51C* was originally identified as OC-susceptibility gene (30); however, later data conferred also increased BC susceptibility (31). Recently, we described two other pathogenic *RAD51C* variants in two OC patients (13). These data indicate that mutations in *RAD51C* may affect ~1% of Czech high-risk BC or OC patients.

Finally, the carriers of variants in FA genes comprised also two basal-like BC patients carrying pathogenic variants in *FANCE* and *BRIP1/FANCJ*, respectively. Both variants were reported in association with esophageal cancer (25) and triple-negative BC (31), respectively.

We found rare truncating variants in several other genes associated with hereditary BC that code for DDSB repair proteins; however, we also identified several truncating variants in the genes coding proteins engaged in other DNA repair pathways. Among others, an interesting candidate is *EXO1*, which codes for exonuclease involved in numerous DNA repair pathways. Besides two indels, we identified and characterized the c.2212-1G>C splicing mutation resulting in six amino acids in-frame deletion (p.V738_K743del), involving the interaction of EXO1 with MSH2 during MMR (32). Contrary to our analysis, Wu et al. (33) characterized the identical variant as a frameshift in a patient with hereditary non-polyposis CRC. Moreover, we further identified also two rare *EXO1* prioritized missense variants [c.325G>A (p.E109K) and c.1105A>C (p.S369R)] in five patients. Clustering of mutations in *EXO1* and presence of mutations in other genes involved for example in NER (*ERCC2*, *ERCC6*) suggest that an impairment of these repair processes by hereditary alterations could increase BC susceptibility. The degree to which these variants may influence BC susceptibility remains to be investigated by further studies.

Variants in non-DNA repair genes

The potentially deleterious hereditary variants were identified in 48/448 non-DNA repair genes, most frequently (in 16 BC patients) in the genes that code for the enzymes of steroid hormone metabolism and signaling. The group primarily included members of the cytochrome p450 superfamily contributing to the estrogen biosynthesis (CYP11A1, CYP17A1, CYP19A1) or catabolism (CYP3A5, CYP1A2) [reviewed in (34)]. Given that estrogens may affect BC etiology, variants in *CYP* genes may influence BC risk.

Variants in *CYP11A1* and *CYP17A1* identified in basal-like patients were previously described in patients suffering from severe congenital adrenal insufficiency (35) (OMIM#613743) and congenital adrenal hyperplasia (36) (OMIM#202110), respectively. Interestingly, Hopper et al. (37) reported p.R239* (c.775C>T) variant in *CYP17A1* in three *BRCA1/2*-negative young sisters with BC and hypothesized that this variant is responsible for dominantly inherited and possibly high-risk BC. Recently, Yang et al. (38) identified c.987delC (p.Y329*) variant in *CYP17A1* in a patient from an HBOC family. We found a novel variant, c.1058dupT (p.L353Ffs*10) in *CYP19A1*, in a patient with a BC and non-Hodgkin lymphoma duplicity whose mother developed bilateral BC. Mutations in similar positions cause aromatase deficiency (OMIM#613546).

Defects in estrogen-catabolizing enzymes suggest a mechanistically more obvious pathophysiological link to BC promotion. As estrogens are known substrates of CYP3A5 and CYP1A2 (34), 11 identified carriers of truncating variants in these genes could potentially have reduced estrogen clearance. We also found a nonsense variant in *NQO2* coding a quinone reductase eliminating estrogen quinones responsible for estrogen-initiated carcinogenesis (39) and one truncating variant in *ESR2* that codes for ERβ with anti-proliferative signaling (40). The high proportion of patients carrying constitutive truncating variants in steroid hormone metabolism genes supports the hypothesis of Hopper et al. (37) suggesting that cancer-causing mutations in these genes may represent a new pathophysiological mechanism linking genetic and environmental interactions in BC susceptibility.

The other functional groups associating the patients with truncating variants in functionally relevant genes were smaller. It is obvious that at least some variants in non-DNA repair genes have very limited (if any) impact on BC susceptibility and they rather represent incidental

findings [e.g. mutations in *ABCC2* was identified in a patient with the Dubin–Johnson syndrome (41) or known *DPYD* mutation related to the fluoropyrimidines toxicity (42)]. Reporting of incidental findings is highly questionable and a matter of debate (43, 44).

Disease and individual characteristics in carriers of truncating variants

Considering the patients and disease characteristics in the carriers of mutations in the major BC predisposing genes, the earlier age at BC diagnosis or more aggressive form of BC subtypes would be expected also in the carriers of mutations in other BC-susceptibility genes. In fact, we did not identify significant differences in clinical and histopathological characteristics between the carriers and non-carriers of truncating variants. This result did not change even when prioritized variants were added into the comparison (data not shown). Similar behavior was also documented recently in a large study of 1824 triple-negative BC patients analyzed by Couch et al. (31) for hereditary mutations in 17 genes, where significant differences in enrichment for family BC/OC history and tumor characteristics were identified only in the carriers of *BRCA1/2* mutations but not in carriers of non-*BRCA1/2* mutations. We suggest that some principal changes in the evaluation of clinical and histopathological characteristics will be required to assess the clinical importance of non-*BRCA1/2* BC-susceptibility genes. Since the frequencies of mutations in these genes are lower by order than that in *BRCA1/2*, an international and consortia effort will be required for such analyses.

Conclusions

Our study identified truncating variants in 32% of patients, and 9% of patients were carriers of a truncating variant in the genes currently analyzed in clinical NGS panels for the cancer risk prediction. The most frequent truncating variants affected FA genes that, together with *BRCA1*, *BRCA2*, and *PALB2*, make this group the most important for cancer susceptibility in BC patients. Our results also show an overrepresentation of the *FANCL* variant c.1096_1099dupATTA in high-risk patients, indicating that this variant may represent a novel BC-susceptibility allele. Moreover, we identified potentially pathogenic variants in several rarely mutated DNA repair genes indicating that despite its low frequency, variants in these genes may influence the development of HBC in Czech patients. We believe that it is important to analyze such genes and in international co-operation to evaluate their contribution to the BC development because they may represent clinically valuable predictors of cancer risk in families of mutation carriers. Interestingly, in other analyzed genes, we found truncating variants in the genes coding the P450 enzymes of steroid hormones metabolism in 5% of BC patients. Therefore, this functional group may contribute to the explanation of so far undisclosed missing heritability in some high-risk BC patients. We are aware that exact role of both c.1096_1099dupATTA and *CYP* genes in BC susceptibility needs to be further clarified by independent and larger studies.

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's web-site.

Acknowledgements

The authors wish to thank all individuals who participated in this study. This study was supported by the grant of Ministry of Health of the Czech Republic NT14054 and by BIOCEV CZ.1.05/1.1.00/02.0109 and Charles University projects PRVOUK-P27/LF1/1, UK 260148/2015, and UNCE 204011.

References

1. Robson M, Offit K. Clinical practice. Management of an inherited predisposition to breast cancer. N Engl J Med 2007: 357 (2): 154–162.
2. Couch FJ, Nathanson KL, Offit K. Two decades after BRCA: setting paradigms in personalized cancer care and prevention. Science 2014: 343 (6178): 1466–1470.
3. Rahman N. Realizing the promise of cancer predisposition genes. Nature 2014: 505 (7483): 302–308.
4. Karami F, Mehdipour P. A comprehensive focus on global spectrum of BRCA1 and BRCA2 mutations in breast cancer. Biomed Res Int 2013: 2013: 928562.
5. Easton DF, Pharoah PD, Antoniou AC et al. Gene-panel sequencing and the prediction of breast-cancer risk. N Engl J Med 2015: 372 (23): 2243–2257.
6. Kean S. Breast cancer. The 'other' breast cancer genes. Science 2014: 343 (6178): 1457–1459.
7. Ripperger T, Gadzicki D, Meindl A, Schlegelberger B. Breast cancer susceptibility: current knowledge and implications for genetic counselling. Eur J Hum Genet 2009: 17 (6): 722–731.
8. Pohlreich P, Stribrna J, Kleibl Z et al. Mutations of the BRCA1 gene in hereditary breast and ovarian cancer in the Czech Republic. Med Princ Pract 2003: 12 (1): 23–29.
9. Pohlreich P, Zikan M, Stribrna J et al. High proportion of recurrent germline mutations in the BRCA1 gene in breast and ovarian cancer patients from the Prague area. Breast Cancer Res 2005: 7 (5): R728–R736.
10. Ticha I, Kleibl Z, Stribrna J et al. Screening for genomic rearrangements in BRCA1 and BRCA2 genes in Czech high-risk breast/ovarian cancer patients: high proportion of population specific alterations in BRCA1 gene. Breast Cancer Res Treat 2010: 124 (2): 337–347.
11. Janatova M, Kleibl Z, Stribrna J et al. The PALB2 gene is a strong candidate for clinical testing in BRCA1- and BRCA2-negative hereditary breast cancer. Cancer Epidemiol Biomarkers Prev 2013: 22 (12): 2323–2332.
12. Kleibl Z, Havranek O, Hlavata I et al. The CHEK2 gene I157T mutation and other alterations in its proximity increase the risk of sporadic colorectal cancer in the Czech population. Eur J Cancer 2009: 45 (4): 618–624.
13. Janatova M, Soukupova J, Stribrna J et al. Mutation analysis of the RAD51C and RAD51D genes in high-risk ovarian cancer patients and families from the Czech Republic. PLoS One 2015: 10 (6): e0127711.
14. Mateju M, Stribrna J, Zikan M et al. Population-based study of BRCA1/2 mutations: family history based criteria identify minority of mutation carriers. Neoplasma 2010: 57 (3): 280–285.
15. Yu W, Clyne M, Khoury MJ, Gwinn M. Phenopedia and Genopedia: disease-centered and gene-centered views of the evolving knowledge of human genetic associations. Bioinformatics 2010: 26 (1): 145–146.
16. Harakalova M, Mokry M, Hrdlickova B et al. Multiplexed array-based and in-solution genomic enrichment for flexible and cost-effective targeted next-generation sequencing. Nat Protoc 2011: 6 (12): 1870–1886.
17. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 2010: 38 (16): e164.

18. Sevcik J, Falk M, Macurek L et al. Expression of human BRCA1 Delta 17–19 alternative splicing variant with a truncated BRCT domain in MCF-7 cells results in impaired assembly of DNA repair complexes and aberrant DNA damage response. Cell Signal 2013: 25 (5): 1186–1193.
19. Ali AM, Kirby M, Jansen M et al. Identification and characterization of mutations in FANCL gene: a second case of Fanconi anemia belonging to FA-L complementation group. Hum Mutat 2009: 30 (7): E761–E770.
20. Schroeder C, Faust U, Sturm M et al. HBOC multi-gene panel testing: comparison of two sequencing centers. Breast Cancer Res Treat 2015: 152 (1): 129–136.
21. Kottemann MC, Smogorzewska A. Fanconi anaemia and the repair of Watson and Crick DNA crosslinks. Nature 2013: 493 (7432): 356–363.
22. Walden H, Deans AJ. The Fanconi anemia DNA repair pathway: structural and functional insights into a complex disorder. Annu Rev Biophys 2014: 43 (1): 257–278.
23. Meetei AR, de Winter JP, Medhurst AL et al. A novel ubiquitin ligase is deficient in Fanconi anemia. Nat Genet 2003: 35 (2): 165–170.
24. Meetei AR, Yan Z, Wang W. FANCL replaces BRCA1 as the likely ubiquitin ligase responsible for FANCD2 monoubiquitination. Cell Cycle 2004: 3 (2): 179–181.
25. Akbari MR, Malekzadeh R, Lepage P et al. Mutations in Fanconi anemia genes and the risk of esophageal cancer. Hum Genet 2011: 129 (5): 573–582.
26. Delahaye-Sourdeix M, Anantharaman D, Timofeeva MN et al. A rare truncating BRCA2 variant and genetic susceptibility to upper aerodigestive tract cancer. J Natl Cancer Inst 2015: 107 (5): djv037.
27. Kiiski JI, Pelttari LM, Khan S et al. Exome sequencing identifies FANCM as a susceptibility gene for triple-negative breast cancer. Proc Natl Acad Sci U S A 2014: 111 (42): 15172–15177.
28. Peterlongo P, Catucci I, Colombo M et al. FANCM c.5791C>T nonsense mutation (rs144567652) induces exon skipping, affects DNA repair activity and is a familial breast cancer risk factor. Hum Mol Genet 2015: 24 (18): 5345–5355.
29. Smith CG, Naven M, Harris R et al. Exome resequencing identifies potential tumor-suppressor genes that predispose to colorectal cancer. Hum Mutat 2013: 34 (7): 1026–1034.
30. Loveday C, Turnbull C, Ruark E et al. Germline RAD51C mutations confer susceptibility to ovarian cancer. Nat Genet 2012: 44 (5): 475–476.
31. Couch FJ, Hart SN, Sharma P et al. Inherited mutations in 17 breast cancer susceptibility genes among a large triple-negative breast cancer cohort unselected for family history of breast cancer. J Clin Oncol 2015: 33 (4): 304–311.
32. Rasmussen LJ, Rasmussen M, Lee B et al. Identification of factors interacting with hMSH2 in the fetal liver utilizing the yeast two-hybrid system. *In vivo* interaction through the C-terminal domains of hEXO1 and hMSH2 and comparative expression analysis. Mutat Res 2000: 460 (1): 41–52.
33. Wu Y, Berends MJ, Post JG et al. Germline mutations of EXO1 gene in patients with hereditary nonpolyposis colorectal cancer (HNPCC) and atypical HNPCC forms. Gastroenterology 2001: 120 (7): 1580–1587.
34. Blackburn HL, Ellsworth DL, Shriver CD, Ellsworth RE. Role of cytochrome P450 genes in breast cancer etiology and treatment: effects on estrogen biosynthesis, metabolism, and response to endocrine therapy. Cancer Causes Control 2015: 26 (3): 319–332.
35. Hiort O, Holterhus PM, Werner R et al. Homozygous disruption of P450 side-chain cleavage (CYP11A1) is associated with prematurity, complete 46, XY sex reversal, and severe adrenal failure. J Clin Endocrinol Metab 2005: 90 (1): 538–541.
36. Hwang DY, Hung CC, Riepe FG et al. CYP17A1 intron mutation causing cryptic splicing in 17alpha-hydroxylase deficiency. PLoS One 2011: 6 (9): e25492.
37. Hopper JL, Hayes VM, Spurdle AB et al. A protein-truncating mutation in CYP17A1 in three sisters with early-onset breast cancer. Hum Mutat 2005: 26 (4): 298–302.
38. Yang X, Wu J, Lu J et al. Identification of a comprehensive spectrum of genetic factors for hereditary breast cancer in a Chinese population by next-generation sequencing. PLoS One 2015: 10 (4): e0125571.
39. Gaikwad NW, Yang L, Rogan EG, Cavalieri EL. Evidence for NQO2-mediated reduction of the carcinogenic estrogen ortho-quinones. Free Radic Biol Med 2009: 46 (2): 253–262.
40. Caiazza F, Ryan EJ, Doherty G, Winter DC, Sheahan K. Estrogen receptors and their implications in colorectal carcinogenesis. Front Oncol 2015: 5: 19.
41. Pacifico L, Carducci C, Poggiogalle E et al. Mutational analysis of ABCC2 gene in two siblings with neonatal-onset Dubin Johnson syndrome. Clin Genet 2010: 78 (6): 598–600.
42. Kleibl Z, Fidlerova J, Kleiblova P et al. Influence of dihydropyrimidine dehydrogenase gene (DPYD) coding sequence variants on the development of fluoropyrimidine-related toxicity in patients with high-grade toxicity and patients with excellent tolerance of fluoropyrimidine-based chemotherapy. Neoplasma 2009: 56 (4): 303–316.
43. Blackburn HL, Schroeder B, Turner C, Shriver CD, Ellsworth DL, Ellsworth RE. Management of incidental findings in the era of next-generation sequencing. Curr Genomics 2015: 16 (3): 159–174.
44. Christenhusz GM, Devriendt K, Dierickx K. Disclosing incidental findings in genetics contexts: a review of the empirical ethical research. Eur J Med Genet 2013: 56 (10): 529–540.