Příloha 1 – Plný text manuskriptu: Lhotova, K., L. Stolarova, P. Zemankova, M. Vocka, M. Janatova, M. Borecka, M. Cerna, S. Jelinkova, J. Kral, Z. Volkova, M. Urbanova, P. Kleiblova, E. Machackova, L. Foretova, J. Hazova, P. Vasickova, F. Lhota, M. Koudova, L. Cerna, S. Tavandzis, J. Indrakova, L. Hruskova, M. Kosarova, R. Vrtel, V. Stranecky, S. Kmoch, M. Zikan, L. Macurek, Z. Kleibl, and J. Soukupova. 2020. "***Multigene Panel Germline Testing of 1333 Czech Patients with Ovarian Cancer***." *Cancers (Basel)* 12 (4). doi: 10.3390/cancers12040956.

Article

Multigene Panel Germline Testing of 1333 Czech Patients with Ovarian Cancer

Klara Lhotova [1,2], Lenka Stolarova [1], Petra Zemankova [1,2], Michal Vocka [3], Marketa Janatova [1,2], Marianna Borecka [1,2], Marta Cerna [1], Sandra Jelinkova [1], Jan Kral [1], Zuzana Volkova [1], Marketa Urbanova [2], Petra Kleiblova [2], Eva Machackova [4], Lenka Foretova [4], Jana Hazova [4], Petra Vasickova [4], Filip Lhota [5], Monika Koudova [5], Leona Cerna [5], Spiros Tavandzis [6], Jana Indrakova [6], Lucie Hruskova [7], Marcela Kosarova [8], Radek Vrtel [9], Viktor Stranecky [10], Stanislav Kmoch [10], Michal Zikan [11], Libor Macurek [12], Zdenek Kleibl [1] and Jana Soukupova [1,2,*]

[1] Institute of Biochemistry and Experimental Oncology, First Faculty of Medicine, Charles University, 128 53 Prague, Czech Republic; Klara.Lhotova@lf1.cuni.cz (K.L.); lenka.stolarova@lf1.cuni.cz (L.S.); petra.boudova@lf1.cuni.cz (P.Z.); mjana@lf1.cuni.cz (M.J.); boreckam@gmail.com (M.B.); marta.cerna@lf1.cuni.cz (M.C.); sandra.jelinkova@lf1.cuni.cz (S.J.); jan.kral@lf1.cuni.cz (J.K.); zuzana.klusonova@lf1.cuni.cz (Z.V.); zdekleje@lf1.cuni.cz (Z.K.)

[2] Institute of Biology and Medical Genetics, First Faculty of Medicine, Charles University and General University Hospital in Prague, 128 00 Prague, Czech Republic; Marketa.Urbanova@vfn.cz (M.U.); pekleje@lf1.cuni.cz (P.K.)

[3] Department of Oncology, First Faculty of Medicine, Charles University and General University Hospital in Prague, 128 08 Prague, Czech Republic; Michal.Vocka@vfn.cz

[4] Department of Cancer Epidemiology and Genetics, Masaryk Memorial Cancer Institute, 656 53 Brno, Czech Republic; emachack@mou.cz (E.M.); foretova@mou.cz (L.F.); hazova@mou.cz (J.H.); vasickova@mou.cz (P.V.)

[5] Department of Medical Genetics, Centre for Medical Genetics and Reproductive Medicine, Gennet, 170 00 Prague, Czech Republic; Filip.Lhota@gennet.cz (F.L.); Monika.Koudova@gennet.cz (M.K.); Leona.Cerna@gennet.cz (L.C.)

[6] Department of Medical Genetics, AGEL Laboratories, AGEL Research and Training Institute, 741 01 Novy Jicin, Czech Republic; spiros.tavandzis@lab.agel.cz (S.T.); jana.indrakova@lab.agel.cz (J.I.)

[7] Department of Medical Genetics, GHC Genetics, 110 00 Prague, Czech Republic; hruskova@ghc.cz

[8] Department of Medical Genetics, Pronatal, 147 00 Prague, Czech Republic; kosarova@pronatal.cz

[9] Department of Medical Genetics, University Hospital Olomouc, Faculty of Medicine and Dentistry, Palacky University Olomouc, 779 00 Olomouc, Czech Republic; radek.vrtel@fnol.cz

[10] Research Unit for Rare Diseases, Department of Pediatrics and Adolescent Medicine, First Faculty of Medicine, Charles University and General University Hospital in Prague, 12808 Prague, Czech Republic; vstra@lf1.cuni.cz (V.S.); skmoch@lf1.cuni.cz (S.K.)

[11] Department of Gynecology and Obstetrics, Hospital Na Bulovce and First Faculty of Medicine, Charles University, 180 81 Prague, Czech Republic; michal.zikan@lf1.cuni.cz

[12] Laboratory of Cancer Cell Biology, Institute of Molecular Genetics of the Czech Academy of Sciences, 142 20 Prague, Czech Republic; libor.macurek@img.cas.cz

* Correspondence: jana.soukupova@lf1.cuni.cz; Tel.: +420-224965739

Received: 15 March 2020; Accepted: 10 April 2020; Published: 13 April 2020

Abstract: Ovarian cancer (OC) is the deadliest gynecologic malignancy with a substantial proportion of hereditary cases and a frequent association with breast cancer (BC). Genetic testing facilitates treatment and preventive strategies reducing OC mortality in mutation carriers. However, the prevalence of germline mutations varies among populations and many rarely mutated OC predisposition genes remain to be identified. We aimed to analyze 219 genes in 1333 Czech OC patients and 2278 population-matched controls using next-generation sequencing. We revealed germline mutations in 18 OC/BC predisposition genes in 32.0% of patients and in 2.5% of controls. Mutations in *BRCA1/BRCA2*, *RAD51C/RAD51D*, *BARD1*, and mismatch repair genes conferred high OC risk (OR > 5). Mutations in

BRIP1 and *NBN* were associated with moderate risk (both OR = 3.5). *BRCA1/2* mutations dominated in almost all clinicopathological subgroups including sporadic borderline tumors of ovary (BTO). Analysis of remaining 201 genes revealed somatic mosaics in *PPM1D* and germline mutations in *SHPRH* and *NAT1* associating with a high/moderate OC risk significantly; however, further studies are warranted to delineate their contribution to OC development in other populations. Our findings demonstrate the high proportion of patients with hereditary OC in Slavic population justifying genetic testing in all patients with OC, including BTO.

Keywords: ovarian cancer; next-generation sequencing; predisposition genes; cancer risk; mutation

1. Introduction

Ovarian cancer (OC) is the most severe gynecologic malignancy with stable incidence and mortality. The most frequent OC types (85–95%) are epithelial tumors, which are high-grade (HG) serous in 70% of cases [1,2]. Because of the nonspecific symptoms and a lack of presymptomatic screening modalities, most women are diagnosed with an advanced disease, having a dismal 25% 5-year survival rate [3].

The overall OC lifetime risk oscillates around 2% in the general female population in developed countries. Central and Eastern Europe, including the Czech Republic, represented a region with the highest OC incidence (11.9 ASRW per 100,000 females) and mortality (6.0 ASRW per 100,000 females) worldwide in 2018 (http://gco.iarc.fr). In the Czech Republic alone, annual OC incidence and mortality in 2018 reached 9.5 and 6.7 ASRW per 100,000 females, respectively.

Genetic predisposition for OC is unusually high and is reported in up to 25% of cases [4–6]. The most frequent germline mutations affect the *BRCA1* and *BRCA2* genes, conferring 24% and 8.4% OC lifetime risks, respectively [7]. The *BRCA1* and *BRCA2* mutation carriers frequently but not exclusively develop HG serous OC [8]. Carriers of mutations in these major OC predisposition genes have also very high risk of breast cancer (BC) development. A high OC risk has also been associated with germline mutations in *RAD51C, RAD51D,* Lynch syndrome genes, and *STK11*; a moderate OC risk with *BRIP1* [9–13]. Risks associated with germline mutations in genes with anticipated BC and/or OC predisposition (incl. *ATM, BARD1, CDH1, CHEK2, NBN, PALB2, PTEN,* and *TP53*) and in other candidate genes remain to be determined [14–17]. The identification of presymptomatic women at high risk who can benefit from risk-reducing salpingo-oophorectomy (RRSO) is of critical importance, as demonstrated by the reduced OC mortality in *BRCA1* and *BRCA2* mutation carriers undergoing preventive surgery [18].

In this report, we aim to establish an association of germline mutations with OC in the Czech patients belonging to the Slavic population that has not been systematically analyzed for OC predisposition. Seven Czech genetic laboratories participated in the analysis of 1333 Czech OC patients by the identical procedure using CZECANCA panel (CZEch CAncer paNel for Clinical Application) targeting 219 genes [19]. Prevalence of variants in genes affected in OC patients was assessed in 2278 population-matched controls. This analysis enabled us to comprehensively determine mutations frequency and clinicopathological characteristics of OC in carriers of mutations in genes with known OC predisposition but also to analyze contribution of population-specific variants in other candidate genes to OC predisposition.

2. Results

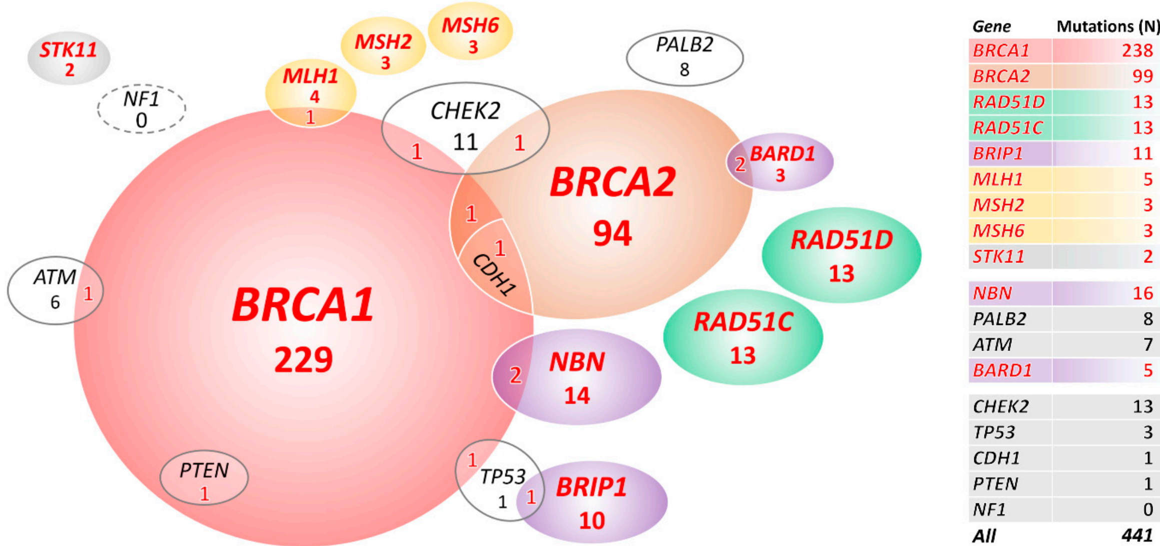
2.1. Description of Study Population

Altogether, samples obtained from 1333 OC patients diagnosed at seven centers were analyzed by the identical panel NGS using the CZECANCA panel targeting 219 cancer-predisposition and candidate genes and were evaluated centrally by the identical bioinformatics pipeline. From 1333

analyzed OC patients, 1045 (78.4%) women were diagnosed with OC only and 288 (21.6%) women with double primary tumors, including BC (210 patients; 15.8%) or other tumors (78 patients; 5.9%). The median age at OC diagnosis was 53.7 years (range 15–86 years). Almost half (47.6%) of the patients had a negative family cancer history. From 1120 OC patients with known histology, 728 (65.0%) women developed serous adenocarcinoma with prevailing HG tumors. Sixty percent of cases represented patients with advanced disease (stages III–IV). The clinicopathological characteristics are provided in Table S1.

2.2. Mutations in 18 Known/Anticipated Hereditary BC/OC Genes

We primarily focused on mutations in 18 BC/OC genes listed in the NCCN Guidelines for Genetic/Familial High-Risk Assessment: Breast, Ovarian, and Pancreatic (Version 1.2020; 4 December 2019). We identified 441 mutations in 427/1333 (32.0%) OC patients and 58/2278 (2.5%) mutation carriers among population-matched controls (PMC) in 18 known/anticipated BC/OC genes (Figure 1, Table 1, and Table S2). Thirteen multiple mutation carriers (Figure 1) identified among patients only (characterized in Table S3) were excluded from the subsequent analyses.



Gene	Mutations (N)
BRCA1	238
BRCA2	99
RAD51D	13
RAD51C	13
BRIP1	11
MLH1	5
MSH2	3
MSH6	3
STK11	2
NBN	16
PALB2	8
ATM	7
BARD1	5
CHEK2	13
TP53	3
CDH1	1
PTEN	1
NF1	0
All	***441***

Figure 1. Overall, 427 mutation carriers of 441 mutations in 18 known/anticipated breast cancer (BC)/ovarian cancer (OC) predisposition genes. In total, 399 carriers in genes significantly associated with OC in our study are highlighted in red letters. *STK11* is highlighted as rarely mutated but established OC predisposition gene.

Carriers of germline mutations in 10 genes (including Lynch syndrome genes analyzed as a group together) had significantly increased OC risk (Table 1 in bold). We found the prevailing *BRCA1* or *BRCA2* germline alterations in 323/1320 (24.5%) patients and in 12/2278 (0.5%) PMC. Further, 65/1320 (4.9%) OC patients carried a mutation in 8 other genes significantly associated with OC risk in our study (including 2 carriers of mutations in *STK11*, an established high-risk OC gene that did not reach significant association in our study due to low frequency of mutation carriers in patients; Figure 1). We found only 19/2278 (0.8%) carriers of mutations in these 8 genes in PMC.

Table 1. Mutation frequencies in 1320 ovarian cancer cases and in 2278 population-matched controls (PMC).

Gene	1320 OC Patients [a] N Mutations (%)	2278 PMC N mutations (%)	OR (95% CI); p [a]
Increased OC risk [b]			
BRCA1 **[c]**	**229 (17.35)**	**5 (0.22)**	**95.2 (40.1–295.2); 1.83 × 10^{-97}**
BRCA2 **[c]**	**94 (7.12)**	**7 (0.31)**	**24.9 (11.6–63.6); 1.16 × 10^{-33}**
RAD51D	**13 (0.98)**	**2 (0.09)**	**11.3 (2.6–103.4); 9.66 × 10^{-5}**
RAD51C	**13 (0.98)**	**4 (0.18)**	**5.7 (1.7–23.8); 0.001**
BRIP1 **[c]**	**10 (0.76)**	**5 (0.22)**	**3.5 (1.1–13); 0.03**
MLH1 **[c]**	**4 (0.3)**	**1 (0.04)**	6.9 (0.7–340.4); 0.06 [d]
MSH2	**3 (0.23)**	**0**	**0.049 [d]**
MSH6	**3 (0.23)**	**0**	**0.049 [d]**
STK11	2 (0.15)	0	0.13
Potentially increase or insufficient evidence OC risk [b]			
NBN **[c]**	**14 (1.06)**	**7 (0.31)**	**3.5 (1.3–10.2); 0.006**
PALB2	8 (0.61)	9 (0.40)	1.5 (0.5–4.5); 0.45
ATM [c]	6 (0.45)	8 (0.35)	1.3 (0.4–4.3); 0.78
BARD1 **[c]**	**3 (0.23)**	**0**	**0.049**
No increased risk of OC [b]			
CHEK2 [c]	11 (0.83)	8 (0.35)	2.4 (0.9–6.8); 0.06
TP53 [c]	1 (0.08)	2 (0.09)	0.9 (0–16.6); 1
CDH1 [c]	0	0	-
PTEN [c]	0	0	-
NF1	0	0	-

[a] Prevalence of mutations in all 1333 patients (including 13 multiple mutation carriers) is provided in Table S2. [b] Gene classification according to the NCCN guidelines version 2020.1. [c] Excluding 13 multiple mutation carriers described in Figure 1 and Table S3. [d] When analyzed Lynch syndrome genes collectively: OR = 22.63 (95% CI 3.4–958.5); $p = 1.95 \times 10^{-05}$.

The copy number variation (CNV) analysis in 18 OC/BC genes revealed 37 large genomic rearrangements in 37/1333 (2.8%) patients. They affected seven genes (23×*BRCA1*, 4×*BRIP1*, 4×*CHEK2*, 2×*MLH1*, 2×*STK11*, 1×*PALB2*, and 1×*CDH1*) and accounted for 8.4% (37/441) of all pathogenic mutations in these genes. Except 1 whole gene duplication of *MSH6* (classified as VUS), we found no CNV in analyzed controls in these 18 genes.

2.3. Clinical and Histopathological Characteristics of Mutation Carriers

Subsequently, we described the clinicopathological characteristics of the mutation carriers in 10 genes associated with OC risk (Figure 2 and Table S4). Multiple mutation carriers (Table S3) were excluded from this analysis.

2.3.1. Age at OC Diagnosis

The highest mutation frequency was found in patients diagnosed with OC at 40–49 and 50–59 years (37.4% and 40.7%, respectively) and the lowest in patients diagnosed before the age of 30 (8.3%; Figure 2A). Interestingly, the mutation frequency in the group of the oldest patients (≥70 years) was twice higher than in the youngest (<30 years) patients' subgroup (p = 0.013 for difference).

This difference was primarily caused by *BRCA1*/*BRCA2* mutations (3.6% vs. 18.1% in patients <30 vs. ≥70 years), as the frequency of *non-BRCA* genes mutations was similar (4.8% vs. 4.3%). The median age at diagnosis was significantly different in *BRCA1* (51.0 years; range 23–78) and *BRCA2* (58.4 years; range 27–78) mutation carriers ($p = 8.5\times10^{-10}$), respectively. The median age at diagnosis in other genes with at least 10 identified mutation carriers increased gradually from *RAD51C* (52.2 years; range 25–69) to *NBN* (54.5 years; range 18–76), *RAD51D* (56.0 years; range 36–69), and *BRIP1* (58.0 years; range 30–71). We observed a younger median age at diagnosis in carriers of mutations in Lynch syndrome genes 46.0 years (range 35–73).

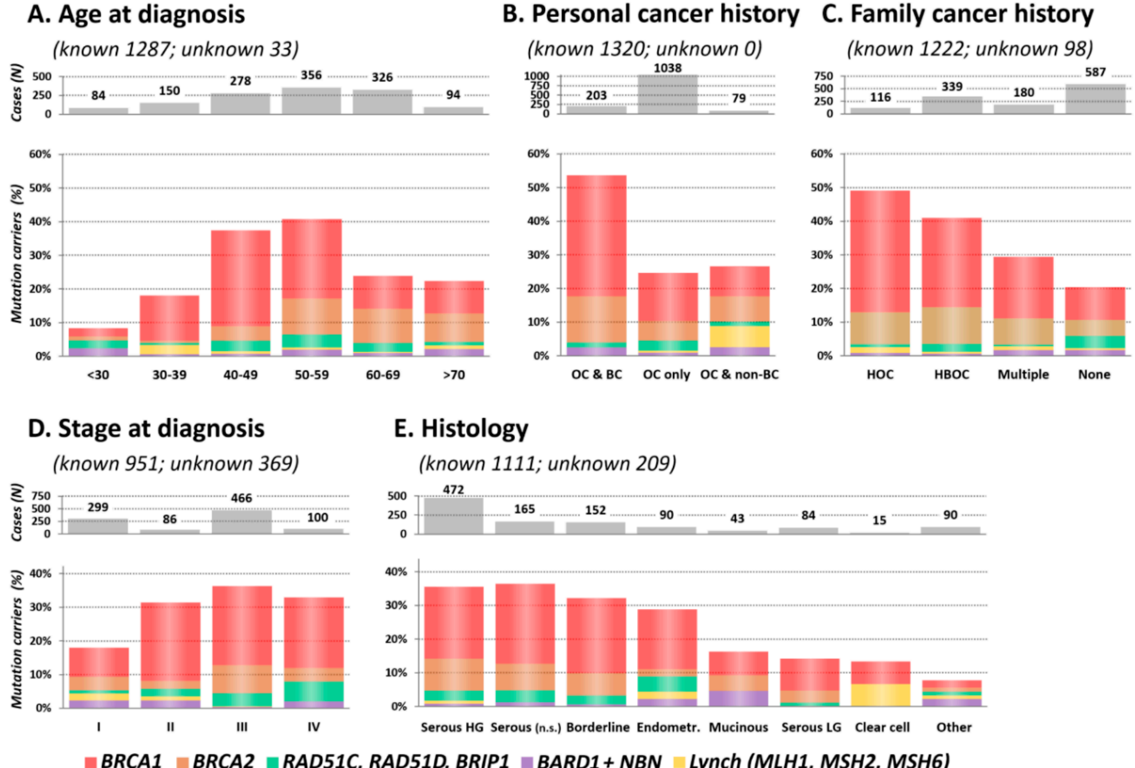


Figure 2. Proportion of mutation carriers in clinicopathological subgroups, including (**A**) Age at OC diagnosis; (**B**) Personal cancer history; (**C**) Family cancer history; (**D**) Stage at diagnosis; (**E**) Histology in 1320 OC patients.

2.3.2. Personal and Family Cancer History

The highest proportion of mutations (109/203; 53.7%) was detected in double primary OC and BC patients, while in patients diagnosed with OC only and double primary OC and non-BC cancer, it reached 256/1038 (24.7%) and 21/79 (26.6%), respectively (Figure 2B). The frequency of mutations in patients from hereditary OC families (HOC) was 49.1% (57/116; Figure 2C). Decreasing proportion of mutation carriers in other family cancer history categories (41.0% in HBOC and 29.4% in multiple cancer) was dominantly caused by decreasing *BRCA1* mutation prevalence. Nevertheless, in 587 OC patients without a positive family cancer history, we still identified 120 (20.4%) carriers of pathogenic mutations.

2.3.3. Stage and Histology

Almost 60% of patients were diagnosed at FIGO stage III or IV (Figure 2D). In contrast, 6/8 informative Lynch syndrome gene mutation carriers were diagnosed with stage I tumors.

The mutation rate stratified OC into two histological clusters. The high mutation rate subgroup included 879 patients with HG/unspecified serous, borderline, and endometrioid tumors with 303 (34.5%) carriers, while the low mutation rate subgroup included 232 patients with low-grade (LG)

serous, mucinous, clear cell, and other tumors with 28 (12.1%) carriers. *BRCA1/2* mutations in HG serous carcinomas were more than twice as frequent (146/472; 30.9%) as in LG serous ones (11/84; 13.1%). Interestingly, the distribution of *BRIP1/RAD51C/RAD51D* mutations among histological types was similar to that of *BRCA1/2*. The lowest proportion of mutations (7/90; 7.8%) was found in rare histological cancer types (herein denominated as "Other").

2.4. Mutations in Additional 201 Analyzed Genes

Finally, we reviewed the presence of germline variants in additional 201 genes targeted by the CZECANCA panel [19]. This analysis revealed 230 mutations in 89 genes in 208 (15.6%) patients (Table S5). Of these, 149 (11.2%) patients carried mutations in "additional" genes exclusively while 59 (4.4%) patients carried a mutation in "additional" genes alongside a mutation in one of the 10 OC risk genes. Mutations in these "additional" genes were rare and their prevalence was significantly higher in patients over controls in only four genes (Table 2). However, only mutations in *PPM1D* were significantly associated with OC risk ($p = 0.003$) following Bonferroni correction and exclusion of carriers of mutations in OC predisposition genes. All *PPM1D* mutations were mosaic with MAF = 14%–60% and MAF = 17%–19% in patients and controls, respectively. It should be noted that blood for genetic testing was sampled after the application of chemotherapy in all *PPM1D* positive patients (in average at 38 months after treatment; ranged 4 months–7.1 years). Seven out of 15 *PPM1D* mutation carriers harbored an additional mutation in another DNA repair gene (3×*BRCA2*, 1×*PALB2*, 1×*EXO1*, and 1×*PMS1*). MAF of *PPM1D* mutations correlated neither with age at OC diagnosis nor with the time from the last chemotherapy (Table S6). Mutations in *PPM1D* and *SHPRH* were significantly associated only with age > 60 years ($p = 0.001$), whereas frequency of *NAT1* mutations in particular categories was similar (Table S7). Uncorrected p values were marginally significant also for germline variants in *MMP8* and *FANCG* in OC patients when carriers of mutations in 10 BC/OC predisposition genes significantly associating with OC risk in our study were excluded (Table 2).

Table 2. Additional 201 analyzed genes significantly associated with OC risk in the group of all OC patients and in a subgroup of 934 patients without mutations in 10 established OC predisposition genes.

Gene	Patients N Mutations (%)	2278 PMC N Mutations (%)	OR (95% CI); p (Bonferroni Corrected *p*)
All 1333 OC patients			
PPM1D	16 (1.20)	2 (0.09)	13.82 (3.24–124.22); 7.4×10^{-6} (0.001)
NAT1	13 (0.98)	5 (0.22)	4.48 (1.49–16.07); 0.003 (n.s.)
SHPRH	5 (0.38)	1 (0.04)	8.57 (0.96–404.83); 0.028 (n.s.)
934 OC patients without mutations in 10 genes significantly associated with OC in our study			
PPM1D	12 (1.28)	2 (0.09)	14.80 (3.28–136.67); 1.7×10^{-5} (0.003)
NAT1	8 (0.86)	5 (0.22)	3.96 (1.13–15.30); 0.026 (n.s.)
MMP8	6 (0.64)	4 (0.18)	3.67 (0.87–17.74); 0.041 (n.s.)
FANCG	5 (0.53)	2 (0.09)	6.12 (1.00–64.45); 0.025 (n.s.)

n.s., nonsignificant.

3. Discussion

The analysis of 1333 Czech OC patients and 2278 population-matched controls provides the most comprehensive view of the genetic architecture of OC predisposition in the Slavic population. From 18 OC/BC predisposition genes listed in current NCCN breast/ovarian familial cancer guidelines, mutations in 10 genes were significantly associated with OC risk in our population being present in 399/1333 (29.9%) OC patients and 31/2278 (1.4%) PMC (Figure 1). Mutations in eight remaining genes were extremely rare (*CDH1*, *PTEN*, *STK11*, and *TP53*) or absent (*CDKN2A* and *NF1*) or did not significantly differ in frequency among cases and controls (*ATM*, *PALB2*, and *CHEK2*). Mutations in *BRCA1/2*, *RAD51C/D*, and Lynch syndrome genes were associated with a high OC risk, while mutations in *BRIP1* were associated with a moderate OC risk in our study (Table 1), in concordance with previous

reports [9,10,20,21]. The *BRCA1* and *BRCA2* mutations, present in 84.0% of all mutation carriers, were by far the most frequent alterations found in 17.9% and 7.4% of our patients, respectively. Mutations in other eight genes leaded by *RAD51C/RAD51D/BRIP1* affected additional 5.0% of patients, as shown also by others recently [5,6,22]. Germline mutations in Lynch syndrome genes together associated with high OC risk. Mutations in *MLH1* prevailed similarly as in Lynch syndrome patients diagnosed with colorectal cancer [23].

In contrast to previous studies, our results suggest increased OC risk in carriers of *NBN* and *BARD1* mutations [12,24]. We did not find significant increase of OC risk for carriers of mutations in *ATM* and *PALB2*, as noticed previously [12,24,25]. However, further analyses considering very large population-matched studies or studies considering families of mutation carriers can better disclose moderate risk associations, as shown for *PALB2* mutations recently [26].

Overrepresentation of mutations in the *CHEK2* gene in OC patients in this study was marginally nonsignificant in contrast to our previous report where we identified moderately increased OC risk for *CHEK2* mutation carriers [27]. However, last four *CHEK2* coding exons were not targeted in our gene panel omitting possible deleterious *CHEK2* alterations identified in our previous study in which last four coding exons were analyzed separately in both cases and controls. Mutations in *NF1* were absent and were extremely rare in *CDH1* and *PTEN*, just like *STK11* mutations found in a patient with nonepithelial OC, a characteristic Peutz–Jeghers syndrome manifestation [9]. Altogether, the high overall frequency of mutations in OC predisposition genes in our study is in agreement with some previous studies [4–6,28] and may contribute to a high OC incidence in our population.

Multigene testing revealed 13 carriers of multiple pathogenic mutations (1.0% of patients). Similar frequency of individuals with this multilocus inherited neoplasia alleles syndrome (MINAS) [29] was shown also in previous analyses of OC patients [30,31].

We analyzed available phenotype characteristics in 1320 OC patients with one pathogenic mutation at the most in 10 genes associated with OC risk in our study (Figure 2). While the highest prevalence of *BRCA1/2* mutation carriers was in patients diagnosed with double primary OC and BC, mutations in *RAD51C/RAD51D/BRIP1* prevailed in patients diagnosed with OC only (Figure 2B); nevertheless, their distribution among histological subtypes was similar to that in *BRCA1/2* mutation carriers (Figure 2E). In contrast to Castera et al. who found mutations in *RAD51C/RAD51D/BRIP1* dominantly in French OC patients with a positive family OC history [32], we identified mutations in these genes in 1/116 (0.9%) and 22/587 (3.7%) carriers in HOC patients and in patients with a negative family cancer history, respectively. Further, we have noticed a surprisingly high frequency of OC-predisposing mutations in older patients. Their prevalence in patients ≥ 60 years was 23.6%, whereas Harter et al. found in this age group 18.9% mutation carriers even though frequency of mutation carriers in patients <60 years in both studies was comparable (32.6% and 33.2%, respectively) [28]. *BRCA1* mutations dominated in patients <60 years over *BRCA2* mutations, while in patients ≥ 60 years, their frequencies were comparable. Moreover, we revealed 29 *BRCA1/2* mutation carriers (13.9% of patients) in 208 OC patients diagnosed at ≥60 years with no family cancer history, while Morgan and colleagues detected only two (4.3%) *BRCA1/2* mutations in 46 sporadic OC patients ≥ 60 years [33]. Even in the oldest subgroup of our OC patients diagnosed at ≥70 years, the frequency of *BRCA1/2* mutation carriers exceeded 18%, while in other studies, *BRCA1/2* mutations' frequency in this age category was below 10% [34,35]. This high frequency of *BRCA1/2* mutations in our patients ≥70 years contrasted with a low frequency in women diagnosed at <30 years (18.1% vs. 3.6%; $p = 0.003$; Figure 2A). The difference was even more apparent in "sporadic" OC cases (with no family cancer history), where *BRCA1/2* mutations were found in 6 out of 45 (13.3%) women ≥70 years but in none of 52 cases diagnosed at <30 years. It should be emphasized that although rare histological OC types were more frequent in the subgroup of 52 patients diagnosed with sporadic OC at <30 years, 32 (65.3%) of 49 informative cases developed invasive epithelial OC.

Mutations in OC predisposition genes significantly prevailed in subgroups with high-grade/nonspecified serous, borderline, and endometrioid tumors over subgroup with low-grade serous,

mucinous, clear cell, or other rare histologic types (Figure 2E). Surprisingly, the overall mutation frequency in patients with borderline tumors was comparable with that of in HG serous OC (32.2% and 36.7%, respectively; Figure 2E). Thus, we compared mutation frequency in patients with no family cancer history diagnosed with these histological tumor types, and we found that although the mutation frequency in sporadic borderline tumors was half in comparison to sporadic HG serous (Figure 3), it still largely exceeded 10% in both hereditary and sporadic cases, justifying the genetic testing of borderline tumors. The large proportion of borderline tumors with positive family cancer history in our study suggested that this OC subtypes belong to a possible manifestation of a cancer predisposition. However, our observation needs to be confirmed in other populations as current reports about borderline tumors in *BRCA1/2* mutation carriers are limited.

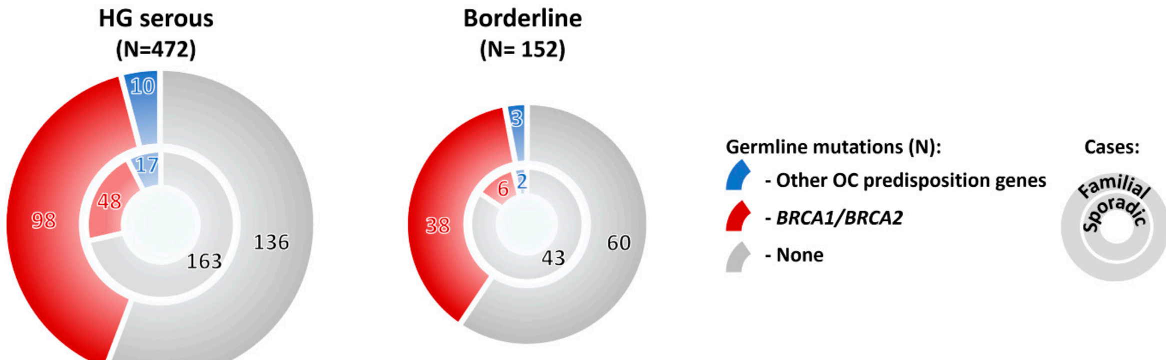


Figure 3. Frequency of mutations in 10 BC/OC predisposition genes significantly associated with OC in our study in OC patients with high-grade (HG) serous and borderline tumors, respectively. The patients were subdivided into subgroups with positive (familial cases) and negative (sporadic cases) family cancer history, respectively.

The multigene panel enabled us to identify other candidate genes associating with increased OC risk. We noticed many rare truncating variants episodically affecting various genes and clustering into *PPM1D*, *NAT1*, and *SHPRH* in OC patients. The *PPM1D* gene, coding for WIP1 phosphatase, was the only candidate associated with OC risk following multiple testing correction. Similarly to the previous studies describing its mosaic variants in OC patients [36–38], we also found mosaic gain-of-function mutations resulting in increased WIP1 phosphatase activity [38]. All *PPM1D* mutations in our patients were identified in postchemotherapy treatment blood samples suggesting their somatic origin [39]. Germline mutations in *NAT1* have not been analyzed for OC predisposition so far. However, several polymorphisms in *NAT1* (coding for arylamine N-acetyltransferase 1 engaged in carcinogen metabolism and detoxification) were shown to modify the risk of various cancers [40,41]. The *SHPRH* gene codes for E3 ubiquitin-protein ligase targeting PCNA upon DNA damage [42]. Contribution of *SHPRH* germline variants to OC risk remains elusive. Overall, low mutation frequencies found in gene candidates in our study precluded its precise OC risk estimations and will require large, multiethnic, case-control studies, segregation analyses in affected families, and functional analyses. Alongside variants clustering to a few candidate genes, we identified rare mutations in a gene family coding for Fanconi anemia (FA) proteins involved in the repair of DNA interstrand crosslinks [43]. Several FA genes belong to established OC predisposition genes, including *BRCA1* (*FANCS*), *BRCA2 (FANCD1*), *RAD51C (FANCO)*, *PALB2 (FANCN)*, and *BRIP1 (FANCJ)*. Except these, we found rare mutations in other FA genes (*FANCG*, *FANCD2*, and *FANCA*) in 11 (0.83%) of 1333 OC patients compared to 5 in 2278 PMC (0.2%), with cumulative OR = 3.78 (95% CI 1.21–13.91; $p = 0.02$). Interestingly, these rare mutations were detected almost exclusively in patients without mutations in other OC predisposition genes.

The strengths of this study include an identical NGS analysis and bioinformatics pipeline in all patients, a careful curation of clinical data, and an ethnically homogeneous set of patients and controls representing the largest sample set from the region of Central and Eastern Europe. Despite

that, the number of individuals still did not allow the precise OC risk calculations in rarely mutated genes. Although all OC cases in the Czech Republic are eligible for genetic testing, OC patients with positive family cancer history and earlier-onset individuals were enriched in our study, especially in a small subgroup enrolled before 2015 (in the Center A only).

Whether the high prevalence of clinically important germline mutations in OC patients justifies population-wide screening is a vivid matter of debate [44–48]. We emphasize that we found *BRCA1/2* mutations in 14.5% of OC patients with no family cancer history who would currently not be revealed presymptomatically without population screening. We assume that careful application of germline testing in all OC patients and their relatives would reduce OC burden in our population. Moreover, the mutations in *BRCA1/2* [49,50] and other OC predisposition genes [51,52] represent valuable predictive biomarkers improving OC chemotherapy.

4. Materials and Methods

Analyzed patients (*N* = 1333) were enrolled in 2010–2018 and included all OC cases regardless of familial cancer history or OC histology subtypes. As knowledge about germline mutations' frequency in women diagnosed with BTO is limited, we included these histological subtypes to our study. Clinicopathological data were obtained during genetic counselling or retrieved from the patients' records. OC patients with a positive cancer family history were stratified into (i) hereditary ovarian cancer (HOC) families with OC and other nonbreast cancer (BC) in the family history; (ii) hereditary breast and/or ovarian cancer (HBOC) families with BC and OC or other cancer in the family history, and (iii) multiple cancer families with non-OC and non-BC in the family history. Index patients were tested in seven centers: (A) First Faculty of Medicine, Charles University, Prague (*N* = 637); (B) Masaryk Memorial Cancer Institute, Brno (*N* = 357); (C) Gennet, Prague (*N* = 273); (D) AGEL Laboratories, Novy Jicin (*N* = 34); (E) GHC Genetics (*N* = 12); (F) Pronatal (*N* = 11), and (G) University Hospital Olomouc (*N* = 9).

Population-matched controls (PMC; *N* = 2278) included 616 noncancer controls collected in centers A (*N* = 344), B (*N* = 150), and D (*N* = 122), and 1662 unselected controls provided by the National Center for Medical Genomics (http://ncmg.cz). The noncancer controls were volunteers (78 males and 538 females) aged ≥ 60 years without a personal or family cancer history (in first-degree relatives). The unselected controls (1170 males and 492 females; median age 57 years, range 18–88 years) were unrelated individuals analyzed by whole-exome sequencing (WES) for various noncancer conditions.

All patients and controls were Caucasians of a Czech origin. Written informed consent was obtained from all patients and controls. The study was approved by the Ethics Committee of the General University Hospital in Prague; ethics approval number was 92/14. The study was performed in accordance with the Declaration of Helsinki.

4.1. Next-Generation Sequencing

Germline blood-derived DNA was analyzed by the CZECANCA (CZEch CAncer paNel for Clinical Application; custom-made SeqCap EZ choice panel; Roche) panel NGS targeting 219 genes on MiSeq (Illumina), as described in details previously [19]. Sequencing reads were aligned by Novoalign v2.08.03 to the human reference genome (hg19). Variants were identified using GATK and Pindel, CNVs using CNV score [19]. The entire diagnostic pipeline was successfully tested using European Molecular Genetics Quality Network schemes (EMQN) and validated as we have described recently [19].

4.2. Variant Classification

We first analyzed 18 genes considered clinically relevant to the HBOC syndrome (MIM #604370) by NCCN, namely, *ATM*, *BARD1*, *BRCA1*, *BRCA2*, *BRIP1*, *CDH1*, *CHEK2*, *MLH1*, *MSH2*, *MSH6*, *NBN*, *PALB2*, *PTEN*, *RAD50*, *RAD51C*, *RAD51D*, *STK11*, and *TP53.* Germline variants (with frequency ≤ 0.01 and ≤0.05 in 1000 Genomes project and noncancer PMC, respectively) were classified into three

groups: i) pathogenic/likely pathogenic, ii) variants of unknown significance (VUS), and iii) likely benign/benign, based on recommendations from the ENIGMA consortium (https://enigmaconsortium.org). All nonsense/frameshift/splicing (± 1–2 bp) mutations/CNVs were considered pathogenic/likely pathogenic unless classified as other in the ClinVar database; whole gene duplications were considered VUS. The other types of mutations were considered pathogenic/likely pathogenic only if classified as such in ClinVar by at least two submitters. TP53 variants were classified using the IARC TP53 database (http://p53.iarc.fr/), *CHEK2* VUS using a recently published functional assay [27].

Subsequently, we analyzed variants in another 201 genes targeted by the CZECANCA panel. Nonsense/frameshift/splicing (± 1–2 bp) mutations/CNVs (except whole gene duplications) with frequency ≤0.01 and ≤0.05 in 1000 Genomes project and in noncancer PMC, respectively, were considered pathogenic.

All pathogenic/likely pathogenic mutations in patients and noncancer PMC were confirmed by Sanger sequencing and CNVs by MLPA (if available) or by qPCR (protocol available on request), and they were submitted to ClinVar under the submission ID SUB5822876.

4.3. Statistical Analysis

The odds ratio (OR) for particular gene was calculated using Fisher's exact test, and p values <0.05 were considered significant. The multiple mutation carriers were excluded from the OR calculations. For the identification of other OC candidate genes, the Bonferroni correction was employed. The associations between mutation status and clinicopathological characteristics were estimated using Fisher's exact test, and p values <0.05 were considered significant.

5. Conclusions

Our study demonstrated that nearly one in three OC patients carries a pathogenic mutation in genes significantly associated with OC. The mutation frequency exceeded 10% in all clinicopathological subgroups, regardless of the age at diagnosis, clinical or histopathological characteristics, with an exception of women diagnosed with OC before the age of 30 or with rare histological OC subtypes. Importantly, we found that the high mutation prevalence included borderline tumors justifying genetic testing of all OC patients, including women diagnosed with borderline tumors. Surprisingly, *BRCA1/2* mutations were not associated with sporadic OC in very young women (≤30 years). Besides the established OC predisposition genes, *NBN* and *BARD1* were significantly associated with a moderate OC risk; however, further studies will be required to specify the associated OC risk and to identify the value of the detected genetic mutations in terms of disease prognosis and therapy prediction. Hence, analyses of rarely mutated BC/OC predisposition genes that failed to increase OC risk in our study are further warranted to evaluate their association with OC in future larger dataset and/or in frame of international consortia. These should include also other candidate alterations with predictive and/or prognostic potential.

Supplementary Materials: The following are available online at http://www.mdpi.com/2072-6694/12/4/956/s1, Table S1: the clinicopathological characteristics of 1333 ovarian cancer patients, Table S2: mutation frequencies in ovarian cancer cases and population-matched controls, Table S3: clinical and pathological characteristics in multiple mutation carriers, Table S4: clinicopathological characteristics of mutation carriers in HBOC genes listed in NCCN guidelines (mutation carriers in a group of 1320 ovarian cancer patients (after exclusion of 13 multiple mutation carriers shown in Table S3) (10 genes significantly associated with OC risk in our study are highlighted), Table S5: mutations in 201 additional analyzed genes with associated OC risk (significantly associated genes ($p < 0.05$) are highlighted), Table S6: characteristics of 15 *PPM1D* mutation carriers, and Table S7: clinicopathological characteristics of mutation carriers in *NAT1*, *PPM1D*, and *SHPRH* significantly associated with OC risk.

Author Contributions: Conceptualization, J.S.; data curation, M.V., L.F., and M.K.; funding acquisition, P.K., Z.K., and J.S.; investigation, K.L., L.S., M.J., M.B., M.C., S.J., J.K., Z.V., M.U., P.K., E.M., L.F., J.H., P.V., F.L., M.K., L.C., S.T., J.I., L.H., M.K., R.V., S.K., M.Z., L.M., Z.K., and J.S.; methodology, K.L., Z.K., and J.S.; project administration, J.S.; resources, P.K., R.V., S.K., Z.K., and J.S.; software, P.Z. and V.S.; validation, K.L., L.S., M.J., M.B., M.C., E.M., F.L., S.T., and J.S.; writing—original draft, K.L., Z.K., and J.S.; writing—review and editing, L.S., P.Z., M.V., M.J., M.B., P.K., E.M., L.F., F.L., and L.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Ministry of Health of the Czech Republic [Grant numbers 16-29959A, NV17-32030A, NV18-03-00024, MH CZ – DRO (FNOl, 00098892)] and by the Charles University projects [SVV 260516, PROGRES Q28/LF1, CZ.02.1.01/0.0/0.0/16_013/0001634].

Acknowledgments: We thank our patients for contribution in this study. We thank our clinical colleagues for their valuable comments to the manuscript and Jan Flemr for language editing.

Conflicts of Interest: The authors have declared no conflicts of interest.

References

1. Levanon, K.; Crum, C.; Drapkin, R. New insights into the pathogenesis of serous ovarian cancer and its clinical impact. *J. Clin. Oncol.* **2008**, *26*, 5284–5293. [CrossRef] [PubMed]
2. Kim, J.; Park, E.Y.; Kim, O.; Schilder, J.M.; Coffey, D.M.; Cho, C.H.; Bast, R.C., Jr. Cell Origins of High-Grade Serous Ovarian Cancer. *Cancers (Basel)* **2018**, *10*, 433. [CrossRef] [PubMed]
3. Siegel, R.L.; Miller, K.D.; Jemal, A. Cancer statistics, 2016. *CA A Cancer J. Clin.* **2016**, *66*, 7–30. [CrossRef] [PubMed]
4. Walsh, T.; Casadei, S.; Lee, M.K.; Pennil, C.C.; Nord, A.S.; Thornton, A.M.; Roeb, W.; Agnew, K.J.; Stray, S.M.; Wickramanayake, A.; et al. Mutations in 12 genes for inherited ovarian, fallopian tube, and peritoneal carcinoma identified by massively parallel sequencing. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 18032–18037. [CrossRef] [PubMed]
5. Koczkowska, M.; Krawczynska, N.; Stukan, M.; Kuzniacka, A.; Brozek, I.; Sniadecki, M.; Debniak, J.; Wydra, D.; Biernat, W.; Kozlowski, P.; et al. Spectrum and Prevalence of Pathogenic Variants in Ovarian Cancer Susceptibility Genes in a Group of 333 Patients. *Cancers (Basel)* **2018**, *10*, 442. [CrossRef] [PubMed]
6. Krivokuca, A.; Boljevic, I.; Jovandic, S.; Magic, Z.; Mandic, A.; Tomasevic, Z.; Brankovic-Magic, M. Germline mutations in cancer susceptibility genes in high grade serous ovarian cancer in Serbia. *J. Hum. Genet.* **2019**. [CrossRef]
7. Offit, K. BRCA mutation frequency and penetrance: New data, old debate. *J. Natl. Cancer Inst.* **2006**, *98*, 1675–1677. [CrossRef]
8. Norquist, B.M.; Harrell, M.I.; Brady, M.F.; Walsh, T.; Lee, M.K.; Gulsuner, S.; Bernards, S.S.; Casadei, S.; Yi, Q.; Burger, R.A.; et al. Inherited Mutations in Women With Ovarian Carcinoma. *JAMA Oncol.* **2016**, *2*, 482–490. [CrossRef]
9. Loveday, C.; Turnbull, C.; Ruark, E.; Xicola, R.M.; Ramsay, E.; Hughes, D.; Warren-Perry, M.; Snape, K.; Breast Cancer Susceptibility, C.; Eccles, D.; et al. Germline RAD51C mutations confer susceptibility to ovarian cancer. *Nat. Genet.* **2012**, *44*, 475–476, author reply 476. [CrossRef]
10. Loveday, C.; Turnbull, C.; Ramsay, E.; Hughes, D.; Ruark, E.; Frankum, J.R.; Bowden, G.; Kalmyrzaev, B.; Warren-Perry, M.; Snape, K.; et al. Germline mutations in RAD51D confer susceptibility to ovarian cancer. *Nat. Genet.* **2011**, *43*, 879–882. [CrossRef]
11. Watson, P.; Butzow, R.; Lynch, H.T.; Mecklin, J.P.; Jarvinen, H.J.; Vasen, H.F.; Madlensky, L.; Fidalgo, P.; Bernstein, I.; International Collaborative Group on, H. The clinical features of ovarian cancer in hereditary nonpolyposis colorectal cancer. *Gynecol. Oncol.* **2001**, *82*, 223–228. [CrossRef]
12. Ramus, S.J.; Song, H.; Dicks, E.; Tyrer, J.P.; Rosenthal, A.N.; Intermaggio, M.P.; Fraser, L.; Gentry-Maharaj, A.; Hayward, J.; Philpott, S.; et al. Germline Mutations in the BRIP1, BARD1, PALB2, and NBN Genes in Women With Ovarian Cancer. *J. Natl. Cancer Inst.* **2015**, *107*. [CrossRef] [PubMed]
13. Banno, K.; Kisu, I.; Yanokura, M.; Masuda, K.; Ueki, A.; Kobayashi, Y.; Hirasawa, A.; Aoki, D. Hereditary gynecological tumors associated with Peutz-Jeghers syndrome (Review). *Oncol. Lett.* **2013**, *6*, 1184–1188. [CrossRef] [PubMed]
14. Lilyquist, J.; LaDuca, H.; Polley, E.; Davis, B.T.; Shimelis, H.; Hu, C.; Hart, S.N.; Dolinsky, J.S.; Couch, F.J.; Goldgar, D.E. Frequency of mutations in a large series of clinically ascertained ovarian cancer cases tested on multi-gene panels compared to reference controls. *Gynecol. Oncol.* **2017**, *147*, 375–380. [CrossRef] [PubMed]
15. Carter, N.J.; Marshall, M.L.; Susswein, L.R.; Zorn, K.K.; Hiraki, S.; Arvai, K.J.; Torene, R.I.; McGill, A.K.; Yackowski, L.; Murphy, P.D.; et al. Germline pathogenic variants identified in women with ovarian tumors. *Gynecol. Oncol.* **2018**, *151*, 481–488. [CrossRef]

16. Schubert, S.; van Luttikhuizen, J.L.; Auber, B.; Schmidt, G.; Hofmann, W.; Penkert, J.; Davenport, C.F.; Hille-Betz, U.; Wendeburg, L.; Bublitz, J.; et al. The identification of pathogenic variants in BRCA1/2 negative, high risk, hereditary breast and/or ovarian cancer patients: High frequency of FANCM pathogenic variants. *Int. J. Cancer* **2019**, *144*, 2683–2694. [CrossRef]
17. Daly, M.B.; Pilarski, R.; Berry, M.; Buys, S.S.; Farmer, M.; Friedman, S.; Garber, J.E.; Kauff, N.D.; Khan, S.; Klein, C.; et al. NCCN Guidelines Insights: Genetic/Familial High-Risk Assessment: Breast and Ovarian, Version 2.2017. *J. Natl. Compr. Cancer Netw. JNCCN* **2017**, *15*, 9–20. [CrossRef]
18. Domchek, S.M.; Friebel, T.M.; Neuhausen, S.L.; Wagner, T.; Evans, G.; Isaacs, C.; Garber, J.E.; Daly, M.B.; Eeles, R.; Matloff, E.; et al. Mortality after bilateral salpingo-oophorectomy in BRCA1 and BRCA2 mutation carriers: A prospective cohort study. *Lancet. Oncol.* **2006**, *7*, 223–229. [CrossRef]
19. Soukupova, J.; Zemankova, P.; Lhotova, K.; Janatova, M.; Borecka, M.; Stolarova, L.; Lhota, F.; Foretova, L.; Machackova, E.; Stranecky, V.; et al. Validation of CZECANCA (CZEch CAncer paNel for Clinical Application) for targeted NGS-based analysis of hereditary cancer syndromes. *PLoS ONE* **2018**, *13*, e0195761. [CrossRef]
20. King, M.C.; Marks, J.H.; Mandell, J.B.; New York Breast Cancer Study, G. Breast and ovarian cancer risks due to inherited mutations in BRCA1 and BRCA2. *Science* **2003**, *302*, 643–646. [CrossRef]
21. Rafnar, T.; Gudbjartsson, D.F.; Sulem, P.; Jonasdottir, A.; Sigurdsson, A.; Jonasdottir, A.; Besenbacher, S.; Lundin, P.; Stacey, S.N.; Gudmundsson, J.; et al. Mutations in BRIP1 confer high risk of ovarian cancer. *Nat. Genet.* **2011**, *43*, 1104–1107. [CrossRef] [PubMed]
22. Bonache, S.; Esteban, I.; Moles-Fernandez, A.; Tenes, A.; Duran-Lozano, L.; Montalban, G.; Bach, V.; Carrasco, E.; Gadea, N.; Lopez-Fernandez, A.; et al. Multigene panel testing beyond BRCA1/2 in breast/ovarian cancer Spanish families and clinical actionability of findings. *J. Cancer Res. Clin. Oncol.* **2018**, *144*, 2495–2513. [CrossRef]
23. Lynch, H.T.; de la Chapelle, A. Hereditary colorectal cancer. *N. Engl. J. Med.* **2003**, *348*, 919–932. [CrossRef] [PubMed]
24. Lu, H.M.; Li, S.; Black, M.H.; Lee, S.; Hoiness, R.; Wu, S.; Mu, W.; Huether, R.; Chen, J.; Sridhar, S.; et al. Association of Breast and Ovarian Cancers With Predisposition Genes Identified by Large-Scale Sequencing. *JAMA Oncol.* **2018**. [CrossRef] [PubMed]
25. Kurian, A.W.; Ward, K.C.; Howlader, N.; Deapen, D.; Hamilton, A.S.; Mariotto, A.; Miller, D.; Penberthy, L.S.; Katz, S.J. Genetic Testing and Results in a Population-Based Cohort of Breast Cancer Patients and Ovarian Cancer Patients. *J. Clin. Oncol.* **2019**, JCO1801854. [CrossRef]
26. Yang, X.; Leslie, G.; Doroszuk, A.; Schneider, S.; Allen, J.; Decker, B.; Dunning, A.M.; Redman, J.; Scarth, J.; Plaskocinska, I.; et al. Cancer Risks Associated With Germline PALB2 Pathogenic Variants: An International Study of 524 Families. *J. Clin. Oncol.* **2020**, *38*, 674–685. [CrossRef] [PubMed]
27. Kleiblova, P.; Stolarova, L.; Krizova, K.; Lhota, F.; Hojny, J.; Zemankova, P.; Havranek, O.; Vocka, M.; Cerna, M.; Lhotova, K.; et al. Identification of deleterious germline CHEK2 mutations and their association with breast and ovarian cancer. *Int. J. Cancer* **2019**. [CrossRef]
28. Harter, P.; Hauke, J.; Heitz, F.; Reuss, A.; Kommoss, S.; Marme, F.; Heimbach, A.; Prieske, K.; Richters, L.; Burges, A.; et al. Prevalence of deleterious germline variants in risk genes including BRCA1/2 in consecutive ovarian cancer patients (AGO-TR-1). *PLoS ONE* **2017**, *12*, e0186043. [CrossRef] [PubMed]
29. Whitworth, J.; Skytte, A.B.; Sunde, L.; Lim, D.H.; Arends, M.J.; Happerfield, L.; Frayling, I.M.; van Minkelen, R.; Woodward, E.R.; Tischkowitz, M.D.; et al. Multilocus Inherited Neoplasia Alleles Syndrome: A Case Series and Review. *JAMA Oncol.* **2016**, *2*, 373–379. [CrossRef]
30. Stradella, A.; Del Valle, J.; Rofes, P.; Feliubadalo, L.; Grau Garces, E.; Velasco, A.; Gonzalez, S.; Vargas, G.; Izquierdo, A.; Campos, O.; et al. Does multilocus inherited neoplasia alleles syndrome have severe clinical expression? *J. Med. Genet.* **2018**. [CrossRef]
31. Whitworth, J.; Smith, P.S.; Martin, J.E.; West, H.; Luchetti, A.; Rodger, F.; Clark, G.; Carss, K.; Stephens, J.; Stirrups, K.; et al. Comprehensive Cancer-Predisposition Gene Testing in an Adult Multiple Primary Tumor Series Shows a Broad Range of Deleterious Variants and Atypical Tumor Phenotypes. *Am. J. Hum. Genet.* **2018**, *103*, 3–18. [CrossRef]
32. Castera, L.; Harter, V.; Muller, E.; Krieger, S.; Goardon, N.; Ricou, A.; Rousselin, A.; Paimparay, G.; Legros, A.; Bruet, O.; et al. Landscape of pathogenic variations in a panel of 34 genes and cancer risk estimation from 5131 HBOC families. *Genet. Med.* **2018**, *20*, 1677–1686. [CrossRef] [PubMed]

33. Morgan, R.D.; Burghel, G.J.; Flaum, N.; Bulman, M.; Clamp, A.R.; Hasan, J.; Mitchell, C.L.; Schlecht, H.; Woodward, E.R.; Lallo, F.I.; et al. Prevalence of germline pathogenic BRCA1/2 variants in sequential epithelial ovarian cancer cases. *J. Med. Genet.* **2019**, *56*, 301–307. [CrossRef] [PubMed]
34. Plaskocinska, I.; Shipman, H.; Drummond, J.; Thompson, E.; Buchanan, V.; Newcombe, B.; Hodgkin, C.; Barter, E.; Ridley, P.; Ng, R.; et al. New paradigms for BRCA1/BRCA2 testing in women with ovarian cancer: Results of the Genetic Testing in Epithelial Ovarian Cancer (GTEOC) study. *J. Med. Genet.* **2016**, *53*, 655–661. [CrossRef]
35. Rust, K.; Spiliopoulou, P.; Tang, C.Y.; Bell, C.; Stirling, D.; Phang, T.; Davidson, R.; Mackean, M.; Nussey, F.; Glasspool, R.M.; et al. Routine germline BRCA1 and BRCA2 testing in patients with ovarian carcinoma: Analysis of the Scottish real-life experience. *BJOG* **2018**, *125*, 1451–1458. [CrossRef] [PubMed]
36. Ruark, E.; Snape, K.; Humburg, P.; Loveday, C.; Bajrami, I.; Brough, R.; Rodrigues, D.N.; Renwick, A.; Seal, S.; Ramsay, E.; et al. Mosaic PPM1D mutations are associated with predisposition to breast and ovarian cancer. *Nature* **2013**, *493*, 406–410. [CrossRef] [PubMed]
37. Akbari, M.R.; Lepage, P.; Rosen, B.; McLaughlin, J.; Risch, H.; Minden, M.; Narod, S.A. PPM1D mutations in circulating white blood cells and the risk for ovarian cancer. *J. Natl. Cancer Inst.* **2014**, *106*, djt323. [CrossRef]
38. Kleiblova, P.; Shaltiel, I.A.; Benada, J.; Sevcik, J.; Pechackova, S.; Pohlreich, P.; Voest, E.E.; Dundr, P.; Bartek, J.; Kleibl, Z.; et al. Gain-of-function mutations of PPM1D/Wip1 impair the p53-dependent G1 checkpoint. *J. Cell Biol.* **2013**, *201*, 511–521. [CrossRef]
39. Pharoah, P.D.P.; Song, H.; Dicks, E.; Intermaggio, M.P.; Harrington, P.; Baynes, C.; Alsop, K.; Australian Ovarian Cancer Study, G.; Bogdanova, N.; Cicek, M.S.; et al. PPM1D Mosaic Truncating Variants in Ovarian Cancer Cases May Be Treatment-Related Somatic Mutations. *J. Natl. Cancer Inst.* **2016**, *108*. [CrossRef]
40. Hein, D.W.; Fakis, G.; Boukouvala, S. Functional expression of human arylamine N-acetyltransferase NAT1*10 and NAT1*11 alleles: A mini review. *Pharm. Genom.* **2018**, *28*, 238–244. [CrossRef]
41. Butcher, N.J.; Minchin, R.F. Arylamine N-acetyltransferase 1: A novel drug target in cancer development. *Pharm. Rev.* **2012**, *64*, 147–165. [CrossRef] [PubMed]
42. Seelinger, M.; Otterlei, M. Helicase-Like Transcription Factor HLTF and E3 Ubiquitin Ligase SHPRH Confer DNA Damage Tolerance through Direct Interactions with Proliferating Cell Nuclear Antigen (PCNA). *Int. J. Mol. Sci.* **2020**, *21*, 693. [CrossRef] [PubMed]
43. Nalepa, G.; Clapp, D.W. Fanconi anaemia and cancer: An intricate relationship. *Nat. Rev. Cancer* **2018**, *18*, 168–185. [CrossRef] [PubMed]
44. Zhang, L.; Bao, Y.; Riaz, M.; Tiller, J.; Liew, D.; Zhuang, X.; Amor, D.J.; Huq, A.; Petelin, L.; Nelson, M.; et al. Population genomic screening of all young adults in a health-care system: A cost-effectiveness analysis. *Genet. Med.* **2019**. [CrossRef]
45. Best, A.F.; Tucker, M.A.; Frone, M.N.; Greene, M.H.; Peters, J.A.; Katki, H.A. A Pragmatic Testing-Eligibility Framework for Population Mutation Screening: The Example of BRCA1/2. *Cancer Epidemiol Biomark. Prev.* **2019**, *28*, 293–302. [CrossRef] [PubMed]
46. Turnbull, C.; Sud, A.; Houlston, R.S. Cancer genetics, precision prevention and a call to action. *Nat. Genet.* **2018**, *50*, 1212–1218. [CrossRef]
47. Gabai-Kapara, E.; Lahad, A.; Kaufman, B.; Friedman, E.; Segev, S.; Renbaum, P.; Beeri, R.; Gal, M.; Grinshpun-Cohen, J.; Djemal, K.; et al. Population-based screening for breast and ovarian cancer risk due to BRCA1 and BRCA2. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 14205–14210. [CrossRef]
48. Manchanda, R.; Patel, S.; Gordeev, V.S.; Antoniou, A.C.; Smith, S.; Lee, A.; Hopper, J.L.; MacInnis, R.J.; Turnbull, C.; Ramus, S.J.; et al. Cost-effectiveness of Population-Based BRCA1, BRCA2, RAD51C, RAD51D, BRIP1, PALB2 Mutation Testing in Unselected General Population Women. *J. Natl. Cancer Inst.* **2018**, *110*, 714–725. [CrossRef]
49. George, A.; Kaye, S.; Banerjee, S. Delivering widespread BRCA testing and PARP inhibition to patients with ovarian cancer. *Nat. Rev. Clin. Oncol.* **2017**, *14*, 284–296. [CrossRef]
50. Ledermann, J.; Harter, P.; Gourley, C.; Friedlander, M.; Vergote, I.; Rustin, G.; Scott, C.L.; Meier, W.; Shapira-Frommer, R.; Safra, T.; et al. Olaparib maintenance therapy in patients with platinum-sensitive relapsed serous ovarian cancer: A preplanned retrospective analysis of outcomes by BRCA status in a randomised phase 2 trial. *Lancet. Oncol.* **2014**, *15*, 852–861, Correction in **2015**, *16*, e158. [CrossRef]

51. Chandran, E.A.; Kennedy, I. Significant Tumor Response to the Poly (ADP-ribose) Polymerase Inhibitor Olaparib in Heavily Pretreated Patient With Ovarian Carcinosarcoma Harboring a Germline RAD51D Mutation. *JCO Precis. Oncol.* **2018**, 1–4. [CrossRef]
52. Ngoi, N.Y.L.; Tay, D.; Heong, V.; Thian, Y.L.; Ong, P.Y.; Ow, S.G.W.; Jeyasekharan, A.D.; Lim, Y.W.; Lim, S.E.; Lee, S.C.; et al. Reversal of Bowel Obstruction with Platinum-Based Chemotherapy and Olaparib in Recurrent, Short Platinum-Free Interval, RAD51C Germline Mutation–Associated Ovarian Cancer. *JCO Precis. Oncol.* **2018**, 1–8. [CrossRef]

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

Příloha 2 – Kleiblova, P., L. Stolarova, K. Krizova, F. Lhota, J. Hojny, P. Zemankova, O. Havranek, M. Vocka, M. Cerna, K. Lhotova, M. Borecka, M. Janatova, J. Soukupova, J. Sevcik, M. Zimovjanova, J. Kotlas, A. Panczak, K. Vesela, J. Cervenkova, M. Schneiderova, M. Burocziova, K. Burdova, V. Stranecky, L. Foretova, E. Machackova, S. Tavandzis, S. Kmoch, L. Macurek, and Z. Kleibl. 2019b. "***Identification of deleterious germline CHEK2 mutations and their association with breast and ovarian cancer***." *Int J Cancer*. doi: 10.1002/ijc.32385.

IJC
International Journal of Cancer

Identification of deleterious germline *CHEK2* mutations and their association with breast and ovarian cancer

Petra Kleiblova[1,2*], Lenka Stolarova[1*], Katerina Krizova[3], Filip Lhota[1], Jan Hojny[1], Petra Zemankova[1], Ondrej Havranek[4,5], Michal Vocka[6], Marta Cerna[1], Klara Lhotova[1], Marianna Borecka[1], Marketa Janatova[1], Jana Soukupova[1], Jan Sevcik[1], Martina Zimovjanova[6], Jaroslav Kotlas[2], Ales Panczak[2], Kamila Vesela[2], Jana Cervenkova[7], Michaela Schneiderova[8], Monika Burocziova[3], Kamila Burdova[3], Viktor Stranecky[9], Lenka Foretova[10], Eva Machackova[10], Spiros Tavandzis[11], Stanislav Kmoch[9], Libor Macurek[3] and Zdenek Kleibl[1]

[1]Institute of Biochemistry and Experimental Oncology, First Faculty of Medicine, Charles University, Prague, Czech Republic
[2]Institute of Biology and Medical Genetics, First Faculty of Medicine, Charles University and General University Hospital, Prague, Czech Republic
[3]Laboratory of Cancer Cell Biology, Institute of Molecular Genetics of the ASCR, Prague, Czech Republic
[4]BIOCEV, First Faculty of Medicine, Charles University, Prague, Czech Republic
[5]Department of Hematology, First Faculty of Medicine, Charles University and General University Hospital, Prague, Czech Republic
[6]Department of Oncology, First Faculty of Medicine, Charles University and General University Hospital, Prague, Czech Republic
[7]Department of Radiology, First Faculty of Medicine, Charles University and General University Hospital, Prague, Czech Republic
[8]First Department of Surgery, First Faculty of Medicine, Charles University and General University Hospital, Prague, Czech Republic
[9]Research Unit for Rare Diseases, Department of Pediatrics and Adolescent Medicine, First Faculty of Medicine, Charles University and General University Hospital, Prague, Czech Republic
[10]Department of Cancer Epidemiology and Genetics, Masaryk Memorial Cancer Institute, Brno, Czech Republic
[11]Department of Medical Genetics, AGEL Laboratories, AGEL Research and Training Institute, Novy Jicin, Czech Republic

Germline mutations in checkpoint kinase 2 (*CHEK2*), a multiple cancer-predisposing gene, increase breast cancer (BC) risk; however, risk estimates differ substantially in published studies. We analyzed germline *CHEK2* variants in 1,928 high-risk Czech breast/ovarian cancer (BC/OC) patients and 3,360 population-matched controls (PMCs). For a functional classification of VUS, we developed a complementation assay in human nontransformed RPE1-*CHEK2*-knockout cells quantifying CHK2-specific phosphorylation of endogenous protein KAP1. We identified 10 truncations in 46 (2.39%) patients and in 11 (0.33%) PMC ($p = 1.1 \times 10^{-14}$). Two types of large intragenic rearrangements (LGR) were found in 20/46 mutation carriers. Truncations significantly increased unilateral BC risk (OR = 7.94; 95%CI 3.90–17.47; $p = 1.1 \times 10^{-14}$) and were more frequent in patients with bilateral BC (4/149; 2.68%; $p = 0.003$), double primary BC/OC (3/79; 3.80%; $p = 0.004$), male BC (3/48; 6.25%; $p = 8.6 \times 10^{-4}$), but not with OC (3/354; 0.85%; $p = 0.14$). Additionally, we found 26 missense VUS in 88 (4.56%) patients and 131 (3.90%) PMC ($p = 0.22$). Using our functional assay, 11 variants identified in 15 (0.78%) patients and 6 (0.18%) PMC were scored deleterious ($p = 0.002$). Frequencies of functionally intermediate and neutral variants did not differ between patients

Key words: breast cancer, ovarian cancer, germline mutations, CHEK2, VUS, KAP1, functional assay

Abbreviations: BC: breast cancer; CHEK2: checkpoint kinase 2; CI: confidence interval; CNV: copy number variant(s); CPG: cancer-predisposing genes; CZECANCA: CZEch CAncer paNel for Clinical Application; DHPLC: denaturing high performance liquid chromatography; EGFP: enhanced green fluorescent protein; ExAC: exome aggregation consortium; HRMA: high-resolution melting analysis; IGV: integrative genomics viewer; KAP1: KRAB-associated protein 1; KO: knockout; LGR: large genomic rearrangement(s); MLPA: multiplex ligation-dependent probe amplification; NFE: non-Finnish European; NGS: next-generation sequencing; OC: ovarian cancer; OR: odds ratio; PMC: population-matched controls; RPE1: retinal pigmented epithelial cells; SNV: single nucleotide variant

Additional Supporting Information may be found in the online version of this article.

Conflict of interest: The authors declare no potential conflicts of interest.

Grant sponsor: Akademie Věd České Republiky; **Grant sponsor:** Strategie AV21, Qualitas; **Grant sponsor:** Grantová Agentura, Univerzita Karlova; **Grant numbers:** GAUK 762216, PRIMUS/17/MED/9, PROGRES Q26/LF1, PROGRES Q28/LF1, SVV2019/260367, UNCE/MED/016; **Grant sponsor:** Ministerstvo Školství, Mládeže a Tělovýchovy; **Grant numbers:** NPU II LQ1604, LM2015091; **Grant sponsor:** Ministerstvo Zdravotnictví Ceské Republiky; **Grant numbers:** NV15-27695A, NV15-28830A, NV16-29959A, NV19-03-00279

*P.K. and L.S. authors contributed equally to this work

DOI: 10.1002/ijc.32385

History: Received 6 Feb 2019; Accepted 24 Apr 2019; Online 3 May 2019

Correspondence to: Zdenek Kleibl, Institute of Biochemistry and Experimental Oncology, First Faculty of Medicine, Charles University, U Nemocnice 5, 120 00 Prague 2, Czech Republic, E-mail: zdekleje@lf1.cuni.cz; or Libor Macurek, Institute of Molecular Genetics of the ASCR, Videnska 1083, 142 20 Prague 4, Czech Republic, E-mail: libor.macurek@img.cas.cz

and PMC. Functionally deleterious *CHEK2* missense variants significantly increased BC risk (OR = 3.90; 95%CI 1.24–13.35; $p = 0.009$) and marginally OC risk (OR = 4.77; 95%CI 0.77–22.47; $p = 0.047$); however, carriers low frequency will require evaluation in larger studies. Our study highlights importance of LGR detection for *CHEK2* analysis, careful consideration of ethnicity in both cases and controls for risk estimates, and demonstrates promising potential of newly developed human nontransformed cell line assay for functional *CHEK2* VUS classification.

What's new?
The tumor suppressor gene checkpoint kinase 2 (*CHEK2*) encodes a protein that serves an important role in DNA repair. However, *CHEK2* is also vulnerable to mutations that potentially impact breast cancer risk. Using a functional cell-based assay, the authors of the present study show that truncating and missense *CHEK2* variants are associated with risk of both breast and ovarian cancer. One-third of truncating mutations involved large genomic rearrangements. In addition, *CHEK2* mutations predisposed women to specific breast cancer types, and *CHEK2* mutation carriers with a family history of cancer were at increased risk of developing second primary cancers.

Introduction

Approximately 10% of breast cancer (BC) and 20% of ovarian cancer (OC) cases arise as a hereditary disease in patients carrying a pathogenic mutation in BC/OC-predisposing genes.[1,2] The clinical utility of pathogenic mutations in major BC/OC genes (*BRCA1* and *BRCA2*) is well established but it remains less certain for a growing group of cancer-predisposing genes (CPG) whose germline mutations confer a moderate cancer risk (*ATM*, *CHEK2*, *PALB2*).[3] This problem is becoming even more critical with the introduction of multigene panel next-generation sequencing (NGS) into the routine genetic analysis of high-risk BC/OC individuals.[4]

Germline *CHEK2* mutations have been linked with susceptibility to several malignancies including BC.[5] The *CHEK2* gene codes for serine/threonine CHK2 kinase involved in DNA damage response (DDR). Activated by a DNA lesion, ATM kinase catalyzes CHK2 T68 phosphorylation promoting CHK2 homodimerization through its forkhead-associated domains and kinase domain autophosphorylation.[6,7] Activated CHK2 phosphorylates multiple proteins involved in DNA repair and DDR, including BRCA1/BRCA2 and p53.[8,9] Another CHK2 substrate is KRAB-associated protein 1 (KAP1, alias TIF1β, TRIM28) a universal corepressor required for transcriptional repression mediated by the KRAB protein superfamily. CHK2-mediated KAP1 S473 phosphorylation reduces its transcription repression resulting in wide effects on gene expression.[10] Although the role of the ATM–CHK2–p53 pathway in the DNA damage-induced cell cycle checkpoint is redundant, CHK2 participates in p53-dependent cell death.[11–14]

The association of germline *CHEK2* variants with BC was assessed early in studies genotyping European founder mutations including the truncating mutation c.1100delC and the missense variant c.470T>C (p.I157T).[5] Subsequent meta-analyses demonstrated that while c.1100delC represents a moderate-risk variant for unselected (OR = 2.7; 95% confidence interval [CI] 2.1–3.4), early onset (OR = 2.6; 95%CI 1.3–5.5) and familial BC (OR = 4.8; 95% CI 3.3–7.2),[15] p.I157T is a low-risk variant with OR <1.5 for all BC subgroups.[16] Other founder variants include the spliceogenic mutation c.444+1G>A (IVS2+1G>A) and a large genomic rearrangement (LGR) with exon 9–10 deletion (c.909-2028_1095+330del5395) identified in Slavic populations,[17] and the Ashkenazi Jewish founder missense mutation c.1283C>T (p.S428F).[18]

Only few early studies analyzed the entire *CHEK2* coding sequence and revealed that c.1100delC and p.I157T represent only a fraction of *CHEK2* variants in BC patients.[19–22] Recent panel NGS analyses in large cohorts have shown that the *CHEK2* mutation rate is one of the highest among non-*BRCA1/BRCA2* genes in BC in individuals of Ashkenazi Jewish or European ancestry.[23–26] However, the classification of most missense variants remains uncertain,[27] their assessment is problematic,[4] and nearly one-third of *CHEK2* variants are reported discordantly.[28]

In contrast to BC, the association of *CHEK2* germline variants with OC risk is disputable. While several case–control studies have not significantly associated the c.1100delC mutation with OC development,[29,30] recent panel NGS analyses in 4,439 and 6,001 OC samples from the US identified *CHEK2* as the third most frequently affected susceptibility gene.[31,32]

In our study, we identified germline *CHEK2* variants in 1,928 high-risk BC/OC patients and 3,360 population-matched controls (PMCs). Subsequently, we have developed a cell-based assay utilizing a human RPE1 cell line model with endogenous *CHEK2* knockout to functionally classify the identified variants of unknown significance (VUS). This strategy enabled us to identify deleterious germline *CHEK2* mutations, to evaluate cancer risk in their carriers and to describe the clinical and histopathological characteristics of breast tumors in mutation carriers.

Methods

Detailed information is provided in Supporting Information Methods.

Subjects

The patient group included 1,928 BC/OC patients (herein denoted as *all patients*) referred by clinical geneticists for a CPG-mutation analysis performed at the Laboratory of Oncogenetics, First Faculty of Medicine, Charles University, in 1997–2017. Overall, 424/1,928 patients carried a mutation in other (i.e., non-*CHEK2*) cancer-predisposing gene for BC (*BRCA1*, *BRCA2*, *PALB2*, *TP53*) or OC (*BRCA1*, *BRCA2*, *RAD51C*, *RAD51D*, *MLH1*, *MSH2*, *MSH6*) and were denoted herein as *other CPG-mutated*. Remaining 1,504/1,928 patients were negative for mutations in aforementioned genes (herein denoted as *other CPG-wt*). All participants signed an informed consent approved by the local ethical committee. Clinical and histopathological data (Supporting Information Table S1) were obtained during genetic counseling or retrieved from the patients' records.

The set of 3,360 adult PMCs comprised 720 samples of noncancer individuals, 369 samples of adult blood donors, 609 noncancer controls aged >60 years without cancer in first-degree relatives and 1,662 individuals analyzed by exome sequencing at the National Center for Medical Genomics (http://ncmg.cz). In total, PMC set included 1,593 female (with median age 66 years, range 20–98 years) and 1,767 male (with median age 60 years, range 18–94 years) controls. All patients and controls were Caucasians, of the Czech origin.

Mutation analyses

Until 2015, mutation analyses of the entire *CHEK2* coding sequence in BC patients were performed by a high-resolution melting analysis (HRMA) of all coding exons. LGRs were analyzed by a multiplex ligation-dependent probe amplification (MLPA), as described previously.[33] All OC patients' samples, samples from BC patients enrolled since 2015, and samples from all identified *CHEK2* variant carriers were analyzed by a CZECANCA panel (CZEch CAncer paNel for Clinical Application; custom-made SeqCap EZ choice panel, Roche) targeting 219 genes with MiSeq (Illumina) NGS as described recently.[34] The coverage uniformity enabled to evaluate CNVs at 100× average coverage. *CHEK2* variants identified in patients were also sequenced at the mRNA (cDNA) level to determine a potential impact on splicing. NGS-analysis performed in 2,271/3,360 (67.6%) PMC samples (609 noncancer controls and 1,662 NCMG controls) included SNV/indels and CNV analyses. In remaining 1,089/3,360 (32.4%) PMC samples (720 noncancer individuals and 369 blood donors), entire *CHEK2* coding sequence was analyzed by HRMA, similarly as in patients and mutation-specific PCR/HRMA was used for identification of two *CHEK2* LGRs identified in our population (see Supporting Information Methods for details). The consequences of the identified missense variants were predicted by *in silico* tools: Align-GVGD, MutationTaster, CADD, SIFT, PolyPhen-2, Spidex and GERP.

Cell lines

To generate RPE1-*CHEK2*-KO cells, hTERT-RPE1 cells were transfected with a *CHEK2*-CRISPR/Cas9-KO plasmid (Santa Cruz Biotechnology, Santa Cruz, CA; sc-400,438) and a *CHEK2*-HDR plasmid (1:1) and selected by puromycin (7.5 μg/ml) for 3 weeks. The integration of an HDR cassette into the *CHEK2* locus was confirmed by sequencing and a loss of CHK2 expression by immunoblotting (all used antibodies are described in Supporting Information Methods). To remove the HDR cassette, cells were transfected with Cre vector (Santa Cruz, sc-418,923) and RFP-negative cells were selected by flow cytometry. For stable complementation of CHK2, RPE1-*CHEK2*-KO cells were transfected with a linearized pcDNA4-EGFP-*CHEK2* plasmid, selected with zeocin for 3 weeks and single clones were expanded. Plasmid DNA was transfected using polyethylenimine HCl MAX (MW 40000, Polysciences, Warrington, PA) at a 1:5 ratio and growth media were changed after 3 hr. Silencer Select siRNA oligonucleotides (5 nM, Ambion) were transfected using RNAiMAX (Life Technologies, Carlsbad, CA) according to the manufacturer's instructions.

Plasmids

CHEK2 mutants were generated using QuickChange II Site-Directed Mutagenesis (Agilent Technologies, Santa Clara, CA). Wild-type or mutated *CHEK2* was amplified by PCR and cloned in frame into pcDNA4-EGFP or pGEX-6P-1 plasmids using a Gibson assembly kit (NEB). All mutants were verified by Sanger sequencing. A DNA fragment corresponding to the GVKRSRSGEGEV peptide (containing S473) from human KAP1 was ligated in frame into a pGEX-6P-1 plasmid. Alternatively, a fragment corresponding to T2A-EGFP was ligated into the XbaI site of pcDNA4, and subsequently a fragment corresponding to wild-type or mutant FLAG-*CHEK2* was cloned into *Hind*III/*Xho*I sites resulting in a plasmid for bicistronic expression of FLAG-CHK2 and EGFP.

Immunofluorescence microscopy, cell-based assay for the detection of CHK2 activity

RPE1-*CHEK2*-KO cells transfected with an empty EGFP plasmid, wild-type or mutant EGFP-*CHEK2* were seeded on glass coverslips and fixed by 4% paraformaldehyde 48 hr after transfection. Cells were permeabilized by 0.2% Triton X-100 in PBS for 20 min and blocked with 3% BSA in PBS at room temperature. The coverslips were incubated with the KAP1-pS473 antibody for 1 hr at room temperature, three times washed with PBS and incubated with the goat-antimouse Alexa568 antibody and DAPI. After the PBS washing, the coverslips were mounted using Vectashield H-1000 and imaged using a Scan^R microscope (Olympus, Waltham, MA) equipped with an ORCA-285 camera and a 40×/1.3 NA objective. The total intensity of the KAP1-pS473 signal per nucleus was determined in cells expressing low levels of GFP. Three independent experiments were performed and >300 cells were quantified per condition in each experiment. The KAP1-pS473 signal in cells expressing only EGFP typically reached <10% of the signal in cells expressing wild-type CHK2 and was subtracted as a background. The KAP1-pS473 signal measured in cells expressing mutant CHK2 was normalized to wild-type CHK2-expressing cells. The activities of the analyzed

variants were classified as normal, intermediate or deleterious based on mean pS473 reaching >50%, 25–50% and <25% of wild-type CHK2, respectively.

In vitro kinase assays

Escherichia coli BL21 transformed with wild-type or mutant pGEX-6P-1-*CHEK2* plasmids were induced at A_{600} = 0.6 by 0.2 mM IPTG and grown for 5 hr at 37°C. The bacteria were lysed in ice-cold PBS supplemented with 0.1% TX-100 and 1 mM PMSF and sonicated 2 × 30 sec. Cleared lysates were incubated with Glutathione Sepharose 4 Fast Flow beads (GE Healthcare, Chicago, IL) for 5 hr at 4°C. Bound proteins were eluted with 10 mM reduced glutathione in 50 mM Tris pH 8.0 and mixed with 30% glycerol. Protein concentration was determined by a BCA assay (Pierce, Puyallup, WA). Purified CHK2 was incubated in a kinase buffer (10 mM HEPES pH 7.4, 2.5 mM β-glycerolphosphate, 2 mM EDTA, 1 mM EGTA, 4 mM $MgCl_2$, 100 μM ATP) with GST-KAP1 substrate (2 μg) for 20 min at 30°C and its phosphorylation was detected by immunoblotting using KAP1-pS473 antibody. Alternatively, wild-type or mutant EGFP-CHK2 was immunoprecipitated from transfected HEK293 cells using GFP-Trap (Chromotek, Munich, Germany), treated with λ-phosphatase (200 U/reaction, Santa Cruz). Beads were washed three times with PBS and incubated for 20 min at 30°C with GST-KAP1 in the kinase buffer supplemented with PhosSTOP inhibitor (Roche, Basel, Switzerland). Alternatively, CHK2 kinase activity was measured in crude bacterial lysates *in vitro* using Omnia kinase assay kit (Life Technologies) as described previously.[19]

Statistical analysis

The patients were stratified according to (*i*) functional classes of germline *CHEK2* variants (deleterious, intermediate, neutral), (*ii*) the presence of a mutation in other (i.e., non-*CHEK2*) CPG and (*iii*) cancer and histopathological characteristics. Associations between the *CHEK2* mutation status and cancer diagnoses were analyzed using 3,360 PMC. The strength of the associations was estimated by the odds ratio (OR) in Fisher's exact test and *p* values <0.05 were considered significant.

Results

Germline CHEK2 variants are more frequent in cancer patients than in PMC

We analyzed germline *CHEK2* variants in 1,928 high-risk Czech BC/OC patients and 3,360 PMCs. We identified 36 distinct nonsynonymous variants (Table 1) in 131/1,928 (6.79%) patients and 142/3,360 (4.23%) PMC ($p = 7.4 \times 10^{-5}$).

Ten different frame-shift and splicing mutations ("All truncations" in Table 1) were found in 46 patients (2.39%) and 11 PMC (0.33%; $p = 1.3 \times 10^{-11}$). The most prevalent alterations were LGRs, present in 20 (1.04%) patients and four PMC (0.12%). LGRs included a recurrent exon 9–10 (5,395 bp) deletion and a novel exon 8 (5,601 bp) deletion. The c.1100delC mutation was found in seven (0.36%) patients and three PMC (0.09%). We identified three spliceogenic variants altering the mRNA sequence: c.444+1G>A, recurrent, population-specific c.846+4_846+7delAGTA (resulting in in-frame exon 7 skipping), and c.1260-8A>G (splice acceptor-shift with 7b exonization; Supporting Information Fig. S1). Variants reported as pathogenic in the ClinVar database, causing a frame-shift or truncating the kinase domain were considered pathogenic. Five of 46 patients with a truncating *CHEK2* mutation (four with female BC and one with double primary BC/OC) carried an additional pathogenic mutation in *BRCA1* or *BRCA2* (but not in another CPG). These patients were assigned into a group of 424 other CPG-mutation carriers.

Twenty-six distinct missense variants were found in 88 (4.56%) patients and 131 (3.90%) PMC (*p* = 0.22; Table 1). The most frequent variant was p.I157T with comparable prevalence in patients (58 carriers; 3.01%) and PMC (104 carriers; 3.10%; *p* = 0.93). Functional consequences of the detected missense variants predicted *in silico* yielded contradictory results (Supporting Information Table S2). While MutationTaster, CADD, and GERP predicted all SNVs as deleterious (except a maximum of 3/26 scored as neutral), the remaining four prediction tools, Align-GVGD, SIFT, PolyPhen2 and Spidex, were 100% and ≥75% concordant for 4/26 and 16/26 variants, respectively. Since the clinical significance of the detected SNVs was described as uncertain or conflicting in the ClinVar database (Table 1), we subjected them to subsequent functional analyses.

Functional assays identified deleterious CHEK2 missense variants

To evaluate the enzymatic activity of the identified CHK2 protein variants, we developed a cell-based assay quantifying KAP1-S473 phosphorylation in nontransformed human RPE1 cells. First, we verified the specificity of a monoclonal antibody against phosphorylated KAP1-S473 by immunoblotting and immunofluorescence microscopy (Supporting Information Fig. S2A). Next, we used the CRISPR/Cas9 technology to inactivate *CHEK2* in RPE1 cells (RPE1-*CHEK2*-KO; Fig. 1*a*, Supporting Information Figure S2B). A complete loss of CHK2 as well as RNAi-mediated CHK2 depletion impaired KAP1-S473 phosphorylation in RPE1 cells after ionizing radiation exposure. In contrast, CHK2 loss did not affect the phosphorylation of KAP1 at S824, an established ATM kinase site (Fig. 1*a*). A similar effect was also observed after treating the cells with neocarzinostatin and etoposide (Supporting Information Fig. S2C), suggesting that CHK2 phosphorylates KAP1 at S473 after the induction of DNA damage in general. A stable expression of EGFP-CHK2 in RPE1-*CHEK2*-KO cells rescued the phosphorylation of KAP1 at S473 after exposure to ionizing radiation, further confirming that CHK2 phosphorylates KAP1 after genotoxic stress (Fig. 1*b*). Finally, we transiently expressed the wild-type or mutant CHK2 isoforms in RPE1-*CHEK2*-KO cells and quantified the level of KAP1-S473 phosphorylation by immunofluorescence microscopy (Fig. 2*a*). We supplemented this cell-based model with a semiquantitative measurement of KAP1-pS473 in a cell-free *in vitro* assay using purified CHK2 and GST-KAP1 peptide as a substrate (Fig. 2*b*).

Cancer Genetics and Epigenetics

Table 1. The prevalence of *CHEK2* germline variants

Variant; cDNA (reference: NM_007194.3)	Variant; protein	rs number	ClinVar class	Unilateral FBC (n = 1,298)	Bilateral FBC (n = 149)	MBC (n = 48)	BC and OC (n = 79)	OC only (n = 354)	All patients (n = 1,928)	PMC (n = 3,360)
TRUNCATING mutations ([a]frame-shift; [b]in-frame)										
c.100–101delCA[a]	p.Q34Vfs*42	NA	NA	–	–	–	1	–	1	–
c.277delT[a]	p.W93Gfs*17	rs786203458	5	3	2	–	–	–	5	–
c.283C›T[a]	p.R95*	rs587781269	5	–	–	–	–	–	–	1
c.366delA[a]	p.E122Dfs*8	rs1555927302	5	1	–	–	–	–	1	–
c.444+1G›A[a]	p.E149Ifs*6	rs121908698	5	4	1	–	–	–	5	2
c.846+4_846+7delAGTA[b]	p.D265_H282del	rs764884641	3	7	–	–	–	–	7	–
c.846+1888_908+987del5601[a]	p.P283Dfs*8	NA	NA	2	–	–	–	–	2	–
c.909-2028_1095+330del5395[a]	p.M304Lfs*16	NA	5	11	1	3	1	2	18	4
c.1100delC[a]	p.T367Mfs*15	rs555607708	5	5	–	–	1	1	7	3
c.1260-8A›G[a]	p.L421Ifs*4	rs863224747	3	1	–	–	–	–	1	1
All truncations (%)				33 (2.54)[1]	4 (2.68)	3 (6.25)	3 (3.80)	3 (0.85)	46 (2.39)[1]	11 (0.33)
p-value (Fisher exact test)				9.4×10^{-11}	0.003	8.6×10^{-4}	0.004	0.14	1.3×10^{-11}	Ref.
Missense *CHEK2* mutations classified as DELETERIOUS										
c.190G›A	p.E64K	rs141568342	3–4	3	–	–	–	1	4	2
c.503C›T	p.T168I	rs730881684	3	–	–	1	1	–	2	–
c.520C›G	p.L174V	rs876659400	3	1	–	–	–	–	1	–
c.917G›C	p.G306A	rs587780192	3–4	1	–	–	–	–	1	2
c.980A›G	p.Y327C	rs587780194	3	1	–	–	–	–	1	–
c.1037G›A	p.R346H	rs730881688	3	1	–	–	–	–	1	–
c.1180G›A	p.E394K	rs587780169	3	1	–	–	–	–	1	–
c.1183G›C	p.V395 L	rs587780170	3	–	–	–	–	1	1	–
c.1270T›C	p.Y424H	rs139366548	3	–	1	–	–	–	1	–
c.1274C›T	p.P425L	rs1555913537	3	1	–	–	–	–	1	–
c.1421G›A	p.R474H	rs121908706	3	–	–	–	–	1	1	2
All deleterious missense variants (%)				9 (0.69)	1 (0.67)	1 (2.08)	1 (1.27)	3 (0.85)	15 (0.78)	6 (0.18)
p-value (Fisher exact test)				0.009	0.26	0.09	0.15	0.047	0.002	Ref.
Missense *CHEK2* variants classified as INTERMEDIATE										
c.470T›C	p.I157T	rs17879961	3–5	38	6	2	3	9	58	104
c.688G›T	p.A230S	rs748636216	3	–	–	–	–	–	–	1
c.715G›A	p.E239K	rs121908702	3	–	–	–	–	–	–	2
c.1067C›T	p.S356L	rs121908703	3	–	–	–	–	–	–	1
c.1217G›A	p.R406H	rs200649225	2–3	–	–	–	–	–	–	1
All intermediate missense variants (%)				38 (2.93)[2]	6 (4.03)[2]	2 (4.17)	3 (3.80)	9 (2.54)	58 (3.01)[2]	109 (3.24)
p-value (Fisher exact test)				0.64	0.63	0.67	0.74	0.63	0.68	Ref.

(Continues)

Table 1. The prevalence of *CHEK2* germline variants (Continued)

Variant; cDNA (reference: NM_007194.3)	Variant; protein	rs number	ClinVar class	Unilateral FBC (*n* = 1,298)	Bilateral FBC (*n* = 149)	MBC (*n* = 48)	BC and OC (*n* = 79)	OC only (*n* = 354)	All patients (*n* = 1,928)	PMC (*n* = 3,360)
Missense *CHEK2* variants classified as NEUTRAL										
c.7C›T	p.R3W	rs199708878	3	–	–	1	–	–	1	–
c.538C›T	p.R180C	rs77130927	1–3	1	–	–	–	–	1	3
c.539G›A	p.R180H	rs137853009	3	1	–	–	–	–	1	1
c.541C›T	p.R181C	rs137853010	3	–	–	–	–	–	–	3
c.542G›A	p.R181H	rs121908701	3	1	–	1	–	–	2	–
c.1091T›C	p.I364T	rs774179198	3	1	–	–	–	–	1	–
c.1309A›G	p.K437E	rs764238637	3	1	–	–	–	–	1	–
c.1312G›T	p.D438Y	rs200050883	3	3	–	–	–	1	4	2
c.1427C›T	p.T476M	rs142763740	3–4	2	–	–	–	1	3	3
c.1525C›T	p.P509S	rs587780179	3	1	–	–	–	1	2	4
All neutral missense variants (%)				11 (0.85)	–	2 (4.17)	–	3 (0.85)	16 (0.83)	16 (0.48)
p-value (Fisher exact test)				0.14		0.03		0.42	0.14	Ref.
All CHEK2 missense variants (%)				58 (4.47)	7 (4.70)	5 (10.42)	4 (5.06)	14 (3.95)[3]	88 (4.56)[3]	131 (3.90)
p-value (Fisher exact test)				0.38	0.52	0.04	0.55	0.96	0.22	Ref.

The prevalence of individual variants (divided into subgroups of truncating mutations and missense variants classified according to the results of an RPE1-*CHEK2*-KO cell-based analysis as deleterious, intermediate and neutral; Fig. 2*a*). It is displayed for all patients, their subgroups (unilateral female BC [FBC], bilateral FBC, male BC [MBC], double primary BC and OC and OC only) and population-matched controls (PMC; used as the reference). NA, not available.

[1]Include a FBC compound heterozygote of c.277delT and c.444+1G›A.

[2]Include two p.I157T homozygotes (with unilateral and bilateral FBC all diagnosed at ‹50 years, respectively).

[3]Four other compound heterozygotes in patients group were carriers of p.D265_H282del+p.D438Y, c.5601del+p.I157T, c.1100delC+p.I157T and p.E64K+p.I157T. The NM_007194.3 CHEK2 transcription variant A was used as the reference.

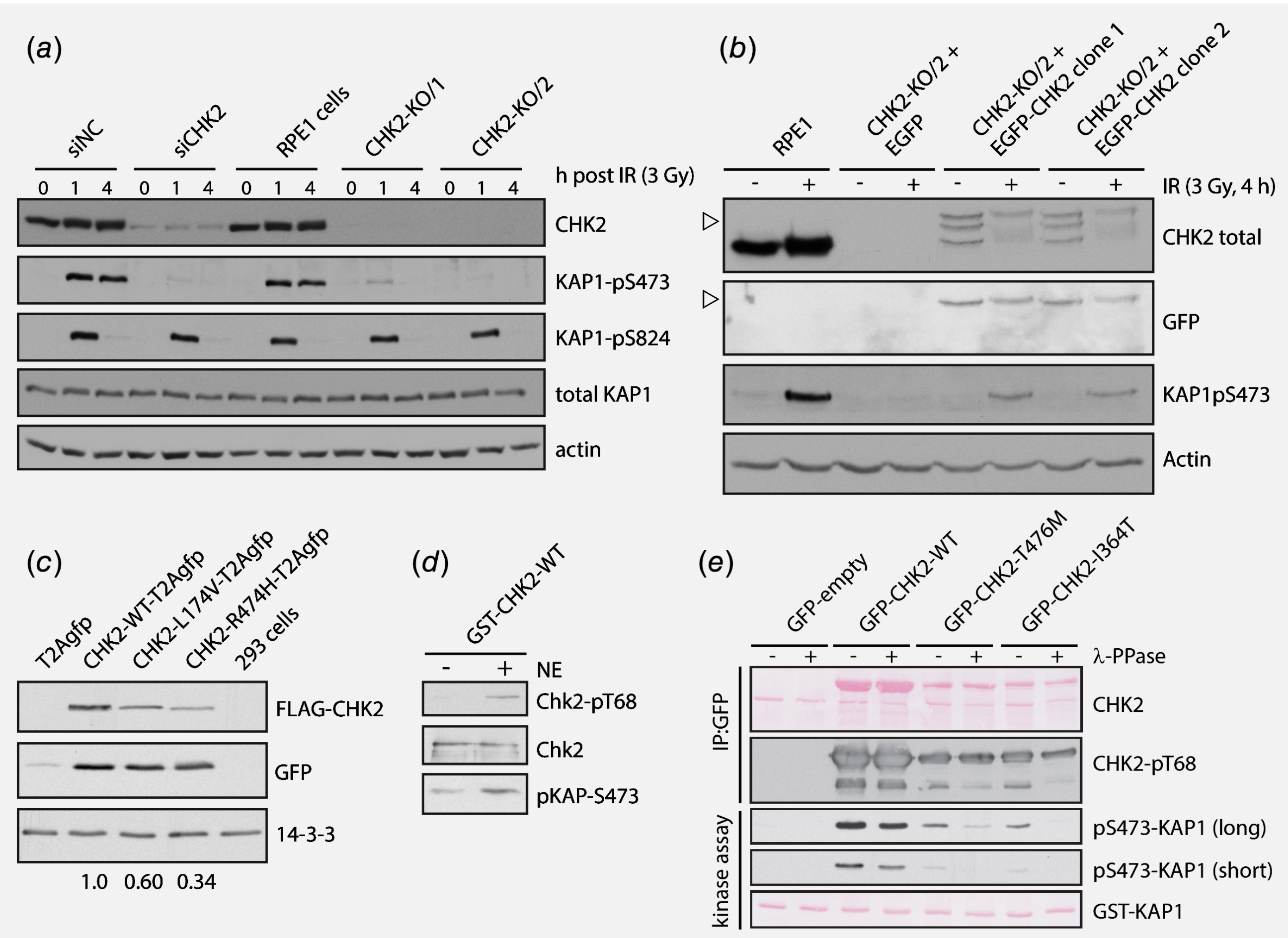


Figure 1. Characterization of a model system for functional analysis of *CHEK2* variants in RPE1-*CHEK2*-KO cells. Comparison of CHK2 depletion and knockout (*a*). RPE1 cells were transfected with control (siNC) or CHK2 siRNA (siCHK2) and assayed in parallel with parental RPE1 (RPE1 cells) and two clones of RPE1-*CHEK2*-KO cells (KO/1 and KO/2, respectively). Cells were harvested 0, 1 or 4 hr after exposure to IR (3 Gy) and analyzed by immunoblotting. Rescue of the *CHEK2* knockout (*b*). RPE1-*CHEK2*-KO cells were transfected with EGFP or EGFP-*CHEK2*-WT plasmids, selected with zeocin and exposed or not to IR. Parental RPE1 cells are shown for comparison. Arrowheads indicate the position of EGFP-CHK2. Impact of *CHEK2* mutations on protein stability (*c*). HEK293 cells were mock-treated or transfected with plasmids (1 μg) coding T2Agfp, CHK2-WT-T2Agfp, CHK2-L174V-T2Agfp or CHK2-R474H-T2Agfp and whole cell lysates were harvested after 20 hr. Numbers indicate the level of FLAG-CHK2 normalized to the level of GFP. Impact of CHK2 phosphorylation on its activity *in vitro* (*d*). Wild-type CHK2 purified from bacteria was incubated or not with nuclear extract from HCT116 cells (NE) in the presence of ATP at 30°C. After the addition of PBS, CHK2 was purified again using glutathione beads. Eluted CHK2 was incubated with KAP1 substrate and phosphorylation was assayed by KAP1-pS473 antibody. Impact of CHK2 phosphorylation on its activity (*e*). HEK293 cells were transfected with plasmids coding EGFP, EGFP-CHK2-WT, EGFP-CHK2-T476M, EGFP-CHK2-I364T. After 48 hr proteins were immunoprecipitated by GFP-Trap and treated or not with λ-phosphatase. Kinase activity was measured in the presence of phosphatase inhibitors and using GST-KAP1 as a substrate (shown in short and long exposition, respectively). [Color figure can be viewed at wileyonlinelibrary.com]

The results of the KAP1 cell-based analysis were in full agreement with KAP1 *in vitro* assays for 13 out of 26 tested missense variants, deleterious mutations (p.D265_H282del and c.1100delC) and wild-type CHK2 (Fig. 2, Supporting Information Table S2). Another five SNVs agreed partially between these two KAP1 assays (being intermediate in one and deleterious or neutral in the complementary assay). Eight variants were discrepant between cell based and *in vitro* KAP1 assays. As an example of discrepant results, the p.L174V variant showed only slightly decreased catalytic activity *in vitro*, but failed to phosphorylate KAP1 in cells. A comparison of the expression levels of CHK2-V174 and wild-type CHK2 both expressed from the bicistronic vector together with GFP (Fig. 1*c*) showed a suppressed expression of p.L174V to ~60% of wild-type CHK2, most probably reflecting impaired folding and/or reduced protein stability. Surprisingly, some variants with low *in vitro* activity were still able to phosphorylate KAP1 in human cells to a similar extent as wild-type CHK2. We hypothesized that the CHK2 kinase activity in human cells is influenced by its posttranslational modifications and, therefore, may differ from bacterially expressed CHK2. Indeed, preincubation of CHK2 purified from bacteria with nuclear extract led to CHK2-T68 phosphorylation. Subsequently, modified CHK2 showed higher ability to phosphorylate KAP1-S473 compared to

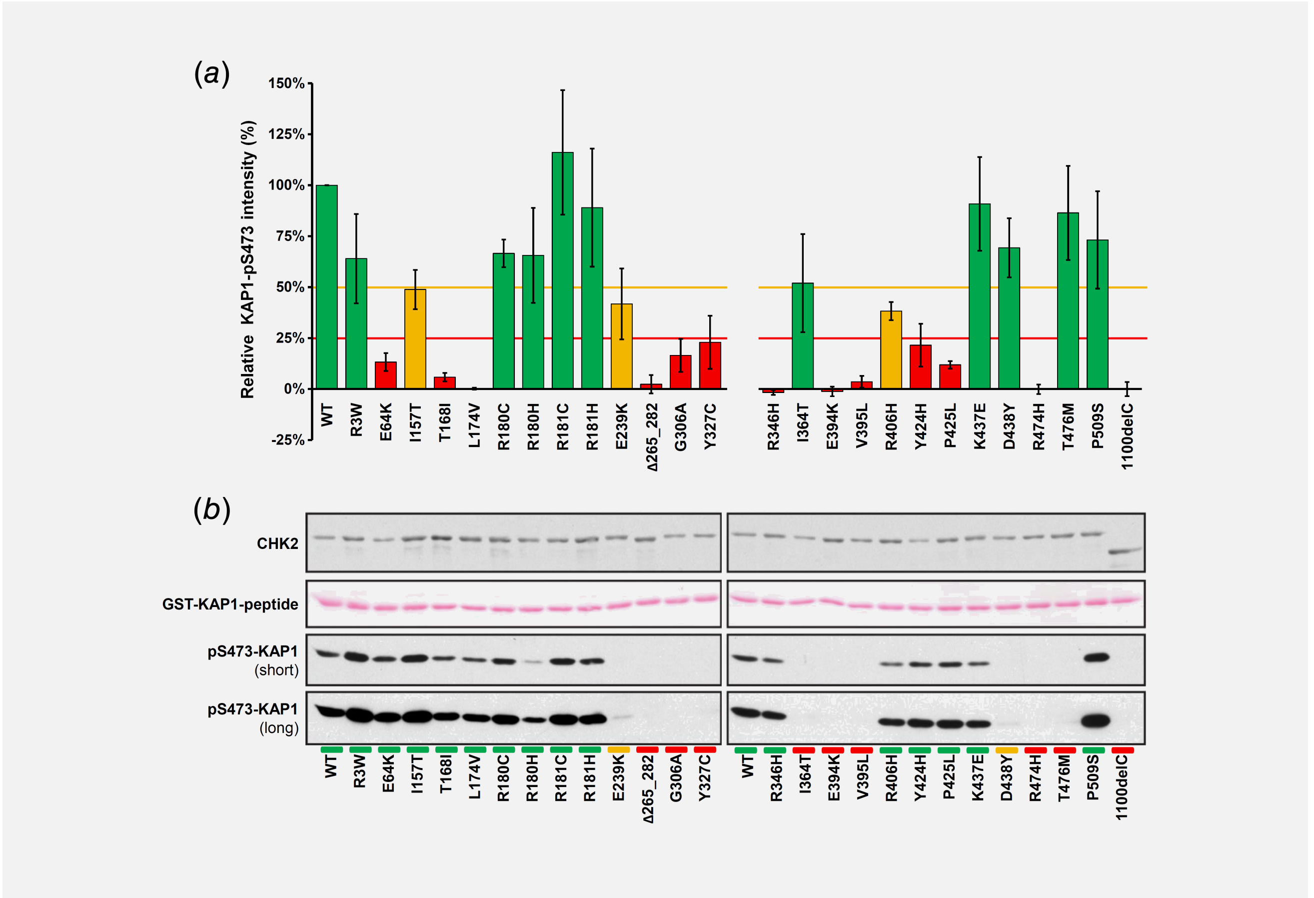


Figure 2. Functional classification of *CHEK2* germline variants was based on RPE1-*CHEK2*-KO cell-based assay. The chart describes relative levels of CHK2-dependent KAP1-S473 phosphorylation in RPE1-*CHEK2*-KO cells (*a*) for detected CHK2 variants. Variants were scored according to the WT (100%) and c.1100delC (0%) CHK2 kinase activity: >50% as "neutral" (green), 25–50% as "intermediate" (yellow) and <25% as "deleterious" (red). Error bars represent standard deviations (SD). Immunoblotting of phosphorylated GST-purified KAP1-peptide at S473 by purified CHK2 isoforms *in vitro* (*b*) was used to complement the assay in RPE1 cells. The individual panels show amounts of particular CHK2 isoforms and GST-KAP1-peptide, and intensity of KAP1-pS473 staining after incubation with purified CHK2 (in short and long exposition, respectively). Colors bars represent classifications from *a*; Δ265_282 means p.D265_H282del. (See online version for color images). *Note*: Variants p.A230S and p.S356L found in PMC (exome samples; not shown in this figure) were functionally classified by the RPE1-CHEK2-KO cell-based assay as intermediate (Supporting Information Table S2). [Color figure can be viewed at wileyonlinelibrary.com]

unmodified CHK2 (Fig. 1*d*). Conversely, the phosphatase treatment of CHK2 immunoprecipitated from HEK293 cells suppressed the *in vitro* activity of p.T476M and p.I364T variants that originally scored well in the cell-based assay (Fig. 1*e*). Our results suggest that posttranslational modifications substantially modulate CHK2 kinase activity and thus the human cell-based assay may better reflect the real CHK2 kinase activity *in vivo*. We also functionally analyzed detected VUS using commercial Omnia kinase *in vitro* assay that fully or partially corresponded to a principally comparable KAP1 *in vitro* assay for 23/26 VUS (Supporting Information Table S2, Fig. S3); however, was unable to dissect VUS discordant between KAP1 assays. Therefore, results from our cell-based assay (Fig. 2*a*), that reflects *in vivo* behavior of analyzed CHK2 variants more appropriately, led us to use solely this assay for the final functional VUS classification (Table 1).

The cell-based assay revealed strongly reduced kinase capacity (<25% of wild-type CHK2) for 11/26 missense variants that were classified as deleterious (Fig. 2*a*). These variants were significantly enriched in patients over PMC (Table 1). A significantly reduced kinase activity was also observed in recurrent c. 846+4_846+7delAGTA (in-frame exon 7 deletion; p.D265_H282del) eliminating the structurally important αC helix (residues 269–280) in the kinase domain.[7] The available pedigrees of patients with deleterious missense variants and c.846+4_846+7delAGTA are provided in the Supporting Information Figure S4. Five missense variants (p.I157T and four VUS identified only in PMC) were functionally classified as intermediate, with kinase activity at 25–50% of wild-type CHK2 in the cell-based assay. Ten missense variants with normal or mildly reduced catalytic activity (retaining >50% of wild-type CHK2) were considered neutral.

CHEK2 mutations are associated with BC and OC risk

We evaluated the association of *CHEK2* germline variants and cancer risk in diagnosis subgroups, considering all 1,928 patients and separately 1,504 patients without other CPG mutation. Regardless of the presence of other CPG mutations, truncating *CHEK2* variants significantly increased cancer risk in all analyzed subgroups except patients with OC only (Table 2). The most significant association was identified for group of 1,298 unilateral female BC patients that included 33 carriers (2.54%) of *CHEK2* truncations (OR = 7.94; 95%CI 3.90–17.47; $p = 9.4 \times 10^{-11}$). Truncations in *CHEK2* had the third highest mutation rate in this subgroup, preceded by *BRCA1* (153 carriers; 11.79%) and *BRCA2*

Table 2. Risk associated with germline *CHEK2* truncating and functionally classified missense variants (deleterious, intermediate and neutral) in all analyzed patients and in a subgroup of patients negatively tested for mutations in other cancer-predisposing genes against frequencies of *CHEK2* variants found in Czech population-matched controls PMC, Table 1

Group of patients	All patients			Other cancer-predisposing genes wt patients		
***CHEK2* variant group**	**Carriers; *N* (%)**	**OR (95%CI)**	***p*-value**	**Carriers; *N* (%)**	**OR (95%CI)**	***p*-value**
Unilateral female BC (I)	*n* = 1,298			*n* = 1,065		
Truncations	33 (2.54)	**7.94 (3.90–17.47)**	**9.4×10^{-11}**	29 (2.72)	**8.52 (4.11–18.97)**	**1.2×10^{-10}**
Deleterious missense	9 (0.69)	**3.90 (1.24–13.35)**	**0.009**	8 (0.75)	**4.23 (1.28–14.82)**	**0.008**
Intermediate missense	38 (2.93)	0.90 (0.60–1.32)	0.64	34 (3.19)	0.98 (0.64–1.47)	0.99
Neutral missense	11 (0.84)	1.79 (0.75–4.11)	0.14	10 (0.94)	1.98 (0.80–4.66)	0.11
Bilateral female BC (II)	*n* = 149			*n* = 104		
Truncations	4 (2.68)	**8.39 (1.92–28.74)**	**0.003**	4 (3.85)	**12.15 (2.77–41.94)**	**8.1×10^{-4}**
Deleterious missense	1 (0.67)	3.77 (0.08–31.42)	0.26	1 (0.96)	5.42 (0.12–45.31)	0.19
Intermediate missense	6 (4.03)	1.25 (0.44–2.88)	0.63	5 (4.81)	1.51 (0.47–3.74)	0.39
Neutral missense	0 (0)	–	–	0 (0)	–	–
Male BC (III)	*n* = 48			*n* = 39		
Truncations	3 (6.25)	**20.21 (3.50–80.00)**	**8.6×10^{-4}**	3 (7.69)	**25.23 (4.34–101.34)**	**4.7×10^{-4}**
Deleterious missense	1 (2.08)	11.87 (0.25–100.83)	0.10	1 (2.56)	14.66 (0.31–125.29)	0.08
Intermediate missense	2 (4.17)	1.30 (0.15–5.07)	0.67	2 (5.13)	1.61 (0.19–6.39)	0.37
Neutral missense	2 (4.17)	**9.07 (0.98–40.41)**	**0.03**	2 (5.13)	**11.26 (1.21–50.79)**	**0.02**
BC and OC (IV)	*n* = 79			*n* = 40		
Truncations	3 (3.80)	**11.99 (2.11–46.6)**	**0.004**	2 (5.00)	**15.97 (1.67–77.08)**	**0.01**
Deleterious missense	1 (1.27)	7.15 (0.15–59.97)	0.15	0 (0)	–	–
Intermediate missense	3 (3.80)	1.18 (0.24–3.67)	0.74	1 (2.50)	0.76 (0.02–4.61)	0.99
Neutral missense	0 (0)	–	–	0 (0)	–	–
OC only (V)	*n* = 354			*n* = 256		
Truncations	3 (0.85)	2.60 (0.46–9.91)	0.14	3 (1.17)	3.61 (0.64–13.78)	0.07
Deleterious missense	3 (0.85)	**4.77 (0.77–22.47)**	**0.047**	3 (1.17)	**6.62 (1.07–31.22)**	**0.02**
Intermediate missense	9 (2.54)	0.78 (0.34–1.55)	0.63	8 (3.13)	0.96 (0.40–1.99)	0.99
Neutral missense	3 (0.84)	1.79 (0.33–6.28)	0.42	2 (0.78)	1.65 (0.18–7.06)	0.37
Any female BC (I + II + IV)	*n* = 1,526			*n* = 1,209		
Truncations	40 (2.62)	**8.19 (4.11–17.75)**	**4.1×10^{-12}**	35 (2.90)	**9.07 (4.49–19.87)**	**2.4×10^{-12}**
Deleterious missense	11 (0.72)	**4.06 (1.37–13.39)**	**0.006**	9 (0.74)	**4.19 (1.33–14.34)**	**0.006**
Intermediate missense	47 (3.08)	0.95 (0.66–1.35)	0.79	40 (3.31)	1.02 (0.69–1.49)	0.92
Neutral missense	11 (0.72)	1.52 (0.64–3.49)	0.30	10 (0.83)	1.74 (0.70–4.10)	0.18
Any OC (IV + V)	*n* = 433			*n* = 296		
Truncations	6 (1.39)	**4.28 (1.29–12.69)**	**0.009**	5 (1.69)	**5.23 (1.41–16.45)**	**0.007**
Deleterious missense	4 (0.92)	**5.21 (1.08–22.06)**	**0.02**	3 (1.01)	**5.72 (0.92–26.94)**	**0.03**
Intermediate missense	12 (2.77)	0.85 (0.42–1.56)	0.77	9 (3.04)	0.94 (0.41–1.87)	0.99
Neutral missense	3 (0.69)	1.46 (0.27–5.12)	0.47	2 (0.68)	1.42 (0.16–6.09)	0.65

The calculations were performed in individual diagnostic subgroups (Roman numerals I–V) and in aggregated groups of any female BC (subgroups I, II and IV) and any OC patients (subgroups IV and V). "Other CPG-wt" group consists of patients without germline mutations in genes predisposing for BC (*BRCA1*, *BRCA2*, *PALB2*, *TP53*) or OC (*BRCA1*, *BRCA2*, *RAD51C*, *RAD51D*, *MLH1*, *MSH2*, *MSH6*). Significant association of *CHEK2* variants with cancer risk is highlighted (in bold). Both aggregated subgroups (Any FBC and Any OC) include patients with double primary BC and OC (IV).

(56 carriers; 4.31%), and followed by *PALB2* (21 carriers; 1.62%) and *TP53* (3 carriers; 0.23%). We also observed a significantly higher prevalence of *CHEK2* truncations in small subgroups of patients with bilateral female BC (4/149; 2.68%; $p = 0.003$), male BC (3/48; 6.25%; $p = 8.6 \times 10^{-4}$) and with double primary BC/OC (3/79; 3.80% $p = 0.004$); however, the low number of patients and mutations limits relevance of calculated ORs. The analysis of two aggregated subgroups of "any female BC" and "any OC" patients (overlapping in patients diagnosed with double primary BC/OC; Table 2) reflected clinically relevant overall risk for BC and OC development in females with *CHEK2* truncations. We found significant associations with both cancer types, which was substantially higher and more significant for "any female BC" (OR = 8.19; 95%CI 4.11–17.75; $p = 4.1 \times 10^{-12}$) than for "any OC" (OR = 4.28; 95%CI 1.29–12.69; $p = 0.009$) subgroups in all patients as well as in patients after excluding those with mutations in other CPG (OR = 9.07; 95%CI 4.49–19.87; $p = 2.4 \times 10^{-12}$ and OR = 5.23; 95%CI 1.41–16.45; $p = 0.007$, respectively).

While the frequencies of functionally deleterious SNV were significantly more frequent in unilateral female BC, OC, any female BC and also any OC subgroups (Tables 1 and 2), the frequencies of functionally neutral or intermediate SNVs did not differ from PMC in any patient subgroup (except for neutral SNVs in a small subgroup of 48 male BC patients). Risks associated with functionally deleterious SNV were lower than risks associated with truncations, except that in OC patients. However, low number of functionally deleterious SNV carriers makes our findings only suggestive but not conclusive.

Twelve out of 54 *BRCA1/BRCA2*-negative *CHEK2* mutation carriers had a VUS in other genes, in which further modification of cancer risk cannot be ruled out (Supporting Information Table S3).

CHEK2 mutations predispose to specific BC types and multiple cancer development

We evaluated histopathological tumor characteristics in 1,209 *other CPG-wt* female BC patients. Breast tumors in *CHEK2* mutation carriers differed from noncarriers, tended to be more frequently of luminal A and less frequently of basal BC subtype, with lower grade and with nonsignificant tendency toward lower clinical stage (Fig. 3; Supporting Information Table S4). Histology, menopausal status and indication criteria for testing did not differ among *CHEK2* mutation carriers and noncarriers. Although the most frequent p.I157T variant did not affect BC risk, its carriers had a similar tendency for BC subtype distribution. Phenotypical characteristics of functionally deleterious missense and truncating *CHEK2* mutation carriers were similar (Supporting Information Table S5).

Second primary cancers (other than BC/OC; Supporting Information Table S3) were diagnosed in *CHEK2* mutation carriers more frequently (10/54; 18.5%) than in carriers of other CPG mutations (25/424; 5.9%; $p = 0.003$) or noncarriers (110/1,403; 7.8%; $p = 0.01$). All 10 *CHEK2* mutation carriers with second cancer (developing 13 tumors together including two cases each of colon, thyroid, renal, head/neck cancers or hematological malignancy, and one case each of lung, urinary bladder or endometrial cancer) had a positive family cancer history.

Discussion

The frequency of germline truncating and splice site *CHEK2* mutation carriers in our study strongly prevailed in all patients over PMC (2.39% *vs.* 0.33%; $p = 1.3 \times 10^{-11}$) but the frequencies of missense variants were comparable (4.56% *vs.* 3.90%; $p = 0.22$). Most missense variants, especially in moderate risk genes (including *CHEK2*) are interpreted as inconclusive VUS, lacking clearly defined risk estimates and representing a major drawback for multigene testing in diagnostic settings.[26,27] Only several reports have described a functional characterization of *CHEK2* VUS by *in vitro*[19,22] or yeast models.[35,36] The *in vitro* assays measure CHK2 kinase catalytic activity over artificial substrate but do not reflect changes in CHK2 intracellular targeting, stability and posttranslational modifications. Moreover, transient CHK2 overexpression can cause its autophosphorylation even in the absence of DNA damage, bypassing necessity for CHK2-T68 phosphorylation and participation of FHA domain on CHK2 activation *in vivo*.[37] Yeast analyses are based on functional complementation of *RAD53*-defective *Saccharomyces cerevisiae* cells by human CHK2 homolog. A growth rate of the yeast cells upon DNA damage correlates with functional competence of the analyzed *CHEK2* variant in this assay. In contrast, our newly developed RPE1-*CHEK2*-KO cell-based assay allowed us to quantify catalytic activity of analyzed *CHEK2* variants in nontransformed human cells in the presence of CHK2 natural upstream activators and downstream substrates.

Altogether, results of functional analysis for 18/26 (69%) of analyzed missense VUS were in full agreement or partially overlapped between our KAP1 cell-based and *in vitro* analyses. Remaining eight variants (p.E64K, p.T168I, p.L174V, p.R346H, p.I364T, p.Y424H, p.P425L, p.T476M) scored discordantly. In subsequent analyses of p.L174V, p.I364T and p.T476M variants, we demonstrated that discordance between results of cell-based and *in vitro* assays resulted from their fundamental differences (Figs. 1*c*–1*e*). Variant p.L174V only mildly decreased KAP1 phosphorylation *in vitro*, but failed to phosphorylate KAP1 in cells. Further analysis revealed that this variant impairs intracellular protein stability explaining its functional defect in cells. This rare FHA domain variant was described once in ClinVar. We identified p.L174V in BC patient diagnosed at 35 years carrying also a pathogenic *BRCA1* mutation (Supporting Information Fig. S4). Variant p.I364T showed low KAP1 phosphorylation *in vitro* but was able to phosphorylate KAP1 in cells. Subsequent analysis demonstrated that CHK2-T364 protein was phosphorylated at T68 when immunoprecipitated from cells and that removing this modification by λ-phosphatase treatment strongly reduced its catalytic activity (Figs. 1*d* and 1*e*) comparable to that in wild-type CHK2. Moreover, Chrisanthar *et al.* described normal dimerization and autophosphorylation, and only mildly reduced kinase activity for p.I364T, concluding a nonaffected

Cancer Genetics and Epigenetics

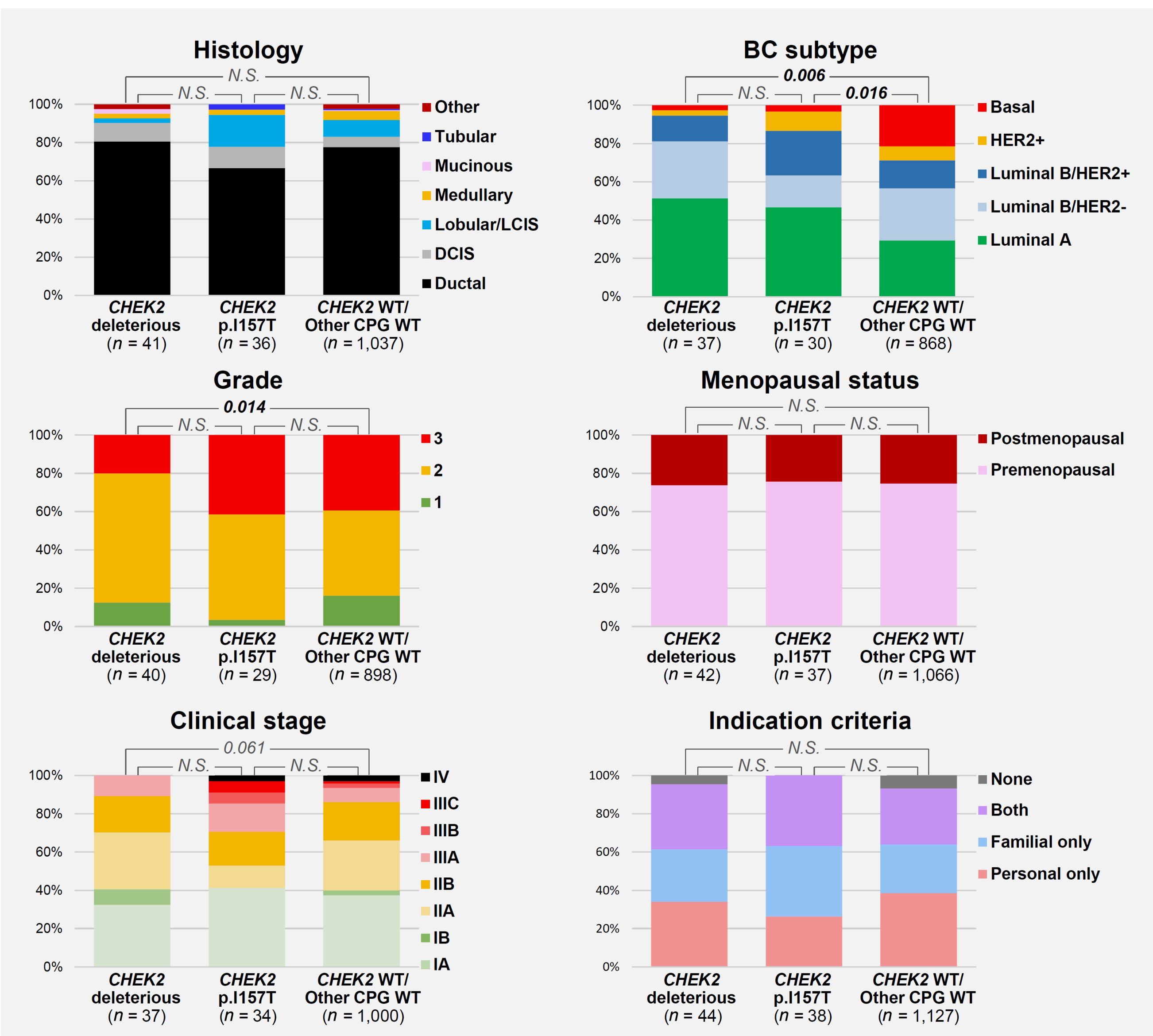


Figure 3. Clinical and histopathological characteristics of female BC patients. A subgroup of 1,209 other CPG-wt patients with any BC were stratified according to the presence of germline deleterious *CHEK2* mutation (truncating or pathogenic missense; $n = 44$), p.I157T ($n = 38$) and *CHEK2*-wt patients ($n = 1{,}127$), respectively. Significant differences between groups are highlighted in bold (N.S. denoted for not significant differences with $p < 0.1$). Numbers in parenthesis (n) characterize number of individuals with known values for particular characteristic. *Note*: "Other CPG-wt" group consists of patients without germline mutation in genes predisposing for BC (*BRCA1*, *BRCA2*, *PALB2*, *TP53*) or OC (*BRCA1*, *BRCA2*, *RAD51C*, *RAD51D*, *MLH1*, *MSH2*, *MSH6*). [Color figure can be viewed at wileyonlinelibrary.com]

kinase function;[38] Delimitsou *et al.* recently scored p.I364T by *S. cerevisiae* assay functionally intermediate (Supporting Information Table S2).[36] We identified this variant in premenopausal BC patient with no cancer diagnosed in first or second-degree relatives. The p.T476M variant behaved similarly as p.I364T, with T68 phosphorylation-dependent kinase activity (Fig. 1*e*). This variant was classified by Delimitsou intermediate, but previous analyses by Roeb *et al.*[35] and Desrichard *et al.*[19] (Supporting Information Table S2) scored p.T476M deleterious by yeast and *in vitro* assays, respectively. We found this variant in three patients and three PMC. Moreover, in concordance with our cell-based assay, the p.T476M was classified as likely benign by Myriad using history weighting algorithm.[39]

Another five discrepant variants were scored in our cell-based assay functionally deleterious. The p.E64K variant affecting SQ/TQ domain was previously analyzed by Wu *et al.*[40] who described its reduced autophosphorylation, CDC25C phosphorylation and severely impaired T68 phosphorylation and concluded

that p.E64K alters SQ/TQ domain conformation impairing CHK2 activation. Two later independent analyses showed mutually opposite results in yeast assays (Supporting Information Table S2).[35,36] We found p.E64K in one OC and three BC patients, including a carrier who developed three primary tumors (Supporting Information Fig. S4); however, two carriers were also identified in PMC, including a male (aged 68) and female (aged 63). We found no additional functional data for p.T168I, a variant localized to the FHA domain, functionally defective also in our Omnia kinase assay (Supporting Information Table S2). We detected p.T168I in a patient carrying a *BRCA2* mutation diagnosed with BC and OC (Supporting Information Fig. S4). Variant p.R346H, affecting kinase domain, was functionally classified deleterious also by Delimitsou *et al.*[36] and our Omnia kinase assay (Supporting Information Table S2). Moreover, in a BCAC study, Southey found an increased BC risk (OR = 5.06; 95%CI 1.09–23.5; $p = 0.017$) for p.R346C variant at the same position[41] and we observed a segregation of p.R346H with BC in analyzed HBC family (Supporting Information Fig. S4). The p.Y424H kinase domain variant was classified functionally defective by two out of three previous yeast-based analyses and in our Omnia kinase assay (Supporting Information Table S2). We detected p. Y424H in patient with double primary premenopausal BC with multiple cancers in family members. The p.P425L variant, affecting P425 participating in CHK2 kinase domain dimerization,[7] showed also partially reduced Omnia kinase assay activity. We found this variant in BC patients diagnosed at 47 years; however, no other relatives were available for the genetic analysis.

Conceptual differences in functional *CHEK2* assays contribute to discrepant findings for individual VUS, especially in variants sensitive to posttranslational CHK2 modifications. Hence, we think that our assay performed in human nontransformed cells provides an opportunity for realistic functional *CHEK2* VUS analysis. Estimated BC risks associated with functionally deleterious, intermediate and neutral variants (Table 2) revealed a lack of risk association for the latter two groups, supporting our correct functional classification. Altogether, functionally deleterious missense mutations were identified in 15 out of 88 *CHEK2* missense variant carriers (Table 1) constituting 20–25% of pathogenic *CHEK2* mutation in BC patients and 40% in OC patients. However, low number of carriers of functionally deleterious variants limited validity of presented data. The extension of our assay to large-scale *CHEK2* VUS analyses with evaluation of clinical data in their carriers will be required to validate our findings, including lower risk associated with functionally deleterious missense variants in comparison to truncations.

To calculate cancer risk for carriers of deleterious *CHEK2* mutations, we considered *all high-risk patients* and, in parallel, a subgroup of *CPG-wt patients.* The *all high-risk patients* group revealed the real proportion of *CHEK2* mutation carriers and associated cancer risk in a realistic context of all individuals indicated for genetic testing according to current guidelines. The analysis of the *CPG-wt* subgroup (raising the proportion of *CHEK2* mutation carriers by excluding 424 other CPG-mutation carriers of whom 90% carried a *BRCA1/BRCA2* mutation) allows to compare our findings with studies analyzing *BRCA1/BRCA2*-wt patients (Table 3).

We are aware that risk calculations have their specific limitations. Analyzed patients' groups were enriched in high-risk patients from multiple cancer families and, in contrast, PMC group share higher proportion of older noncancer individuals. Both factors can contribute to an overestimated risks found in our study. Other *CHEK2* studies also demonstrated higher OR found in analyses involving patients with familial BC (Table 3) indicating that a precise risk estimation will require a representative number of analyzed individuals and appropriately selected PMC. Higher cancer risks found in our study was affected also by high frequency of LGRs whose identification by panel NGS has been considered problematic[34] or omitted[26] in comparable analyses. Our data urge its careful evaluation in *CHEK2* analyses. Although the OR values calculated in our study must be interpreted with caution (especially in case of missense variants), our data clearly show that germline *CHEK2* mutations carriers are significantly enriched especially in the largest group of female BC patients. Interestingly, deleterious *CHEK2* mutations increased risk of male BC. *CHEK2* was the second most frequently mutated CPG in this small subgroup, preceded by *BRCA2* and followed by *BRCA1*, and *PALB2* (data not shown), indicating that germline *CHEK2* mutations contribute to male BC, as suggested previously.[51,53,54]

Deleterious *CHEK2* mutations were associated with a moderately increased OC risk in our study. However, due to the limited numbers of analyzed OC individuals with *CHEK2* mutations (10 in all patients, 4 in the CPG-negative subgroup), these observations need further validation. A substantial proportion of deleterious missense mutations (4/10) in OC patients indicates that their functional classification will be necessary for proper OC risk assessment.

Our analysis confirmed proposed "*CHEK2* mutation-specific" tumor phenotype, characterized by premenopausal, ductal, grade 2, luminal A or luminal B/HER2-negative tumors, reported in other studies.[25,26,46,55] These tumor characteristics lost in carriers of coincidental *BRCA1/BRCA2* mutations having a stronger effect on tumor phenotype. Nurmi *et al.*[42] identified an additive effect of mutations in moderate-penetrance genes, including *CHEK2*, increasing BC risk in Finnish *BRCA1/BRCA2* mutation carriers. The effect of coincidental alterations in other moderate-penetrance CPG with *CHEK2* mutations are unknown; however, the influence of a polygenic risk score on c.1100delC penetrance has been recently documented.[56]

A strongly increased frequency of second cancers of various origin in *CHEK2* mutation carriers and tumors in their relatives corresponds to documented multiorgan cancer susceptibility in *CHEK2* mutations carriers[5,25] and indicates that family cancer history associated with *CHEK2* mutations must be reconsidered to facilitate the selection of potential *CHEK2* mutation carriers for genetic analyses.

The p.I157T variant did not increase cancer risk in our study; an observation we have previously reported for sporadic BC

Table 3. Analyses of germline variants in the *CHEK2* gene or analyses of selected germline *CHEK2* variants in studies (upper part) and selected meta-analyses (lower part) calculating odds ratio for breast cancer development in mutation carriers

References	Pop.	P: patientsC: controls	Analysis	Odds ratio (95%CI); *p*—evaluated group or *CHEK2* variant
Nurmi *et al.*[42]	FI	P: 3156 BC or OC patients C: 2089 PMC	c.319+2T›A	5.40 (1.58–18.45); 0.007—unselected BC 6.04 (1.65–22.10); 0.007—familial BC
Girard *et al.*[43]	FR	P: 1,207 *BRCA1/2*-negative BC females having sister with BC C: 1,199 noncancer PMC	*CHEK2* (panel NGS)	3.0 (1.9–5.0); 1×10^{-5}—any variant 5.8 (2.0–16.9); 0.001—loss of function variant 2.4 (1.4–4.3); 0.002—likely deleterious missense
Hauke *et al.*[26]	DE	P: 5,589 *BRCA1/2*-negative BC C: 2,189 noncancer PMC	*CHEK2* (panel NGS)	3.72 (1.99–6.94); ‹0.0001—truncations
Couch *et al.*[24]	US	P: 29,090 BC C: ExAC-NFE non-TCGA	*CHEK2* (panel NGS)	2.31 (1.88–2.85); 3.04×10^{-17}—c.1100delC 2.26 (1.89–2.72); 1.75×10^{-20}—pathogenic variants (p.I157T, p.S428F excluded) 1.48 (1.31–1.67); 1.75×10^{-10}—any variant (p.I157T, p.S428F included) 1.35 (1.12–1.63); 0.0002; bilateral BC
Decker *et al.*[44]	UK	P: 13,087 BC C: 5,488 PMC	*CHEK2* (4 genes)	3.11 (2.15–4.69); 5.6×10^{-11}—truncations 1.36 (0.99–1.87); 0.066—all rare missense 1.51 (1.02–2.24); 0.047—rare missense in any domain 3.27 (1.66–5.83); 0.0014—bilateral BC
Slavin *et al.*[45]	US	P: 2,266 *BRCA1/2*-neg. Fam. BC C: ExAC	*CHEK2*	1.62 (1.03–2.51); 0.004—truncations
Schmidt *et al.*[46]	BCAC	44,777 BC 42,977 PMC	c.1100delC	2.26 (1.90–2.69); 2.3×10^{-20}—invasive BC 2.55 (2.10–3.10); 4.9×10^{-21}—ER-positive BC 1.32 (0.93–1.88); 0.12—ER-negative BC
Southey *et al.*[41]	BCAC	P: 42,671 C: 42,164 PMC	c.349A›G (p.R117G) c.538C›T (p.R180C) c.715G›A (p.E239K) c.1036C›T (p.R346C) c.1312G›T (p.D438Y)	2.26 (1.29–3.95); 0.003—for variant p.R117G 1.33 (1.05–1.67); 0.016—for variant p.R180C 1.70 (0.73–3.93); 0.210—for variant p.E239K 5.06 (1.09–23.5); 0.017—for variant p.R346C 1.03 (0.62–1.71); 0.910—for variant p.D438Y
Cybulski *et al.*[47]	PL	P: 7,494 *BRCA1*-negative BC C: 4,346 PMC	c.1100delC, c.444+1G›A, del5395	3.6 (2.6–5.1)—BC 3.3 (2.3–4.7)—patients with no BC family history 5.0 (3.3–7.6)—patients with BC in first or second degree relatives 7.3 (3.2–16.8)—patients with BC in first and second degree relatives
Desrichard *et al.*[19]	FR	P: 507 *BRCA1/2*-negative BC C: 513 noncancer PMC	*CHEK2*	4.15 (1.38–12.50); 0.007—all *CHEK2* variants 5.18 (1.49–18.00); 0.004—*CHEK2* mutations (p.K244R ex)
Le Calvez-Kelm *et al.*[20]	US, AU	P: 1303 BC ≤45 years C: 1,109 noncancer females	*CHEK2*	6.18 (1.76–21.8)—truncations 2.20 (1.20–4.01)—rare missense
Weischer *et al.*[48]	DK	P: 1,101 BC C: 4,665 PMC	c.1100delC	3.2 (1.0–9.9)—BC (prospective study) 2.6 (1.3–5.4)—BC (case–control study)

(Continues)

Table 3. Analyses of germline variants in the *CHEK2* gene or analyses of selected germline *CHEK2* variants in studies (upper part) and selected meta-analyses (lower part) calculating odds ratio for breast cancer development in mutation carriers (Continued)

References	Pop.	P: patientsC: controls	Analysis	Odds ratio (95%CI); *p*—evaluated group or *CHEK2* variant
Cybulski *et al.*[5]	PL	P: 1,017 BC C: 4,000 PMC	c.1100delC; c.444+1G›A; p.I157T	2.2; $p = 0.02$—for c.1100delC and c.444+1G›A 1.4; $p = 0.02$—for p.I157T
Dufault *et al.*[21]	DE	P: 516 *BRCA1/2*-negative BC C: 1,315 random PMC	*CHEK2*	3.44 (1.19–9.95); 0.016—c.1100delC 3.9 (1.3–10.9)—c.1100delC and c.1214del4
CHEK2 Breast Cancer Case-Control Consortium[49]	UK, NL, FI, DE, AU	P: 10,860 BC C: 9,065 PMC	c.1100delC	2.34 (1.72–3.20); 1×10^{-7} 2.23 (1.60–3.11)—BC with no BC in first degree relative 3.12 (1.90–5.15)—BC with 1 BC in first degree relative 4.17 (1.26–13.75)—BC with ≥2 BC in first degree relative
Vahteristo, 2002[50]	FI	1,035 unselected BC 1885 PMC	c.1100delC	1.48 (0.83–2.65); 0.182 unselected BC 2.27 (1.11–4.63); 0.021 familial BC 6.17 (1.87–20.32); 0.007 bilateral BC
Liang *et al.*[51]	Meta	P: 118,735 BC C: 195,807	c.1100delC	2.88 (2.65–3.16)—female BC 2.87 (1.85–4.47)—early onset BC 3.21 (2.41–4.29)—familial BC 3.13 (1.94–5.07)—male BC
Liu *et al.*[16]	Meta	P: 19,621 BC C: 27,001	p.I157T	1.48 (1.31–1.66); ‹0.0001—unselected BC 1.48 (1.16–1.89); ‹0.0001—familiar BC 1.47 (1.29–1.66); ‹0.0001—early onset BC 4.17 (2.89–6.03); ‹0.0001—lobular BC
Zhang *et al.*[52]	Meta	P: 9,970/ C:7,526 P: 13,331/C: 10,817 P: 10,543/ C:10,817 P: 4,1,791/C: 50,910	c.444+1G›A p.I157T del5395 c.1100delC	3.07 (2.03–4.63); 9.82×10^{-8}—for variant c.444 +1G›A 1.52 (1.31–1.77); 4.76×10^{-8}—for variant p.I157T 2.53 (1.61–3.97); 6.33×10^{-5}—for variant del5395 3.10 (2.59–3.71); ‹10^{-20}—for variant c.1100delC
Weischer *et al.*[15]	Meta	P: 26,488 C: 27,402	c.1100delC	2.7 (2.1–3.4)—unselected BC 2.6 (1.3–5.5)—early onset BC 4.8 (3.3–7.2)—familial BC

Abbreviations: AU, Australia; BC, breast cancer; BCAC, Breast Cancer Association Consortium; CN, China; DE, Germany; EU, European union; FI, Finland; DK, Denmark; FR, France; meta, meta-analysis; NL, Nederland; PL, Poland; US, USA.

patients.[57] With OR = 1.5 reported in numerous studies (Table 3), is below the threshold considered for moderate-penetrance genes (OR > 2) and together with a high frequency in PMC it negates a clinically considerable effect on BC risk. We noticed a higher proportion of lobular BC in p.I157T carriers (Fig. 3), known from previous studies.[16,58,59] Our functional analysis classified p.I157T as an "intermediate" variant with catalytic activity reaching 48.8% of wild-type CHK2. Hence, an increased cancer risk cannot be ruled out in homozygote p.I157T carriers.

In conclusion, our study demonstrated a substantial clinical relevance of a *CHEK2* analysis in high-risk BC/OC patients, supported by the results of a cell-based functional assay markedly reducing the number of VUS. In addition, the high frequency of non-BC/OC tumors in *CHEK2* mutation carriers and their relatives warrants further investigation by collaborative international efforts.

Acknowledgements

We thank Prof. Yves-Jean Bignon who kindly provided the *CHEK2* plasmid, Prof. Martin Haluzik and Dr. Petra Kavalkova for their help with FluoroStar analysis, Jan Flemr for language editing. We thank all patients and their families for their participation in our study. The study was supported by grants of the Ministry of Health of the Czech Republic NV15-28830A, NV15-27695A, NV16-29959A, NV19-03-00279 and the grants of Charles University PROGRES Q28/LF1 and Q26/LF1, GAUK 762216, SVV2019/260367, PRIMUS/17/MED/9 and UNCE/MED/016. We thank The National Center for Medical Genomics (LM2015091) for providing allelic frequencies in ethnically matched populations (project CZ.02.1.01/0.0/0.0/16_013/0001634) and NPU II LQ1604. LM was partially supported by the Academy of Sciences of the Czech Republic (Strategie AV21, Qualitas).

References

1. Sun J, Meng H, Yao L, et al. Germline mutations in cancer susceptibility genes in a large series of unselected breast cancer patients. *Clin Cancer Res* 2017;23:6113–9.
2. Norquist BM, Harrell MI, Brady MF, et al. Inherited mutations in women with ovarian carcinoma. *JAMA Oncol* 2016;2:482–90.
3. Kleibl Z, Kristensen VN. Women at high risk of breast cancer: molecular characteristics, clinical presentation and management. *Breast* 2016;28: 136–44.
4. Easton DF, Pharoah PD, Antoniou AC, et al. Gene-panel sequencing and the prediction of breast-cancer risk. *N Engl J Med* 2015;372: 2243–57.
5. Cybulski C, Gorski B, Huzarski T, et al. CHEK2 is a multiorgan cancer susceptibility gene. *Am J Hum Genet* 2004;75:1131–5.
6. Matsuoka S, Rotman G, Ogawa A, et al. Ataxia telangiectasia-mutated phosphorylates Chk2 in vivo and in vitro. *Proc Natl Acad Sci U S A* 2000;97:10389–94.
7. Cai Z, Chehab NH, Pavletich NP. Structure and activation mechanism of the CHK2 DNA damage checkpoint kinase. *Mol Cell* 2009;35:818–29.
8. Zannini L, Delia D, Buscemi G. CHK2 kinase in the DNA damage response and beyond. *J Mol Cell Biol* 2014;6:442–57.
9. White DE, Negorev D, Peng H, et al. Rauscher FJ, 3rd. KAP1, a novel substrate for PIKK family members, colocalizes with numerous damage response factors at DNA lesions. *Cancer Res* 2006; 66:11594–9.
10. Hu C, Zhang S, Gao X, et al. Roles of Kruppel-associated box (KRAB)-associated co-repressor KAP1 Ser-473 phosphorylation in DNA damage response. *J Biol Chem* 2012;287:18937–52.
11. Jallepalli PV, Lengauer C, Vogelstein B, et al. The Chk2 tumor suppressor is not required for p53 responses in human cancer cells. *J Biol Chem* 2003;278:20475–9.
12. Takai H, Naka K, Okada Y, et al. Chk2-deficient mice exhibit radioresistance and defective p53-mediated transcription. *EMBO J* 2002;21:5195–205.
13. Hirao A, Kong YY, Matsuoka S, et al. DNA damage-induced activation of p53 by the checkpoint kinase Chk2. *Science* 2000;287:1824–7.
14. Cao L, Kim S, Xiao CY, et al. ATM-CHK2-p53 activation prevents tumorigenesis at an expense of organ homeostasis upon Brca1 deficiency. *EMBO J* 2006;25:2167–77.
15. Weischer M, Bojesen SE, Ellervik C, et al. CHEK2*1100delC genotyping for clinical assessment of breast cancer risk: meta-analyses of 26,000 patient cases and 27,000 controls. *J Clin Oncol* 2008;26:542–8.
16. Liu C, Wang Y, Wang QS, et al. The CHEK2 I157T variant and breast cancer susceptibility: a systematic review and meta-analysis. *Asian Pac J Cancer Prev* 2012;13:1355–60.
17. Walsh T, Casadei S, Coats KH, et al. Spectrum of mutations in BRCA1, BRCA2, CHEK2, and TP53 in families at high risk of breast cancer. *JAMA* 2006;295:1379–88.
18. Walsh T, Mandell JB, Norquist BM, et al. Genetic predisposition to breast cancer due to mutations other than BRCA1 and BRCA2 founder alleles among Ashkenazi Jewish women. *JAMA Oncol* 2017;3:1647–53.
19. Desrichard A, Bidet Y, Uhrhammer N, et al. CHEK2 contribution to hereditary breast cancer in non-BRCA families. *Breast Cancer Res* 2011;13:R119.
20. Le Calvez-Kelm F, Lesueur F, Damiola F, et al. Rare, evolutionarily unlikely missense substitutions in CHEK2 contribute to breast cancer susceptibility: results from a breast cancer family registry case-control mutation-screening study. *Breast Cancer Res* 2011;13:R6.
21. Dufault MR, Betz B, Wappenschmidt B, et al. Limited relevance of the CHEK2 gene in hereditary breast cancer. *Int J Cancer* 2004;110:320–5.
22. Bell DW, Kim SH, Godwin AK, et al. Genetic and functional analysis of CHEK2 (CHK2) variants in multiethnic cohorts. *Int J Cancer* 2007;121: 2661–7.
23. Leedom TP, LaDuca H, McFarland R, et al. Breast cancer risk is similar for CHEK2 founder and non-founder mutation carriers. *Cancer Genet* 2016;209:403–7.
24. Couch FJ, Shimelis H, Hu C, et al. Associations between cancer predisposition testing panel genes and breast cancer. *JAMA Oncol* 2017;3:1190–6.
25. Fan Z, Ouyang T, Li J, et al. Identification and analysis of CHEK2 germline mutations in Chinese BRCA1/2-negative breast cancer patients. *Breast Cancer Res Treat* 2018;169:59–67.
26. Hauke J, Horvath J, Gross E, et al. Gene panel testing of 5589 BRCA1/2-negative index patients with breast cancer in a routine diagnostic setting: results of the German consortium for hereditary breast and ovarian cancer. *Cancer Med* 2018;7: 1349–58.
27. Young EL, Feng BJ, Stark AW, et al. Multigene testing of moderate-risk genes: be mindful of the missense. *J Med Genet* 2016;53:366–76.
28. Espenschied C, Kleiblova P, Richardson M, et al. Classifying variants in the CHEK2 gene: the importance of collaboration. *Eur J Cancer* 2017; 72:S25.
29. Baysal BE, DeLoia JA, Willett-Brozick JE, et al. Analysis of CHEK2 gene for ovarian cancer susceptibility. *Gynecol Oncol* 2004;95:62–9.
30. Suspitsin EN, Sherina NY, Ponomariova DN, et al. High frequency of BRCA1, but not CHEK2 or NBS1 (NBN), founder mutations in Russian ovarian cancer patients. *Hered Cancer Clin Pract* 2009;7:5.
31. Carter NJ, Marshall ML, Susswein LR, et al. Germline pathogenic variants identified in women with ovarian tumors. *Gynecol Oncol* 2018;151:481–8.
32. Kurian AW, Ward KC, Howlader N, et al. Genetic testing and results in a population-based cohort of breast cancer patients and ovarian cancer patients. *J Clin Oncol* 2019. https://doi.org/10.1200/JCO.18.01854.
33. Havranek O, Kleiblova P, Hojny J, et al. Association of Germline CHEK2 gene variants with risk and prognosis of non-Hodgkin lymphoma. *PLoS One* 2015;10:e0140819. [Epub ahead of print].
34. Soukupova J, Zemankova P, Lhotova K, et al. Validation of CZECANCA (CZEch CAncer paNel for clinical application) for targeted NGS-based analysis of hereditary cancer syndromes. *PLoS One* 2018;13:e0195761. [Epub ahead of print].
35. Roeb W, Higgins J, King MC. Response to DNA damage of CHEK2 missense mutations in familial breast cancer. *Hum Mol Genet* 2012;21:2738–44.
36. Delimitsou A, Fostira F, Kalfakakou D, et al. Functional characterization of CHEK2 variants in a *Saccharomyces cerevisiae* system. *Hum Mutat* 2019;40:631–48. https://doi.org/10.1002/humu.23728.
37. Ahn JY, Li X, Davis HL, et al. Phosphorylation of threonine 68 promotes oligomerization and autophosphorylation of the Chk2 protein kinase via the forkhead-associated domain. *J Biol Chem* 2002;277:19389–95.
38. Chrisanthar R, Knappskog S, Lokkevik E, et al. CHEK2 mutations affecting kinase activity together with mutations in TP53 indicate a functional pathway associated with resistance to epirubicin in primary breast cancer. *PLoS One* 2008; 3:e3062.

39. Mundt E, Nix P, Bowles KR, et al. *Complexities in hereditary cancer variant classification: three case examples.* ACMG Annual Clinical Genetics Meeting March 21–25, 2017, Poster#154, Phoenix Convention Center Phoenix, Arizona.
40. Wu X, Dong X, Liu W, et al. Characterization of CHEK2 mutations in prostate cancer. *Hum Mutat* 2006;27:742–7.
41. Southey MC, Goldgar DE, Winqvist R, et al. PALB2, CHEK2 and ATM rare variants and cancer risk: data from COGS. *J Med Genet* 2016;53: 800–11.
42. Nurmi A, Muranen TA, Pelttari LM, et al. Recurrent moderate-risk mutations in Finnish breast and ovarian cancer patients. *Int J Cancer* 2019. https://doi.org/10.1002/ijc.32309.
43. Girard E, Eon-Marchais S, Olaso R, et al. Familial breast cancer and DNA repair genes: insights into known and novel susceptibility genes from the GENESIS study, and implications for multigene panel testing. *Int J Cancer* 2019;144:1962–74.
44. Decker B, Allen J, Luccarini C, et al. Rare, protein-truncating variants in ATM, CHEK2 and PALB2, but not XRCC2, are associated with increased breast cancer risks. *J Med Genet* 2017; 54:732–41.
45. Slavin TP, Maxwell KN, Lilyquist J, et al. The contribution of pathogenic variants in breast cancer susceptibility genes to familial breast cancer risk. *NPJ Breast Cancer* 2017;3:22.
46. Schmidt MK, Hogervorst F, van Hien R, et al. Age- and tumor subtype-specific breast cancer risk estimates for CHEK2*1100delC carriers. *J Clin Oncol* 2016;34:2750–60.
47. Cybulski C, Wokolorczyk D, Jakubowska A, et al. Risk of breast cancer in women with a CHEK2 mutation with and without a family history of breast cancer. *J Clin Oncol* 2011;29:3747–52.
48. Weischer M, Bojesen SE, Tybjaerg-Hansen A, et al. Increased risk of breast cancer associated with CHEK2*1100delC. *J Clin Oncol* 2007;25: 57–63.
49. CHEK2 Breast Cancer Case-Control Consortium. CHEK2*1100delC and susceptibility to breast cancer: a collaborative analysis involving 10,860 breast cancer cases and 9,065 controls from 10 studies. *Am J Hum Genet* 2004;74:1175–82.
50. Vahteristo P, Bartkova J, Eerola H, et al. A CHEK2 genetic variant contributing to a substantial fraction of familial breast cancer. *Am J Hum Genet* 2002;71:432–8.
51. Liang M, Zhang Y, Sun C, et al. Association between CHEK2*1100delC and breast cancer: a systematic review and meta-analysis. *Mol Diagn Ther* 2018;22:397–407.
52. Zhang B, Beeghly-Fadiel A, Long J, et al. Genetic variants associated with breast-cancer risk: comprehensive research synopsis, meta-analysis, and epidemiological evidence. *Lancet Oncol* 2011;12: 477–88.
53. Hallamies S, Pelttari LM, Poikonen-Saksela P, et al. CHEK2 c.1100delC mutation is associated with an increased risk for male breast cancer in Finnish patient population. *BMC Cancer* 2017;17:620.
54. Pritzlaff M, Summerour P, McFarland R, et al. Male breast cancer in a multi-gene panel testing cohort: insights and unexpected results. *Breast Cancer Res Treat* 2017;161:575–86.
55. Schmidt MK, Tollenaar RA, de Kemp SR, et al. Breast cancer survival and tumor characteristics in premenopausal women carrying the CHEK2*1100delC germline mutation. *J Clin Oncol* 2007;25:64–9.
56. Muranen TA, Greco D, Blomqvist C, et al. Genetic modifiers of CHEK2*1100delC-associated breast cancer risk. *Genet Med* 2017;19:599–603.
57. Kleibl Z, Havranek O, Novotny J, et al. Analysis of CHEK2 FHA domain in Czech patients with sporadic breast cancer revealed distinct rare genetic alterations. *Breast Cancer Res Treat* 2008; 112:159–64.
58. Muranen TA, Blomqvist C, Dork T, et al. Patient survival and tumor characteristics associated with CHEK2:p.I157T—findings from the breast cancer association consortium. *Breast Cancer Res* 2016; 18:98.
59. Huzarski T, Cybulski C, Domagala W, et al. Pathology of breast cancer in women with constitutional CHEK2 mutations. *Breast Cancer Res Treat* 2005;90:187–9.

Příloha 3 - Soukupova, J., P. Zemankova, K. Lhotova, M. Janatova, M. Borecka, L. Stolarova, F. Lhota, L. Foretova, E. Machackova, V. Stranecky, S. Tavandzis, P. Kleiblova, M. Vocka, H. Hartmannova, K. Hodanova, S. Kmoch, and Z. Kleibl. 2018. "***Validation of CZECANCA (CZEch CAncer paNel for Clinical Application) for targeted NGS-based analysis of hereditary cancer syndromes***." *PLoS One* 13 (4):e0195761. doi: 10.1371/journal.pone.0195761.

RESEARCH ARTICLE

Validation of CZECANCA (CZEch CAncer paNel for Clinical Application) for targeted NGS-based analysis of hereditary cancer syndromes

Jana Soukupova[1]*, Petra Zemankova[1], Klara Lhotova[1], Marketa Janatova[1], Marianna Borecka[1], Lenka Stolarova[1], Filip Lhota[1,2], Lenka Foretova[3], Eva Machackova[3], Viktor Stranecky[4], Spiros Tavandzis[5], Petra Kleiblova[1,6], Michal Vocka[7], Hana Hartmannova[4], Katerina Hodanova[4], Stanislav Kmoch[4], Zdenek Kleibl[1]*

1 Institute of Biochemistry and Experimental Oncology, First Faculty of Medicine, Charles University, Prague, Czech Republic, **2** Centre for Medical Genetics and Reproductive Medicine, Gennet, Prague, Czech Republic, **3** Department of Cancer Epidemiology and Genetics, Masaryk Memorial Cancer Institute, Brno, Czech Republic, **4** Research Unit for Rare Diseases, Department of Paediatrics and Adolescent Medicine, First Faculty of Medicine, Charles University and General University Hospital in Prague, Prague, Czech Republic, **5** Department of Medical Genetics, AGEL Laboratories, AGEL Research and Training Institute, Novy Jicin, Czech Republic, **6** Institute of Biology and Medical Genetics, First Faculty of Medicine, Charles University and General University Hospital in Prague, Prague, Czech Republic, **7** Department of Oncology, First Faculty of Medicine, Charles University and General University Hospital in Prague, Prague, Czech Republic

* zdekleje@lf1.cuni.cz (ZK); jana.soukupova@lf1.cuni.cz (JS)

Abstract

Background

Carriers of mutations in hereditary cancer predisposition genes represent a small but clinically important subgroup of oncology patients. The identification of causal germline mutations determines follow-up management, treatment options and genetic counselling in patients' families. Targeted next-generation sequencing-based analyses using cancer-specific panels in high-risk individuals have been rapidly adopted by diagnostic laboratories. While the use of diagnosis-specific panels is straightforward in typical cases, individuals with unusual phenotypes from families with overlapping criteria require multiple panel testing. Moreover, narrow gene panels are limited by our currently incomplete knowledge about possible genetic dispositions.

Methods

We have designed a multi-gene panel called CZECANCA (CZEch CAncer paNel for Clinical Application) for a sequencing analysis of 219 cancer-susceptibility and candidate predisposition genes associated with frequent hereditary cancers.

Results

The bioanalytical and bioinformatics pipeline was validated on a set of internal and commercially available DNA controls showing high coverage uniformity, sensitivity, specificity and

OPEN ACCESS

Citation: Soukupova J, Zemankova P, Lhotova K, Janatova M, Borecka M, Stolarova L, et al. (2018) Validation of CZECANCA (CZEch CAncer paNel for Clinical Application) for targeted NGS-based analysis of hereditary cancer syndromes. PLoS ONE 13(4): e0195761. https://doi.org/10.1371/journal.pone.0195761

Editor: Obul Reddy Bandapalli, German Cancer Research Center (DKFZ), GERMANY

Received: October 19, 2017

Accepted: March 28, 2018

Published: April 12, 2018

Copyright: © 2018 Soukupova et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This study was supported by the grants of Ministry of Health (www.mzcr.cz) 15-27695A (JS), 15-28830A (ZK), 16-29959A (ZK), and DRO MMCI, 00209805 (LF), grants of Charles University (www.cuni.cz) PROGRES Q28/LF1 (ZK), and SVV2017/260367 (SK) and Ministry of Education,

Youth and Sports (www.msmt.cz) project CZ.02.1.01/0.0/0.0/16_013/0001634 (SK).

Competing interests: The authors have declared that no competing interests exist.

accuracy. The panel demonstrates a reliable detection of both single nucleotide and copy number variants. Inter-laboratory, intra- and inter-run replicates confirmed the robustness of our approach.

Conclusion

The objective of CZECANCA is a nationwide consolidation of cancer-predisposition genetic testing across various clinical indications with savings in costs, human labor and turnaround time. Moreover, the unified diagnostics will enable the integration and analysis of genotypes with associated phenotypes in a national database improving the clinical interpretation of variants.

Introduction

Hereditary cancer syndromes are heterogeneous diseases characterized by the development of various cancer types in carriers of rare germline mutations in cancer susceptibility genes. These genes dominantly code for tumor suppressor proteins negatively regulating mitotic signals and cell cycle progression, activating apoptotic pathways, or executing DNA repair processes [1].

In general, it is considered that around 5% of all cancer diagnoses arise in hereditary cancer form. However, the percentage of hereditary cancers varies by cancer type, ranging from less than 3% in lung cancer to over 30% in pheochromocytoma [2, 3]. Important features distinguishing hereditary and sporadic cancers include an increased lifetime cancer risk with early disease onset, an increased risk of cancer multiplicity, the accumulation of cancer diagnoses in affected families, and a 50% risk of disease trait transmission to the offspring [1]. Considering these attributes and their consequences in terms of decreased life expectancy, decreased quality of life and increased medical expenses, patients carrying mutations in cancer susceptibility genes and their relatives represent a medically important subgroup with specific needs for increased cancer surveillance, a tailored follow-up and therapy, and rational prevention. However, the primary need is an unequivocal identification of the causative germline variant.

Although cancer inheritance has been suggested for over 150 years, the first gene conferring an increased cancer risk (*Rb*) was discovered only 30 years ago [4]. Hundreds of predisposing or candidate genes have been characterized since then, including the clinically most important "major" cancer susceptibility genes with high penetrance representing a subset of genes whose germline variants confer a high cancer risk (with relative risk (RR) > 5.0) in a substantial proportion of hereditary cancer patients. Pathogenic germline variants in "major" genes occur most commonly in patients with breast, ovarian, and colorectal cancers with variable proportions across populations worldwide. The group of cancer susceptibility genes with moderate penetrance is more extensive and growing steadily [5]. However, the clinical utility for many moderate penetrance genes is currently limited by the insufficient evidence about the degree of cancer risks associated with their germline variants.

The rapid improvement and availability of next-generation sequencing (NGS) technologies enable efficient simultaneous analyses of many cancer susceptibility genes in oncology patients or asymptomatic individuals at risk in routine diagnostics. NGS offers multiple approaches for the investigation of cancer predisposition, including the sequencing of whole genomes, exomes or transcriptomes. At present, however, the most widely used method of detecting clinically informative genetic alterations in the clinical setting is targeted panel NGS, analyzing selected

subsets of genes of interest [6]. Nevertheless, the numbers of genes included in panels differ substantially among laboratories and depend on healthcare systems. While some cancer-specific or multi-cancer panels include only the "major" predisposition genes for which substantial literature exists with regard to their diagnostic relevance, others include larger gene sets consisting of all clinically relevant genes and additional genes for which the evidence of cancer predisposition is still unclear.

NGS-based cancer testing has been rapidly adopted by routine clinical laboratories [7]. Their primary choice resides in the decision whether to use a commercially available NGS panel, or to design custom-made systems. The decision is influenced by clinical demand determining the set of targeted genes, by the spectrum of cancer diagnoses that will be analyzed, by the expected number of analyzed samples, and by costs of the analyses.

Our aim was to develop a universal diagnostic approach suitable for contributing genetic laboratories and allowing sample batching across multiple cancer indications. We focused on i) designing a custom-made multi-cancer panel with the desired sequencing quality and uniformity permitting a reliable variant identification, ii) the development of a robust analytical procedure limiting inter-run and inter-laboratory differences, and iii) the optimization of the bioinformatics pipeline enabling unified variant calling and annotation. The data collected from analyses of high-risk individuals performed in contributing laboratories will be used to create a nationwide genotype–phenotype database improving clinical variant interpretation in high-risk individuals.

Methods

Validation samples

Patient DNA samples. Validation of CZECANCA pipeline included analyses of 389 samples previously tested for the presence of germline variants available from DNA repository of the Institute of Biochemistry and Experimental Oncology. First Faculty of Medicine, Charles University. Of these, 137 samples carried pathogenic SNVs or short indels (in *BRCA1/2*, *PALB2*, *CHEK2*, *ATM*, *NBN*, *DPYD*, *PPM1D*, *RAD51C*, *RAD51D*, or *TP53*), 217 had been tested negatively using previous gene-by-gene analyses based on Sanger sequencing or a protein truncation test (PTT) [8–16], and 35 samples carried intragenic rearrangements in *BRCA1*, *CHEK2*, *PALB2*, or *TP53*, identified by the MLPA (multiplex ligation-dependent probe amplification) analysis [10, 17, 18]. All blood-isolated DNA samples were obtained from individuals that gave their written informed consent with mutation analyses of cancer susceptibility genes and who agreed to use their genetic material for research purposes. The study was approved by Ethics Committee of the First Medical Faculty, Charles University and General University Hospital in Prague. All used samples were anonymized prior analysis.

Human genome reference standards. Five commercially available DNA reference standards (NA12878, NA24149, NA24385, NA24631 and NA24143) were obtained from Coriell Institute for Medical Research. Well described genotypes, including high confident calls for variant and wild-type alleles, is the major advantage of these reference standards. The genotypes and variants in reference samples identified by CZECANCA analysis and obtained from reference variant-call format (VCF) files (available from the Genome in a Bottle (GIAB) website; http://jimb.stanford.edu/giab/), respectively, were compared to compute CZECANCA sensitivity, specificity, and accuracy, as described by Hardwick et al. [19].

Panel design

The multi-cancer panel CZECANCA was designed using the online NimbleDesign software utility (NimbleGen, Roche; http://sequencing.roche.com/products/software/nimbledesign-

software.html). For enrichment, we selected genes with a known predisposition for hereditary breast, ovarian, colorectal, pancreatic, gastric, endometrial, kidney, prostate and skin cancers, together with known DNA repair genes associated (or potentially associated) with cancer susceptibility (a list of 219 selected genes is provided in S1 Table), considering the results of our previous NGS analysis with a broad panel of 581 genes [20]. The primary gene target for probe coverage was represented by all exons (in case of known cancer susceptibility genes) or all coding exons (in other genes), including 10 bases from adjacent intronic regions. The design considered all transcription variants of selected genes available at UCSC website (https://genome.ucsc.edu/; accessed 2015-05-21). The promoter regions of the *BRCA1* and *BRCA2* genes were included into the primary target. The probes were designed using *continuous design* under strict conditions–minimal and maximal *close matches* (number of times in which a probe sequence matches the genome with either $\leq$ 5 insertions or deletions, or gap of $\leq$ 5 bp) were one and three, respectively, allowing us to hybridize the probes up to three targets across the genome. Because of the strict design conditions, some clinically relevant regions were left untargeted for technical reasons such as repeats and homologous regions (see S1 Table). The final panel target size reached 628,069 bases.

Library preparation

Five hundred ng of genomic DNA isolated from peripheral blood and dissolved in TE buffer was used for preferred ultrasound shearing using Covaris E220 (Covaris Inc). As an alternative DNA fragmentation method, we tested enzymatic digestion using Fragmentase (KAPA Biosystems, Roche) with incubation for 25 min at 37˚C according to the manufacturer's instruction. The mean average size of DNA fragments targeted 200 bp. Sizing and quality was controlled using the Agilent High Sensitivity DNA kit on the Agilent 2100 Bioanalyzer System (Agilent).

Libraries were prepared using the KAPA HTP Library Preparation kit (for ultrasound-sheared DNA samples) or KAPA HyperPlus Kit (for Fragmentase-digested DNA samples) according to the manufacturer's instructions (KAPA Biosystems, Roche) with minor modifications including the use of universal in-house prepared adapters, double-indexing primers for ligation-mediated polymerase chain reaction (LM-PCR), and primers for post-capture PCR, as described further. The adapters [Adapter#1: `5′-ACACTCTTTCCCTACACGACGCTCTTCCGATC*T-3′` ("*" denotes for phosphothiolate bond) and Adapter#2: `5′-pGATCGGAAGAGCACACGTCTGAACTCCAGTCAC-3′` ("p" denotes for 5' phosphate)] were hybridized in Tris:NaCl buffer mix (50 mM Tris:HCl pH 7.5; 50 mM NaCl) in 97˚C for 2 min, followed by 72 cycles involving incubation at 97˚C for 1 min (-1˚C per cycle) and 25˚C for 5min. The barcoding of size-selected DNA fragments enabling subsequent sample pooling was performed during LM-PCR with indexing primers [Primer#1: `5′-AATGATACGGCGACCACCGAGATCTACACxxxxxxxxACACTCTTTCCCTACACGACGCTCTTCCGATC*T-3′` and Primer#2: `5′-CAAGCAGAAGACGGCATACGAGATxxxxxxxxGTGACTGGAGTTCAGACGTGTGCTCTTCCGAT*C-3′` ("*" denotes for phosphothiolate bond; "xxxxxxxx" denotes for a sequence of particular indices same as the Illumina Truseq HT index i7 and i5)]. The number of LM-PCR cycles was reduced to six to limit the presence of PCR duplicates. Sizing and quality after the double-sided size selection and LM-PCR were controlled using the Agilent High Sensitivity DNA kit on the Agilent 2100 Bioanalyzer System.

To reach the targeted mean coverage (100X), 30 individual barcoded samples (33 ng each) were pooled for the enrichment (usually two overnight hybridizations; tested for 16–72 hours without a significant effect on enrichment efficacy) using the CZECANCA (NimbleGen SeqCap EZ Choice, Roche) to create a sequencing library. After the enrichment, the library was amplified using Primer 1: `5′-AATGATACGGCGACCACCGAGATCTACAC-3′` and Primer 2:

5′-CAAGCAGAAGACGGCATACGAGAT-3′. The number of post-capture PCR cycles was reduced to 11 to reach the optimal library concentration (2 ng/μl) and to minimalize the number of PCR duplicates.

After the enrichment control using qPCR (NimbleGen SeqCap EZ Library SR User's Guide), the final 18 pM libraries were sequenced on the MiSeq system using MiSeq Reagent Kit v3, 150 cycles (Illumina).

Bioinformatics

Single nucleotide variants (SNVs). The NGS data obtained from sequencing with the CZECANCA were processed using an analysis pipeline based on standard tools. FASTQ files were generated by MiSeq. The quality of raw data was controlled using FastQC v0.11.2 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). FASTQ files were subsequently mapped using Novoalign v2.08.03 to hg19 (http://www.novocraft.com/products/novoalign/) to generate sequence alignment map (SAM) files. SAM files were transformed to binary form (BAM files) using Picard tools v1.129 (https://broadinstitute.github.io/picard/). Raw BAM files were further processed to eliminate PCR duplicates of mapped reads. The quality of mapped bases was checked and recalibrated according to default settings using Genome Analysis Toolkit (GATK) v3.3 (https://software.broadinstitute.org/gatk/). The finalized BAM file was converted using a GATK pipeline to a variant-call format (VCF) containing alternative variants only. ANNOVAR was used to annotate VCF files generated using GATK [21, 22] and to check the presence of each variant in external databases (ExAC, 1000Genome or ClinVar) [23–25]. Predictive values from selected prediction algorithms (for example SIFT [26], Mutation Analyzer [27], MutationTaster [28], LRT [29], PolyPhen-2 [30], phyloP [31], GERP [32], CADD [33] or spidex (https://www.deepgenomics.com/spidex) were added to the annotated alternative variants.

For a comparison with CZECANCA sequencing, the data from routine analyses using the TruSight cancer panel (Illumina), performed in a laboratory of the Masaryk Memorial Cancer Institute in Brno were analyzed by an identical bioinformatics pipeline [34].

The Integrative Genomics Viewer (IGV) was used for visualization and manual inspection of individual BAM files [35].

Medium-size indels. The detection and exact sequence determination of medium-size insertions and tandem duplications (involving approximately half of the sequence reads, depending on the sequencing chemistry used) is very challenging. The identification of these alterations was based on the method of soft-clipped bases using Pindel (http://gmt.genome.wustl.edu/packages/pindel/) [36]. The finalized BAM files served as an input for the analysis. In our case (with mean read size of 75 bp; MiSeq Reagent Kit v3, 150 cycles chemistry) insertion or duplication exceeding 35 bp was considered as a medium-size indel.

Copy number variations (CNVs). An analysis CNVs was performed using the CNVkit (https://pypi.python.org/pypi/CNVkit). The CNVs analysis is coverage-based and therefore required good coverage uniformity. Raw BAM files served as the input for this analysis.

Coverage visualization. The visualization of sequence coverage of the individual samples, enabling a fast visual inspection of coverage limit >20X (for a reliable identification of heterozygotes) across the analyzed genes, was performed by an in-house "Boudalyzer" script written in R language. The coverage is visualized from the finalized BAM files. This tool was used for the generation of manuscript figures showing coverages of the analyzed genes.

Variant interpretation. We used the scoring scheme outlined in ENIGMA guidelines (https://enigmaconsortium.org/) for variant interpretation to classify SNVs and indels as benign (Class 1), likely benign (Class 2), variant of unknown significance (Class 3), likely pathogenic (Class 4) and pathogenic (Class 5) [37]. Identified variants of unknown significance

(VUS) were further prioritized if their minor allele frequency was lower than 1% in ExAC, 1000Genome databases, or in a two sets of population-matched controls containing anonymized genomic data from 530 non-cancer controls analyzed by CZECANCA NGS and from 780 unselected Czech individuals analyzed by an exome sequencing (provided by the National Center for Medical Genomics; http://ncmg.cz). Potentially deleterious VUSes were selected based on concordant results obtained from above-mentioned *in silico* prediction algorithms. These priorized VUS variants were enrolled into the list of variants for subsequent segregation analyses or functional *in vitro* testing performed in selected genes.

The CZECANCA contains 22 genes that are listed in the ACMG recommendation (S1 Table) for the reporting of secondary findings [38].

Results

Target gene coverage

The NGS analysis with CZECANCA targeting the coding sequences of 219 genes (S1 Table) displayed high coverage uniformity. Under standard conditions for routine analyses, we targeted sequencing coverage 100X. In these settings, more than 85% of the targeted regions were covered 100X, 98% of the targeted regions were covered at least 50X and less than 0.2% of targeted regions had coverage below 20X (Fig 1A). The entire coding sequence was fully covered at least 100X in 144/219 targeted genes (65.8%), at least 50X in 190/219 genes (86.8%), and at least 20X in 207/219 targeted genes (94.5%; Fig 2). Coverage did not exceed 300X in any of the captured targets.

Coverage was uniform among samples independently analyzed in the participating laboratories using the described protocol (Fig 3), and also among samples sequenced using separately-synthesized CZECANCA lots (data not shown). The equal coverage uniformity was independent of coverage depth (Fig 1B). The coverage uniformity was partially influenced by the DNA fragmentation approach with better results obtained by ultrasound fragmentation in comparison with enzymatic DNA cleavage. The improved results (more random DNA shearing) obtained with the ultrasound fragmentation protocol were indicated by an analysis of terminal (di)nucleotides in reads from samples prepared by both DNA fragmentation methods, regardless of the laboratory site (Figs 1C and 3). The CZECANCA coverage uniformity substantially surpassed that of the Illumina TruSight Cancer Panel (Fig 3F).

Low-covered regions (uncovered or with coverage $\leq$20X) were constantly observed in 12/219 genes (5.5%; Fig 2, S1 Table). In nine genes, the low–covered regions were mostly limited to a single exon (typically the first exon) representing usually a small fraction of the coding sequence. In three incompletely covered genes (*CHEK2*, *MDC1*, *NF1*), single or several exons were omitted from the CZECANCA design (see Panel design in Methods). The remaining low-covered regions were GC-rich regions with mean GC content of 76.88% (S2 Table) while the average GC content of the CZECANCA targets is 47%.

Sequencing quality was partially influenced by the particular MiSeq sequencer. In standard runs, more than 99% of bases reached a Phred score >20 (i.e. 99% accuracy) and approximately 97% of bases overcame a Phred score of 30 (i.e. 99.9% accuracy). A decrease in PCRs cycles during library preparation reduced the number of PCR duplicates, which finally represented 7–9% of reads. The mean off-target (reads mapped to distance exceeding 250 bp from the nearest bait) across the performed runs was constantly less than 12% of reads.

Reproducibility, specificity and sensitivity analysis

The reproducibility of variant calls was tested using intra-, inter-run, and inter-laboratory replicates. During the sequencing of intra-run replicates, we also evaluated the impact of coverage depth on coverage uniformity and reproducibility.

Fig 1. Coverage parameters from CZECANCA sequencing. (A) The chart expresses the percentages of covered target bases (cov. b.) obtained from 25 analyzed samples from a standard run targeting sequencing coverage 100X. (B) The coverage (at y-axis) of *BRCA1* coding sequence (NM_007294; x-axis; vertical lines represent exon boundaries) in three independent runs targeting sequencing coverages 20X, 100X, or 500X demonstrates coverage uniformity, not influenced by coverage depth. (C) The "randomness" of the DNA shearing approach using ultrasound (US) and enzymatic cleavage was compared by an analysis of the distribution of ending nucleotides and dinucleotides in reads completely mapped to the large exon 11 (chr17:41243452–41246877; 3426bp) in the *BRCA1* gene, representing one of the largest continuous genomic fragments targeted by CZECANCA probes. The chart displays the relativized distribution of terminal nucleotides and dinucleotides in the analyzed region from 12 samples from each laboratory normalized to the average nucleotide and dinucleotide content of the analyzed region. The distribution of last nucleotides and dinucleotides in fragments from samples processed by US oscillate closer to a normalized value (1) than in fragments of samples prepared by the enzymatic cleavage.

https://doi.org/10.1371/journal.pone.0195761.g001

Three individually bar-coded replicates were pooled for enrichment in amounts corresponding to 33 ng (considered as 100%), 24.75 ng (75%), and 16.5 ng (50%), respectively. The subsequent bioinformatics of these samples, considering variants with GATK quality >100 in the targeted regions (exon sequences with 12 bp from adjacent introns), revealed 293 (100%), 292 (99.7%) and 290 (99.0%) variants, respectively (S3 Table). Altogether, 289/293 (98.6%) variants were identified in all replicates, while four variants not detected in DNA-reduced samples were variant homozygotes located in low-covered regions or had GATK quality <100. The

AIP ALK APC APEX1 ATM ATMIN ATR ATRIP AURKA AXIN1 BABAM1 BAP1
BARD1 BML BMPR1A BRAP BRCA1 BRCA2 BRCC3 BRE BRIP1 BUB1B C11orf30 C11orf40
CASP8 CCND1 CDC73 CDH1 CDK4 CDKN1B CDKN1C CDKN2A CEBPA CEP57 CLSPN CSNK1D
CSNK1E CWF19L2 CYLD DCLRE1C DDB2 DHFR DICER1 DMC1 DNAJC21 DPYD EGFR EPCAM
EPHX1 ERCC1 ERCC2 ERCC3 ERCC4 ERCC5 ERCC6 ESR1 ESR2 EXO1 EXT1 EXT2
EYA2 EZH2 FAM175A FAM175B FAN1 FANCA FANCB FANCC FANCD2 FANCE FANCF FANCG
FANCI FANCL FANCM FBXW7 FH FLCN GADD45 GATA2 GPC3 GRB7 HELQ HNF1A
HOXB13 HRAS HUS1 CHEK1 CHEK2 KAT5 KCNJ5 KIT LIG1 LIG3 LIG4 LMO1
LRIG1 MAX MCPH1 MDC1 MDM2 MDM4 MEN1 MET MGMT MLH1 MLH3 MMP8
MPL MRE11A MSH2 MSH3 MSH5 MSH6 MSR1 MUS81 MUTYH NAT1 NBN NCAM1
NELFB NF1 NF2 NFKBIZ NHEJ1 NSD1 OGG1 PALB2 PARP1 PCNA PHB PHOX2B
PIK3CG PLA2G2A PMS1 POLB POLD1 POLE PPM1D PREX2 PRF1 PRKAR1A PRKDC PTEN
PTCH1 PTTG2 RAD1 RAD9A RAD17 RAD18 RAD23B RAD50 RAD51 RAD51AP1 RAD51B RAD51C
RAD51D RAD52 RAD54B RAD54L RB1 RBBP8 RECQL RECQL4 RECQL5 RET RFC1 RFC2
RFC4 RHBDF2 RNF8 RNF146 RNF168 RPA1 RUNX1 SDHAF2 SDHB SETBP1 SETX SHPRH
SLX4 SMAD4 SMARCA4 SMARCB1 SMARCE1 STK11 SUFU TCL1A TELO2 TERF2 TERT TLR2
TLR4 TMEM127 TOPBP1 TP53 TP53BP1 TSC1 TSC2 TSHR UBE2A UBE2B UBE2I UBE2V2
UBE4B UIMC1 VHL WRN WT1 XPA XPC XRCC1 XRCC2 XRCC3 XRCC4 XRCC5
XRCC6 ZNF350 ZNF365

200X 100X 20X

Fig 2. Coverage (y-axis) of coding sequences (x-axis) of 219 CZECANCA target genes from a routine, randomly selected run targeting 100X coverage. Note: Fully covered genes are depicted in green letters, genes with coverage <20X in a single exon are in orange letters, and genes with uncovered regions exceeding single exon or >10% of coding sequence are in red letters. Green horizontal bars (below individual graphs constructed using "Boudalyzer" script) indicate coverage ≥ 20X; red horizontal bars indicate regions covered <20X and uncovered regions.

https://doi.org/10.1371/journal.pone.0195761.g002

analysis demonstrated that alternative nucleotides could still be reliably detected in samples with reduced overall coverage, showing the robustness of the analysis in samples with unequal DNA input (Fig 4A).

A subsequent analysis of inter-run replicates (performed with another DNA sample analyzed in two independent runs) revealed 356 unique variants with GATK quality >100 in at least one replicate (S4 Table). Overall, 354 (99.4%) variants were identified in both inter-run replicates with a strong coverage correlation (Fig 4B).

In addition, the inter-laboratory performance was tested by an NGS analysis of an identical DNA control sample in four laboratories participating in the panel validation (Fig 4C), which revealed 332 unique variants with GATK quality >100 in at least one laboratory, from which we identified 331 (99.7%), 327 (98.5%), 329 (99.1%), and 329 (99.1%) variants in the particular laboratory, respectively. The discordant findings were caused by variants in low-covered regions, with low base Phred quality, or GATK quality <100 (S5 Table).

Sensitivity and specificity were assessed in 354 samples previously tested for the presence of germline variants. All 137 previously identified pathogenic germline mutations in *BRCA1/2* and other susceptibility genes were detected by CZECANCA (S6 Table). Moreover, an analysis

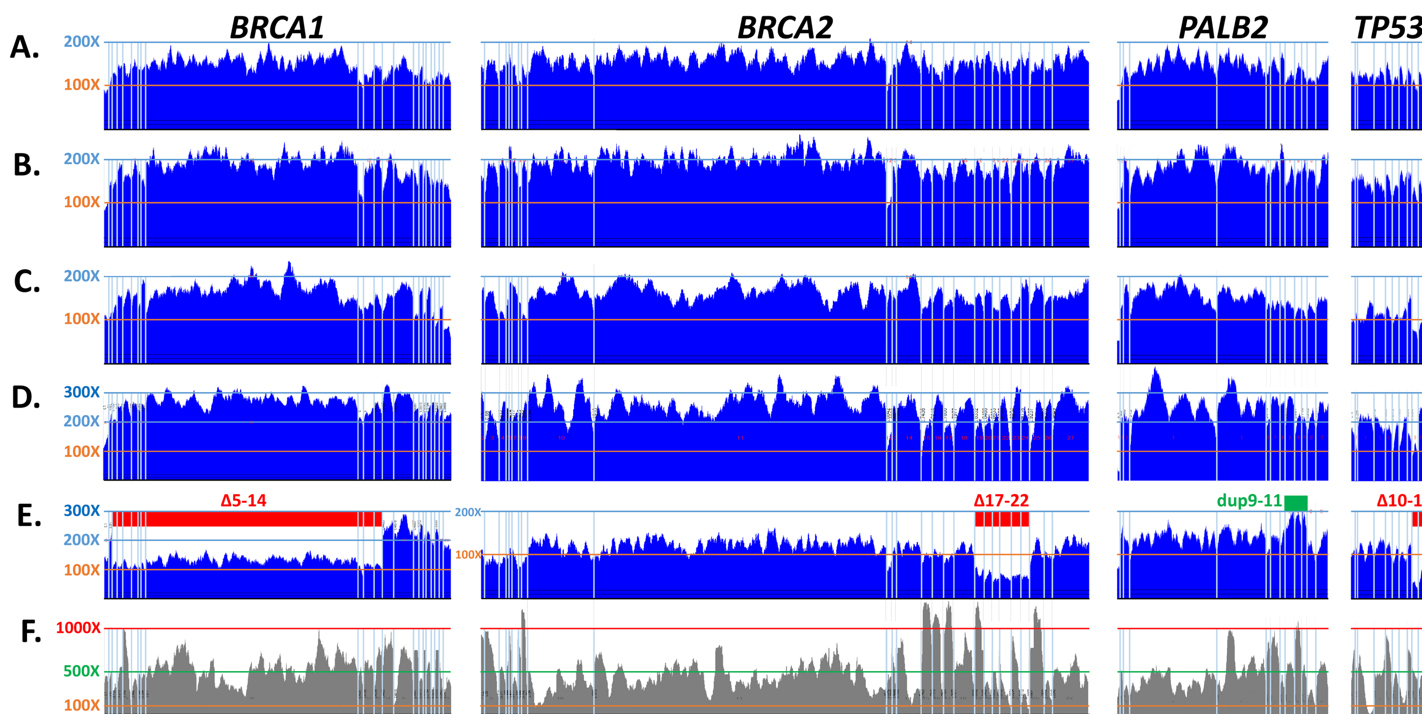


Fig 3. Coverage of selected genes from the CZECANCA (A-E) and TruSight Cancer sequencing (F) panels. The pictures show coverage (at y-axis) alongside the coding sequences of *BRCA1* (NM_007294), *BRCA2* (NM_000059), *PALB2* (NM_024675), and *TP53* (NM_000546), the vertical lines represent exon boundaries. Panels A–D show results obtained from a CZECANCA NGS analysis of various samples performed in four participating laboratories using the ultrasound (A, B) or enzymatic (C, D) DNA fragmentation protocol. Examples of the identified CNV aberrations in the depicted genes (deletions in *BRCA1*, *BRCA2* and *TP53* and duplication in *PALB2*) are shown in panel E. For comparison, panel F demonstrates the uneven coverage of the depicted genes by sequencing using the TruSight Cancer panel (Illumina).

https://doi.org/10.1371/journal.pone.0195761.g003

Fig 4. Analysis of intra-run (A), inter-run (B), and inter-laboratory (C) replicates. The panels show sequencing coverages (y-axis) of the identified variants arranged according to chromosomal localizations (x-axis). We used moving average curves (average of 3 values) to compare trends in coverages. Panel (A) describes the results of an analysis of three independently processed intra-run replicates from an identical DNA sample pooled in 33 ng (considered as 100%), 24.75 ng (75%), and 16.5 ng (50%), respectively. Panel (B) demonstrates variant coverages identified in two independent inter-run (run 8 and 14) replicates. All coverage values of sample #3647 in run 14 were corrected by a factor of 1.3880 to normalize coverages between samples (see S4 Table). Panel (C) shows coverages of variants identified in an inter-laboratory control sequenced in four laboratories (Lab) participating in panel validation (see S5 Table). The coverages of variants identified in Lab 2, 3, and 4 were normalized to the average coverage of Lab 1 for better comparisons of coverages.

https://doi.org/10.1371/journal.pone.0195761.g004

revealed nine additional *BRCA1* or *BRCA2* mutations. Of these, seven mutations were identified in samples previously tested by cDNA sequencing (they had not been detected previously, probably because of nonsense-mediated decay). The pathogenic missense mutation c.3G>A in *BRCA2* was found in a sample negatively analyzed using PTT and the pathogenic *BRCA2* mutation c.5645C>A was found in the carrier of c.5266dupC in *BRCA1* in whom the identification of a pathogenic *BRCA1* variant discontinued subsequent *BRCA2* testing.

Further, we validated the sensitivity of CNVs detection on 35 samples tested positively using the MLPA analysis (S7 Table). All CNVs including 18 samples with large *BRCA1* deletions or duplications, 12 CNVs in *CHEK2*, four in *PALB2* and one in *TP53* were detected using CNVkit software in routine settings targeting 100X coverage (Fig 5A; S8 Table). This analysis also enabled to setup CNVkit thresholds indicating the presence of a deletion or a duplication. To estimate the number of false positive and true positive CNV calls obtained from CNVkit, we further analyzed aggregated results from four consecutive runs performed in two

Fig 5. The panel A show results of CNV analysis revealing large deletions or duplications in four genes in a testing set of 35 samples with previously identified CNVs. The charts show median-normalized values of CNV scores for particular gene bins (default settings in CNVkit software; S8 Table). Values <-0.6 and >0.45 (red dotted lines) were assumed as thresholds indicating a deletion or a duplication, respectively. All shown CNVs were confirmed by MLPA previously (S7 Table). The panels B and C demonstrate frequency of true positive (TP) and false positive (FP) CNV signals from analyses performed in two participating laboratories (laboratory 1 in B and laboratory 3 in C). While 116 samples analyzed in four consecutive runs in B were prepared using the ultrasound (US) fragmentation, 125 other samples in four consecutive runs in C were prepared using the enzymatic (ENZ) fragmentation method. Samples in vivid colors highlight suspected samples that were further analyzed by MLPA analysis and samples in *BRCA1* Δ5–14 (B) and Δ8 (C) denote for true positives. The presence of putative CNVs in *PALB2*, *CHEK2*, and *TP53* were excluded by analysis that revealed heterozygotes in regions with suspected deletions or by an MLPA analysis.

https://doi.org/10.1371/journal.pone.0195761.g005

participating laboratories preparing sequencing libraries by ultrasound shearing and enzymatic digestion, respectively (Fig 5B and 5C). The CNV analysis in *BRCA1* gene revealed that two out of 116 (1.7%) ultrasound-sheared samples (from laboratory 1) and five out of other 125 (4%) enzymatically-digested samples (from laboratory 3) were scored as the samples with suspected deletion or duplication. The *BRCA1* MLPA analysis performed in all samples revealed that one suspected sample from each laboratory was true positive (exon 5–14 del in laboratory 1 and exon 8 del in laboratory 3), remaining suspected samples (one from laboratory 1 and four from laboratory 3) were false positive, and 114/116 in laboratory 1 and 120/125 in laboratory 3 were true negative *BRCA1* samples.

While the minimum coverage for a reliable detection of SNVs was estimated at 20X, the minimum coverage required for a reliable detection of CNVs is higher [39]. However, we have noticed that coverage uniformity is at least of the same importance. While the type of the DNA fragmentation protocol (ultrasound vs. enzymatic digestion) did not influence the sensitivity of SNVs detection (Fig 4C), enzymatic digestion caused difficulties in reliable CNVs detection (with an increased number of CNVkit false positives) when comparing samples with the same coverage. We suppose that the main problem of a CNVs coverage-based analysis of enzymatically fragmented samples is worse coverage uniformity caused by non-random DNA cleavage, as discussed above (Fig 1C). To evaluate the sensitivity of CNVs detection in other targeted genes and to better address the influence of DNA fragmentation protocol on the CNV analysis, we compared results of CNVkit analysis in remaining 20 ACMG genes (except *BRCA1* and *TP53* discussed above) covered by CZECANCA target (Fig 6).

The analysis revealed relative low rate of suspected CNVs (0–4 and 0–23 carriers per gene in samples prepared by ultrasound DNA fragmentation and enzymatic DNA digestion, respectively) and demonstrated that preparation of sequencing libraries using ultrasound digestion substantially decreased the need for subsequent MLPA analyses. With the exception of *BRCA2* in which MLPA analysis was performed in all suspected samples, application of MLPA analysis in remaining genes were directed by the phenotype characteristics of analyzed probands. The only CNV identified in remaining ACMG genes was exon 17 deletion in the tuberin (*TSC2*) gene in a patient with typical skin affections. The CNV analysis of the entire set of CZECANCA target genes is provided in S11 Table. The data indicate that deviations of median-normalized CNVkit values in a run of consecutive bin sets could indicate highly probable presence of a large intragenic deletion or duplication (S1 Fig). The extreme case of such situation provides the analysis of genes localized on X chromosome in male and female probands (S2 Fig) that also demonstrates the dynamic range of analysis in detection of real deletion.

For the detection of medium-size insertions and tandem duplications, we added the Pindel tool to the bioinformatics pipeline in order to identify the 64 bp tandem duplication in *BRCA1* (c.5468-11_5520dup64; NM_007294; Chr17: 41197765–41197830 on Assembly GRCh37) not detected by GATK. The sensitivity of a Pindel analysis was recently confirmed by another GATK-omitted variant, the 38 bp duplication in *CHEK2* (c.845_846+36dup38; NM_007194; Chr22: 29105958–29105995 on Assembly GRCh37), confirmed by Sanger sequencing.

Five DNA reference standards (NA12878, NA24149, NA24385, NA24631 and NA24143) with well-described genotypes were analyzed by CZECANCA pipeline to benchmark the overall workflow performance [19]. Comparison between genotypes identified in CZECANCA analysis and available as reference VCFs showed a high concordance in identification of homozygotes and heterozygotes and also high sensitivity, specificity and accuracy of CZECANCA NGS analysis (Fig 7; S9 Table). Totally, 1,722 true positive variants (332–355 per sample), 252 false positive variants (42–57 per sample), and 13 false negative variants (0–5 per sample) were scored in all analyzed DNA reference standards considering 628,069 bases of CZECANCA target region. All were localized in 84 short genomic regions that comprised in majority homopolymeric or repetitive non-coding sequences creating recurrent sequencing errors in currently used sequencing platforms, as indicated by 7/13 not identified (false negative) variants flanking to position of false positive variants. The subsequent manual IGV inspection revealed that the remaining six false negative variants (all indels) were present with allelic fraction below 15% (filtered out through the bioinformatics pipeline).

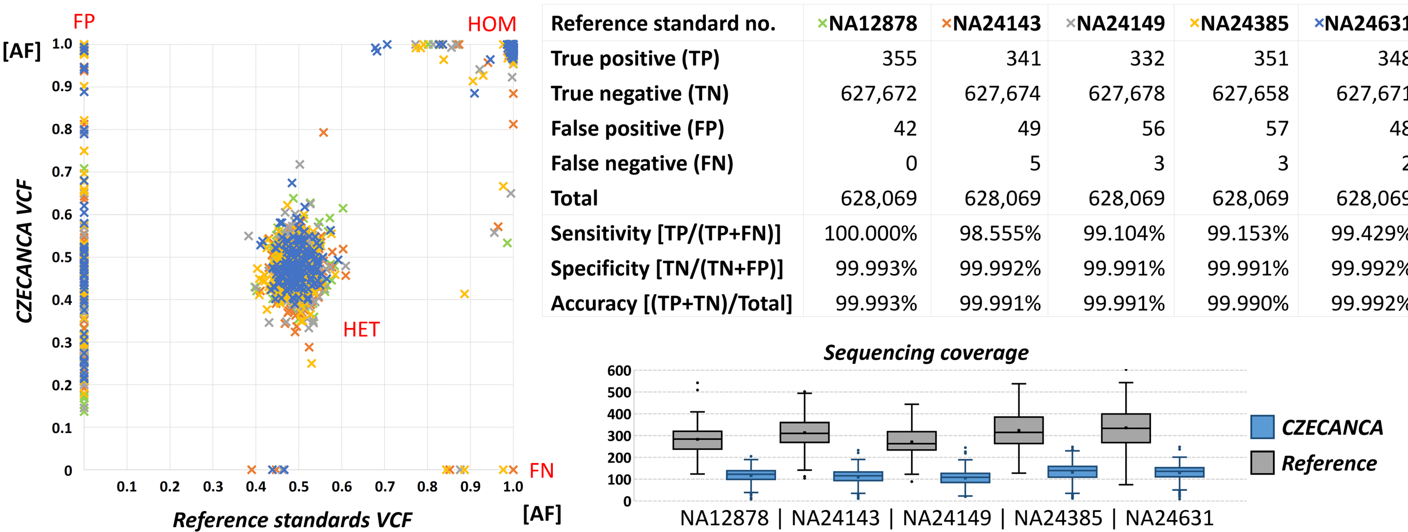
Finally, an external quality assessment of CZECANCA was performed using the pilot NGS germline mutations scheme provided by the European Molecular Genetics Quality Network (EMQN; www.emqn.org). This external quality assessment showed a 100% sensitivity of variant detection (S10 Table).

Fig 6. CNV detection is influenced by a DNA preparation method. Panels show analyses of remaining ACMG genes (not shown in Fig 5B and 5C) from four runs performed in laboratory 1 (116 DNA samples fragmented by ultrasound) and laboratory 3 (125 DNA samples fragmented enzymatically). The numbers in parentheses express number of samples with possible CNVs from all analyzed samples in contributing laboratories. *indicate samples analyzed by MLPA negatively (FP–black) or positively (TP–red). Bin set covering exon 1 in *RET* was excluded from the analysis due to the large coverage variability.

https://doi.org/10.1371/journal.pone.0195761.g006

Discussion

Multi-gene panel NGS has changed the genetic landscape for hereditary cancer syndromes. At present, clinical testing prioritizes the use of smaller cancer-specific panels, usually up to 30 cancer susceptibility genes. A large number of panels is available particularly for breast/ovarian and colorectal cancers, which represent frequent diagnoses with a high contribution of genetic components influencing the disease onset, progression and treatment outcomes [40]. Analyses



Reference standard no.	NA12878	NA24143	NA24149	NA24385	NA24631
True positive (TP)	355	341	332	351	348
True negative (TN)	627,672	627,674	627,678	627,658	627,671
False positive (FP)	42	49	56	57	48
False negative (FN)	0	5	3	3	2
Total	628,069	628,069	628,069	628,069	628,069
Sensitivity [TP/(TP+FN)]	100.000%	98.555%	99.104%	99.153%	99.429%
Specificity [TN/(TN+FP)]	99.993%	99.992%	99.991%	99.991%	99.992%
Accuracy [(TP+TN)/Total]	99.993%	99.991%	99.991%	99.990%	99.992%

Fig 7. Comparison of variant detection (shown as values of variant allelic fraction; AF) in DNA reference standards (NA12878, NA24149, NA24385, NA24631 and NA24143) obtained from CZECANCA analysis (x-axis) and AF from VCF files for these standards downloaded from http://jimb.stanford.edu/giab/ (y-axis). The graph shows all variants with GATK quality >100 reached in CZECANCA analysis (including FP variants) and undetected (FN) variants. Heterozygote variants clustered in the center, while homozygote variants in right upper corner. Variant distribution was partially influenced by the differences in mean sequencing coverage targeting 100X and 300X in CZECANCA and DNA reference standards VCFs, respectively. The number of TP, TN, FP, FN, and total number of variant (= CZECANCA target) was used to calculate of sensitivity, specificity, and accuracy of CZECANCA analysis.

https://doi.org/10.1371/journal.pone.0195761.g007

based on smaller panels mainly simplify the clinical interpretation of the identified genotypes with a reduction of incidental findings. While their use is beneficial in clearly indicated patients with typical phenotype characteristics for a given cancer syndrome, the selection of a proper cancer-specific gene panel is not trivial in individuals with less characteristic features (e.g. patients from multi-cancer families). Moreover, our current knowledge of many cancer syndromes is based on the analyses of mostly prototypical cases, the testing criteria are changing dynamically, and the list of cancer predisposition genes with clinical utility is far less complete. Recently, Pearlman et al. analyzed 450 early-onset colorectal cancer patients and showed that a third (24/72) of mutation-positive patients did not meet the established genetic testing criteria for the gene(s) in which they had a mutation [41]. An analysis of mismatch repair (MMR) genes (traditionally linked to hereditary non-polyposis colorectal cancer) in a set of 34,981 cancer patients in a study by Espenschied et al. revealed that out of 528 patients with MMR mutations, 63 (11.9%) had breast cancer only and thus *MSH6* and *PMS2* mutation carriers may manifest with a hereditary breast and ovarian cancer phenotype [42]. In an analysis of *BRCA1* and *BRCA2* in 1,371 unselected breast cancer cohorts, Grindedal et al. showed that common guidelines identified only 45–90% of mutation carriers [43]. The ultimate solution to identify cancer risks would be an analysis of the whole exome (or even better genome) in all cancer patients; however, the implementation of such a strategy is not realistic at present [44]. We suppose that the use of larger multi-cancer panels (containing hundreds of genes) for an analysis of genetic risk in cancer patients is beneficial for several reasons. i) Such an analysis reveals a complex variation landscape of target genes in different cancers [7]. ii) It reveals carriers of concurrent pathogenic mutations and iii) it enables the testing of affected individuals from multi-cancer families with reasonable costs and turnaround time. Finally, iv) combining all genes of interest in a single panel simplifies and unifies laboratory procedures in a single workflow even if testing for different syndromes.

We have developed the custom-designed CZECANCA multi-cancer panel targeting the coding sequence of 219 cancer susceptibility or candidate genes, enabling the identification of a genetic predisposition in the most frequent hereditary cancer syndromes. Besides the established cancer susceptibility genes, we have decided to include also a subset of genes with low, clinically still unconfirmed utility, although their variants cannot be reported until their clinical evidence is known. These genes code for known interactors of established cancer susceptibility gene products, whose mutations may result in a similar phenotypic outcome. However, we suppose that knowledge obtained through the association of the identified genotypes with the phenotypic characteristics of the analyzed patients may substantially accelerate the process of clinical utility evaluation. Moreover, a subsidiary genetic report could be easily generated from the stored data in case of the approval of new cancer susceptibility genes included in CZECANCA. From the technical point of view, a larger genomic target has a favorable impact on panel complexity, improving its coverage uniformity [45].

The validation of the CZECANCA analytic workflow together with the bioinformatics pipeline is necessary for its implementation into routine diagnostics [46]. The presented analytical workflow was optimized for sequencing using MiSeq Illumina, representing the most frequently used NGS platform currently available in diagnostic laboratories. Genetic testing using gene panels is a cost-effective strategy [47]. The material costs for library preparation and sequencing (chemicals, kits, and disposables) using CZECANCA do not exceed €150 per patient in the standard settings (targeting sequencing coverage 100X). The CZECANCA workflow was intended mainly for medium throughput laboratories. As a universal panel, CZECANCA significantly reduces the turnaround time. The sequencing data for 30 analyzed DNA samples in one sequencing MiSeq run might be available in four days (three days for DNA fragmentation and library preparation, depending on hybridization time, and one day for MiSeq sequencing). We are aware that the low-covered or uncovered regions (affecting 12/219 CZECANCA-targeted genes) may require additional effort and time, when requested for genetic assessment.

The validation showed CZECANCA's high sensitivity, specificity, analytical robustness, and accuracy. We have demonstrated that SNVs and small/medium-size indels could be detected with high confidence. Moreover, we have shown that the uniform coverage (targeting to mean 100X coverage) of a target sequence enabled a robust identification of CNVs without the need of routine MLPA, serving as the method for independent CNVs confirmation or exclusion of false positivities. However, despite that the number of false positive calls was low and we detect no false negative sample in ACMG genes, we are aware that with caution needs to be interpreted positive CNV calls in genes for which MLPA assay (or other method) are not routinely available for confirmatory purposes. When required, presence of false positive signals can be reduced by the use of ultrasound fragmentation providing unbiased DNA shearing over enzymatic lysis and/or increased sequencing coverage.

Another advantage of NGS (over Sanger sequencing) is its ability to identify *cis* or *trans* positions of compound, closely localized heterozygous SNVs. For example, the position of double substitution in the *PALB2* gene creating a stop codon (c.661_662delinsTA; p.Val221*; NM_024675), which required further analyses (e.g. PTT) before the NGS era [10], can be identified directly from sequencing reads (Fig 8). The identification of additional pathogenic mutations during the validation procedure in negatively pre-tested samples indicated that a re-analysis is warranted for at least high-risk patients negatively tested by historical analyses based on indirect prescreening methods (e.g. PTT) or cDNA sequencing [48].

CZECANCA (CZEch CAncer paNel for Clinical Application) is intended to unify cancer predisposition testing in the Czech Republic, helping diagnostics laboratories transform the gene-by-gene strategy to NGS, even if is not a population-specific panel *per se*. NGS-based

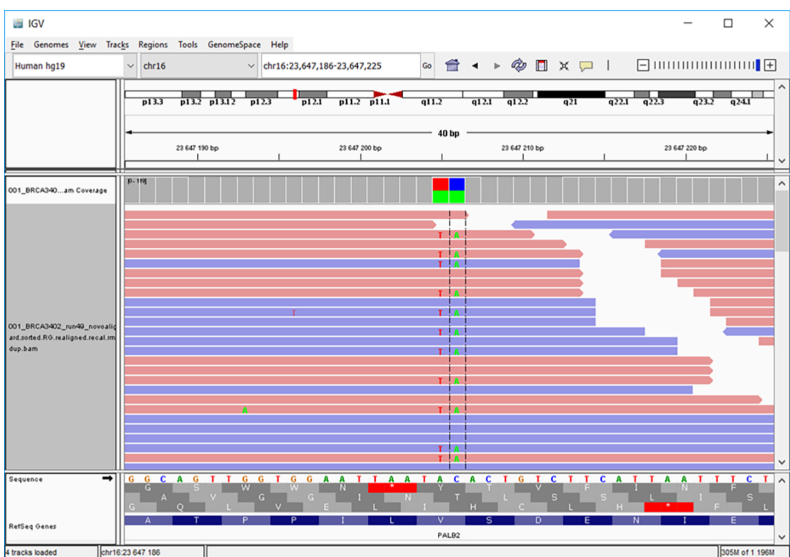


Fig 8. Identification of c.661_662delinsTA double substitution (p.Val221*) in *PALB2* (NM_024675). The BAM file displayed in IGV shows the *cis*-position of both substitutions in approximately 50% of forward (pink bars) and reverse (blue bars) reads, respectively.

https://doi.org/10.1371/journal.pone.0195761.g008

technologies bring new challenges including technological aspects, bioinformatics processing, the management of large datasets, and clinical interpretation of results [46]. The use of a uniform analytical and bioinformatics approach improves the identification of technical and platform-specific sequencing errors, as we demonstrated in inter-run and intra-run comparisons. Moreover, validation of the panel using reference standard DNA samples with known genotypes enabled identification of genomic loci (dominantly homopolymeric regions) providing these recurrent sequencing errors, which could be subsequently easily eliminated by bioinformatics. The use of CZECANCA will help generate a global view of constitutional variants from the perspective of known cancer predisposition and candidate genes in the population. Simultaneously with the sequencing of cancer patients, we aim to sequence non-cancer controls in order to identify and establish the frequency of population-specific neutral variants. The introduction of patients' and control genotypes with associated phenotypes into a nationwide database currently being created will simplify the interpretation of variants, which remains the main challenge at present. In general, NGS-based analyses result in an increased number of incidental findings or variants of unknown significance. The patient must be informed about this possibility before the testing and must have the opt in / opt out possibility clearly formulated in the informed consent. Consensus on what incidental information should be disclosed has yet to be reached. Currently, there is general agreement on reporting mutations in known high-penetrant genes in patients with a typical personal and family cancer history [38]. However, there is no agreement on pathogenic mutations in genes with lower penetrance or on mutations related to autosomal-recessive syndromes. These questions are currently being tackled in cooperating centers on a rather individual basis, depending on the formulation of the informed consents obtained, and on the clinical experience of the indicating geneticists [49].

In conclusion, CZECANCA allows comprehensive testing for a majority of frequent hereditary cancer syndromes while mitigating potential difficulties of incidental findings in non-cancer genes as seen in exome or genome sequencing. The reliability of the procedure enables an unbiased identification of variants present in patients, which together with a correct interpretation of variants is key for the effective management of hereditary cancer patients and their relatives.

Supporting information

S1 Table. List of 219 CZECANCA targeted genes with basic characteristics of their protein products. The primary gene target for the probe coverage was represented by coding sequences (cds) representing all exons (in case of known cancer susceptibility genes) or all coding exons (in other genes), including 10 bases from adjacent intronic regions. The promoter regions of the *BRCA1* and *BRCA2* genes were included into the primary target. Because of the strict design conditions, some clinically important regions were left untargeted (highlighted) for technical reasons such as repeats and homologous regions. (The characteristics of protein products were obtained from string.embl.de and/or genecards.org).
(XLSX)

S2 Table. Regions of interest with low coverage ≤20X. The average coverage is the mean from 10 randomly selected samples.
(XLSX)

S3 Table. Comparison of identified variants in the targeted exonic regions and 12 bp from adjacent introns with GATK quality >100 in three intra-run replicates of sample #2268. The DNA sample pooled for the enrichment in amounts corresponding to 33 ng (e.g. 1/30; considered as 100%), 75% and 50% of this amount, respectively. (Cov = coverage; Q = quality; discordant variants are highlighted).
(XLSX)

S4 Table. Comparison of identified variants in the targeted exonic regions and 12 bp from adjacent introns with GATK quality >100 in two independent run replicates of sample #3647. All values of coverages (Cov) of sample #3647 in run 14 were corrected by a factor of 1.3880 to normalize coverages between samples for presentation in Fig 4B. (Q = quality; discordant variants are highlighted).
(XLSX)

S5 Table. Comparison of identified variants in the targeted exonic regions and 12 bp from adjacent introns with GATK quality >100 in sample #3582 analyzed independently in four participating laboratories(Lab). All values of coverages (Cov) in Lab2, Lab3, and Lab4 were corrected to the coverage of Lab1 by a factor shown in line 336 to normalize coverages between samples for Fig 4C. (discordant variants are highlighted).
(XLSX)

S6 Table. List of variants used for the validation of SNVs detection.
(XLSX)

S7 Table. List of CNVs used for the validation of a large genomic rearrangements analysis.
(XLSX)

S8 Table. CNV scores (from CNVkit software) of bins in *BRCA1*, *PALB2*, *CHEK2*, and *TP53*. The numbers of samples with previously characterized CNVs are highlighted in red. The table show raw values obtained from CNVkit as well as median-normalized values. The normalized values >0.5 (highlighted in green) were indicative for the presence of a duplication, while values <-0.6 (highlighted in yellow) were indicative for a deletion. Data from this table were used for creation of Fig 5.
(XLSX)

S9 Table. Variants identified in five Coriell Institute reference samples sequenced using CZECANCA pipeline and their comparison with VCF files obtained from GIAB website.

The considered targeted region encompasses 628,069 bases of CZECANCA target region. False negative variants are highlighted.
(XLSX)

S10 Table. Variant consensus analysis report from EMQN (NGS pilot 2016) for CZEN-CANCA sequencing of a reference sample.
(XLSX)

S11 Table. Results of CNV analysis performed in two validation sets consisting of four runs from Laboratory 1 (116 samples prepared using the ultrasound DNA fragmentation on Covaris) and four runs from Laboratory 3 (125 other samples prepared using the enzymatic DNA cleavage by Fragmentase). To estimate number of false positive (FP) and false negative (FN) samples, data for CNV analysis of Coriell Institute reference samples (Coriell; 10 samples analyzed in Laboratory 1 and prepared using the ultrasound DNA fragmentation on Covaris) were added. The values in cells represent differences of CNV scores for a given cell (i.e. sample in the coordinate) from the median value of signals from particular sample group (i.e. Coriell—columns Q-Z, Laboratory 1—columns AB-EM, Laboratory 3—columns EO-JI) in a given CNVkit_bin_set_coordinate (column A). Values in cells showing individual analyzed samples from particular sample group exceeding the given CNVkit threshold value for deletion (<-0,6) and duplication (>0,45) are highlighted as red and green cells, respectively. The columns C-O provide several aggregated metrics, that include number of individual samples in which deletion (columns G-I), duplication (J-L), or deletion+duplication (M-O) was found in a given coordinate in particular sample group. Columns C-E enable identification of non-informative bin sets with suspected false positive (FP) signals (indicated by the value = 1) that include regions on X chromosome called in male samples as deletions (highlighted in blue in column B), regions with insufficient coverage or containing pseudogenes (highlighted in orange and yellow, respectively; in column B), or bin sets containing the improbable number of deletions+duplications exceeding the 4% of analyzed samples in a particular sample group.
(XLSX)

S1 Fig. Run of consecutive bin set coordinates with values indicating a deletion (< -0.6; red) or a duplication (> 0.45; green) increases the probability of a real rearrangement. The *BRCA1* and *BRIP1* deletions were confirmed by MLPA analyses, which are currently no available for confirmation of secondary findings in *MSR1* or *ZNF350*. (The graphs expressed normalized CNVkit values shown in S11 Table).
(TIF)

S2 Fig. CNV analysis of genes *BRCC3*, *FANCB*, *GPC3*, and *UBE2A* localized on X chromosome enabled to demonstrate differences in normalized CNVkit values in samples carrying a real 'deletion' in samples prepared by ultrasound DNA fragmentation or enzymatic DNA lysis. The XX and X indicates areas of samples obtained from female and male probands, respectively. (The graphs expressed normalized CNVkit values shown in S11 Table). Upper panel shows normalized CNVkit values in 116 samples analyzed in four runs in laboratory 1. Lower panel shows normalized CNVkit values in 125 other samples analyzed in four runs in laboratory 3.
(TIF)

Acknowledgments

We would like to thank Jaroslav Vohanka and Xavier Miro (Roche) for their valuable advice in probe design, Jana Chrudimska for technical assistance with sequencing, and Jan Flemr for language editing.

Author Contributions

Conceptualization: Jana Soukupova, Marketa Janatova, Zdenek Kleibl.

Data curation: Jana Soukupova, Petra Zemankova, Viktor Stranecky, Michal Vocka, Zdenek Kleibl.

Formal analysis: Jana Soukupova, Marketa Janatova, Lenka Foretova, Petra Kleiblova, Michal Vocka, Zdenek Kleibl.

Funding acquisition: Jana Soukupova, Lenka Foretova, Stanislav Kmoch, Zdenek Kleibl.

Investigation: Jana Soukupova, Petra Zemankova, Klara Lhotova, Marketa Janatova, Marianna Borecka, Lenka Stolarova, Filip Lhota, Eva Machackova, Spiros Tavandzis, Petra Kleiblova, Michal Vocka.

Methodology: Jana Soukupova, Petra Zemankova, Klara Lhotova, Marketa Janatova, Lenka Foretova, Viktor Stranecky, Petra Kleiblova, Hana Hartmannova, Katerina Hodanova, Stanislav Kmoch, Zdenek Kleibl.

Project administration: Jana Soukupova.

Resources: Michal Vocka.

Software: Petra Zemankova, Viktor Stranecky.

Supervision: Jana Soukupova, Marketa Janatova, Lenka Foretova, Viktor Stranecky, Hana Hartmannova, Katerina Hodanova, Stanislav Kmoch, Zdenek Kleibl.

Validation: Jana Soukupova, Petra Zemankova, Klara Lhotova, Marianna Borecka, Lenka Stolarova, Filip Lhota, Eva Machackova, Spiros Tavandzis.

Visualization: Petra Zemankova, Zdenek Kleibl.

Writing – original draft: Jana Soukupova, Zdenek Kleibl.

Writing – review & editing: Jana Soukupova, Petra Zemankova, Klara Lhotova, Marketa Janatova, Marianna Borecka, Lenka Stolarova, Filip Lhota, Lenka Foretova, Eva Machackova, Viktor Stranecky, Spiros Tavandzis, Petra Kleiblova, Michal Vocka, Hana Hartmannova, Katerina Hodanova, Stanislav Kmoch, Zdenek Kleibl.

References

1. Kulkarni A, Carley H. Advances in the recognition and management of hereditary cancer. Br Med Bull. 2016; 120(1):123–38. https://doi.org/10.1093/bmb/ldw046 PMID: 27941041
2. Stoffel EM, Cooney KA. Advances in inherited cancers: Introduction. Semin Oncol. 2016; 43(5):527. https://doi.org/10.1053/j.seminoncol.2016.09.003 PMID: 27899182
3. Rahman N. Mainstreaming genetic testing of cancer predisposition genes. Clin Med. 2014; 14(4):436–9. https://doi.org/10.7861/clinmedicine.14-4-436 PMID: 25099850
4. Rahman N. Realizing the promise of cancer predisposition genes. Nature. 2014; 505(7483):302–8. https://doi.org/10.1038/nature12981 PMID: 24429628
5. Foulkes WD. Inherited susceptibility to common cancers. N Engl J Med. 2008; 359(20):2143–53. https://doi.org/10.1056/NEJMra0802968 PMID: 19005198
6. Feero WG. Clinical application of whole-genome sequencing: proceed with care. JAMA. 2014; 311 (10):1017–9. https://doi.org/10.1001/jama.2014.1718 PMID: 24618961
7. Shah PD, Nathanson KL. Application of Panel-Based Tests for Inherited Risk of Cancer. Annu Rev Genomics Hum Genet. 2017; 18(1):201–27. https://doi.org/10.1146/annurev-genom-091416-035305 PMID: 28504904

8. Pohlreich P, Stribrna J, Kleibl Z, Zikan M, Kalbacova R, Petruzelka L, et al. Mutations of the BRCA1 gene in hereditary breast and ovarian cancer in the Czech Republic. Med Princ Pract. 2003; 12(1):23–9. https://doi.org/10.1159/000068163 PMID: 12566964
9. Pohlreich P, Zikan M, Stribrna J, Kleibl Z, Janatova M, Kotlas 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–R36. https://doi.org/10.1186/bcr1282 PMID: 16168118
10. Janatova M, Kleibl Z, Stribrna J, Panczak A, Vesela K, Zimovjanova M, 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):2323–32. https://doi.org/10.1158/1055-9965.EPI-13-0745-T PMID: 24136930
11. Kleibl Z, Havranek O, Hlavata I, Novotny J, Sevcik J, Pohlreich P, 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):618–24. https://doi.org/10.1016/j.ejca.2008.09.022 PMID: 18996005
12. Soukupova J, Dundr P, Kleibl Z, Pohlreich P. Contribution of mutations in ATM to breast cancer development in the Czech population. Oncol Rep. 2008; 19(6):1505–10. PMID: 18497957
13. Borecka M, Zemankova P, Vocka M, Soucek P, Soukupova J, Kleiblova P, et al. Mutation analysis of the PALB2 gene in unselected pancreatic cancer patients in the Czech Republic. Cancer Genet. 2016; 209(5):199–204. https://doi.org/10.1016/j.cancergen.2016.03.003 PMID: 27106063
14. Kleibl Z, Fidlerova J, Kleiblova P, Kormunda S, Bilek M, Bouskova K, 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):303–16. https://doi.org/10.4149/neo_2009_04_303 PMID: 19473056
15. Kleiblova P, Shaltiel IA, Benada J, Sevcik J, Pechackova S, Pohlreich P, et al. Gain-of-function mutations of PPM1D/Wip1 impair the p53-dependent G1 checkpoint. J Cell Biol. 2013; 201(4):511–21. https://doi.org/10.1083/jcb.201210031 PMID: 23649806
16. Janatova M, Soukupova J, Stribrna J, Kleiblova P, Vocka M, Boudova P, 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. https://doi.org/10.1371/journal.pone.0127711 PMID: 26057125
17. Ticha I, Kleibl Z, Stribrna J, Kotlas J, Zimovjanova M, Mateju M, 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):337–47. https://doi.org/10.1007/s10549-010-0745-y PMID: 20135348
18. Havranek O, Kleiblova P, Hojny J, Lhota F, Soucek P, Trneny M, et al. Association of Germline CHEK2 Gene Variants with Risk and Prognosis of Non-Hodgkin Lymphoma. PLoS One. 2015; 10(10): e0140819. https://doi.org/10.1371/journal.pone.0140819 PMID: 26506619
19. Hardwick SA, Deveson IW, Mercer TR. Reference standards for next-generation sequencing. Nat Rev Genet. 2017; 18(8):473–84. https://doi.org/10.1038/nrg.2017.44 PMID: 28626224
20. Lhota F, Zemankova P, Kleiblova P, Soukupova J, Vocka M, Stranecky V, et al. Hereditary truncating mutations of DNA repair and other genes in BRCA1/BRCA2/PALB2-negatively tested breast cancer patients. Clin Genet. 2016. https://doi.org/10.1111/cge.12748 PMID: 26822949
21. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010; 38(16):e164. https://doi.org/10.1093/nar/gkq603 PMID: 20601685
22. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010; 20(9):1297–303. https://doi.org/10.1101/gr.107524.110 PMID: 20644199
23. Das R, Ghosh SK. Genetic variants of the DNA repair genes from Exome Aggregation Consortium (EXAC) database: significance in cancer. DNA Repair (Amst). 2017; 52:92–102. https://doi.org/10.1016/j.dnarep.2017.02.013 PMID: 28259467
24. Genomes Project C, Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, et al. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012; 491(7422):56–65. https://doi.org/10.1038/nature11632 PMID: 23128226
25. Landrum MJ, Lee JM, Riley GR, Jang W, Rubinstein WS, Church DM, et al. ClinVar: public archive of relationships among sequence variation and human phenotype. Nucleic Acids Res. 2014; 42(Database issue):D980–5. https://doi.org/10.1093/nar/gkt1113 PMID: 24234437

26. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc. 2009; 4(7):1073–81. https://doi.org/10.1038/nprot.2009.86 PMID: 19561590
27. Wildeman M, van Ophuizen E, den Dunnen JT, Taschner PE. Improving sequence variant descriptions in mutation databases and literature using the Mutalyzer sequence variation nomenclature checker. Human Mutat. 2008; 29(1):6–13. https://doi.org/10.1002/humu.20654 PMID: 18000842
28. Schwarz JM, Rodelsperger C, Schuelke M, Seelow D. MutationTaster evaluates disease-causing potential of sequence alterations. Nat Methods. 2010; 7(8):575–6. https://doi.org/10.1038/nmeth0810-575 PMID: 20676075
29. Chun S, Fay JC. Identification of deleterious mutations within three human genomes. Genome Res. 2009; 19(9):1553–61. https://doi.org/10.1101/gr.092619.109 PMID: 19602639
30. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. Nat Methods. 2010; 7(4):248–9. https://doi.org/10.1038/nmeth0410-248 PMID: 20354512
31. Pollard KS, Hubisz MJ, Rosenbloom KR, Siepel A. Detection of nonneutral substitution rates on mammalian phylogenies. Genome Res. 2010; 20(1):110–21. https://doi.org/10.1101/gr.097857.109 PMID: 19858363
32. Cooper GM, Stone EA, Asimenos G, Program NCS, Green ED, Batzoglou S, et al. Distribution and intensity of constraint in mammalian genomic sequence. Genome Res. 2005; 15(7):901–13. https://doi.org/10.1101/gr.3577405 PMID: 15965027
33. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet. 2014; 46(3):310–5. https://doi.org/10.1038/ng.2892 PMID: 24487276
34. Machackova E, Hazova J, Stahlova Hrabincova E, Vasickova P, Navratilova M, Svoboda M, et al. [Retrospective NGS Study in High-risk Hereditary Cancer Patients at Masaryk Memorial Cancer Institute]. Klin Onkol. 2016; 29 Suppl 1:S35–45. PMID: 26691941.
35. Thorvaldsdottir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform. 2013; 14(2):178–92. https://doi.org/10.1093/bib/bbs017 PMID: 22517427
36. Ye K, Schulz MH, Long Q, Apweiler R, Ning Z. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. Bioinformatics. 2009; 25(21):2865–71. https://doi.org/10.1093/bioinformatics/btp394 PMID: 19561018
37. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015; 17(5):405–24. https://doi.org/10.1038/gim.2015.30 PMID: 25741868
38. Kalia SS, Adelman K, Bale SJ, Chung WK, Eng C, Evans JP, et al. Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2.0): a policy statement of the American College of Medical Genetics and Genomics. Genet Med. 2017; 19(2):249–55. https://doi.org/10.1038/gim.2016.190 PMID: 27854360
39. Zhao M, Wang Q, Wang Q, Jia P, Zhao Z. Computational tools for copy number variation (CNV) detection using next-generation sequencing data: features and perspectives. BMC Bioinformatics. 2013; 14(11):S1. https://doi.org/10.1186/1471-2105-14-s11-s1 PMID: 24564169
40. Easton DF, Pharoah PD, Antoniou AC, Tischkowitz M, Tavtigian SV, Nathanson KL, et al. Gene-panel sequencing and the prediction of breast-cancer risk. New Engl J Med. 2015; 372(23):2243–57. https://doi.org/10.1056/NEJMsr1501341 PMID: 26014596
41. Pearlman R, Frankel WL, Swanson B, et al. Prevalence and spectrum of germline cancer susceptibility gene mutations among patients with early-onset colorectal cancer. JAMA Oncol. 2017; 3(4):464–71. https://doi.org/10.1001/jamaoncol.2016.5194 PMID: 27978560
42. Espenschied CR, LaDuca H, Li S, McFarland R, Gau C-L, Hampel H. Multigene Panel Testing Provides a New Perspective on Lynch Syndrome. J Clin Oncol. 2017; 35(22):2568–75. https://doi.org/10.1200/JCO.2016.71.9260 PMID: 28514183
43. Grindedal EM, Heramb C, Karsrud I, Ariansen SL, Maehle L, Undlien DE, et al. Current guidelines for BRCA testing of breast cancer patients are insufficient to detect all mutation carriers. BMC cancer. 2017; 17(1):438. https://doi.org/10.1186/s12885-017-3422-2 PMID: 28637432
44. Majewski J, Schwartzentruber J, Lalonde E, Montpetit A, Jabado N. What can exome sequencing do for you? J Med Genet. 2011; 48(9):580–9. https://doi.org/10.1136/jmedgenet-2011-100223 PMID: 21730106

45. Sims D, Sudbery I, Ilott NE, Heger A, Ponting CP. Sequencing depth and coverage: key considerations in genomic analyses. Nat Rev Genet. 2014; 15(2):121–32. https://doi.org/10.1038/nrg3642 PMID: 24434847

46. Matthijs G, Souche E, Alders M, Corveleyn A, Eck S, Feenstra I, et al. Guidelines for diagnostic next-generation sequencing. Eur J Human Genet. 2016; 24(1):2–5. https://doi.org/10.1038/ejhg.2015.226 PMID: 26508566

47. Azimi M, Schmaus K, Greger V, Neitzel D, Rochelle R, Dinh T. Carrier screening by next-generation sequencing: health benefits and cost effectiveness. Mol Genet Genomic Med. 2016; 4(3):292–302. https://doi.org/10.1002/mgg3.204 PMID: 27247957

48. Moran O, Nikitina D, Royer R, Poll A, Metcalfe K, Narod SA, et al. Revisiting breast cancer patients who previously tested negative for BRCA mutations using a 12-gene panel. Breast Cancer Res Treat. 2017; 161(1):135–42. https://doi.org/10.1007/s10549-016-4038-y PMID: 27798748

49. Soukupová J, Zemánková P, Kleiblová P, Janatová M, Kleibl Z. [CZECANCA: CZEch CAncer paNel for Clinical Application—Design and Optimization of the Targeted Sequencing Panel for the Identification of Cancer Susceptibility in High-risk Individuals from the Czech Republic]. Klin Onkol. 2016; 29 Suppl 1: S46–54. Czech. PMID: 26691942.

Příloha 4 - Soukupová, J., K. Lhotová, P. Zemánková, M. Vočka, M. Janatová, L. Stolařová, M. Borecká, P. Kleiblová, E. Macháčková, L. Foretová, M. Koudová, F. Lhota, S. Tavandzis, M. Zikán, V. Stránecký, K. Veselá, A. Panczak, J. Kotlas, and Z. Kleibl. 2019. "***Contribution of Massive Parallel Sequencing to Diagnosis of Hereditary Ovarian Cancer in the Czech Republic***." Klin Onkol 32 (Supplementum2):72-78. doi: 10.14735/amko2019S72.

proLékaře.cz | 28.9.2020

Přínos masivního paralelního sekvenování pro diagnostiku dědičných forem nádorů ovaria v České republice

Contribution of Massive Parallel Sequencing to Diagnosis of Hereditary Ovarian Cancer in the Czech Republic

Soukupová J.[1], Lhotová K.[1], Zemánková P.[1], Vočka M.[2], Janatová M.[1], Stolařová L.[1], Borecká M.[1], Kleiblová P.[3], Macháčková E.[4], Foretová L.[4], Koudová M.[5], Lhota F.[5], Tavandzis S.[6], Zikán M.[7], Stránecký V.[8], Veselá K.[3], Panczak A.[3,9], Kotlas J.[3], Kleibl Z.[1]

[1] Ústav biochemie a experimentální onkologie, 1. LF UK v Praze
[2] Onkologická klinika 1. LF UK a VFN v Praze
[3] Ústav biologie a lékařské genetiky, 1. LF UK a VFN v Praze
[4] Oddělení epidemiologie a genetiky nádorů, Masarykův onkologický ústav, Brno
[5] Centrum lékařské genetiky a reprodukční medicíny GENNET, Praha
[6] Oddělení lékařské genetiky, Laboratoře AGEL, Nový Jičín
[7] Gynekologicko-porodnická klinika 1. LF UK a Nemocnice na Bulovce, Praha
[8] Ústav dědičných metabolických poruch, 1. LF UK a VFN v Praze
[9] Radioterapeutická a onkologická klinika FN Královské Vinohrady, Praha

Souhrn

Východiska: Karcinom ovaria, závažné nádorové onemocnění s vysokou mortalitou, je v České republice diagnostikováno každým rokem přibližně u 1 000 žen. Riziko vzniku onemocnění je zvýšeno u nosiček mutací v některých nádorových predispozičních genech. S vysokým relativním rizikem (RR > 5) jsou spojeny mutace v genech *BRCA1, BRCA2, BRIP1*, geny Lynchova syndromu, *RAD51C, RAD51D, STK11*; s možným zvýšením rizika mutace v genech *ATM, CHEK2, NBN, PALB2, BARD1*. Cílem práce bylo určit frekvenci mutací v nádorových predispozičních genech v naší populaci. ***Metody a výsledky:*** Celkem 1 057 pacientek s karcinomem ovaria a 617 nenádorových kontrol bylo vyšetřeno pomocí panelového sekvenování nové generace na platformě Illumina. Patogenní mutace ve vysoko rizikových genech, vč. velkých genomových přestaveb, byly v našem souboru zachyceny u 30,6 % pacientek; u neselektovaných pacientek byla frekvence mutací téměř 25 %, u pacientek s negativní rodinnou anamnézou 18 %. Nejčastěji mutovanými predispozičními geny byly *BRCA1* a *BRCA2*, součet frekvence mutací v ostatních ovariálních predispozičních genech odpovídal frekvenci mutací v genu *BRCA2*. Záchyt mutací u pacientek starších 70 let byl více než třikrát vyšší v porovnání s pacientkami ve věku pod 30 let. ***Závěr:*** Karcinom ovaria je heterogenní onemocnění s vysokým podílem dědičné formy onemocnění. Vzhledem k nedostatku adekvátních screeningových modalit pro včasnou diagnostiku onemocnění je identifikace nosiček mutací v ovariálních predispozičních genech klíčová, s vysokým potenciálem k celkovému snížení mortality z důvodu karcinomu ovaria.

Klíčová slova

karcinom ovaria – nádorové geny – mutace – masivní paralelní sekvenování – sekvenování nové generace – panel genů

Tato práce byla podpořena grantem AZV 15-27695A, SVV2019/260367, PROGRES Q28/LF1.

This work was supported by grants AZV 15-27695A, SVV2019/260367, PROGRES Q28/LF1.

Autoři deklarují, že v souvislosti s předmětem studie nemají žádné komerční zájmy.

The authors declare they have no potential conflicts of interest concerning drugs, products, or services used in the study.

Redakční rada potvrzuje, že rukopis práce splnil ICMJE kritéria pro publikace zasílané do biomedicínských časopisů.

The Editorial Board declares that the manuscript met the ICMJE recommendation for biomedical papers.

RNDr. Jana Soukupová, Ph.D.
Ústav biochemie a experimentální onkologie
1. LF UK v Praze
U Nemocnice 5
128 53 Praha 2
e-mail: jproko@lf1.cuni.cz

Obdrženo/Submitted: 7. 3. 2019
Přijato/Accepted: 24. 4. 2019

doi: 10.14735/amko2019S72

proLékaře.cz | 28.9.2020

Summary

Background: Ovarian cancer is a disease with high mortality. Approximately 1,000 women are diagnosed with ovarian cancer in the Czech Republic annually. Women harboring a mutation in cancer-predisposing genes face an increased risk of tumor development. Mutations in *BRCA1, BRCA2, BRIP1*, and Lynch syndrome genes (*RAD51C, RAD51D,* and *STK11*) are associated with a high risk of ovarian cancer, and mutations in *ATM, CHEK2, NBN, PALB2,* and *BARD1* appear to increase the risk. Our aim was to examine the frequency of mutations in cancer-predisposing genes in the Czech Republic. ***Materials and methods:*** We analyzed 1,057 individuals including ovarian cancer patients and 617 non-cancer controls using CZECANCA panel next-generation sequencing on the Illumina platform. Pathogenic mutations in high-risk genes, including CNVs, were detected in 30.6% of patients. The mutation frequency reached 25.0% and 18.2% in subgroups of unselected ovarian cancer patients and patients with a negative family cancer history, respectively. The most frequently mutated genes were *BRCA1* and *BRCA2.* The overall frequency of mutations in non-BRCA genes was comparable to that in *BRCA2.* The mutation frequency in ovarian cancer patients aged >70 years was three times higher than that in patients diagnosed before the age of 30. ***Conclusion:*** Ovarian cancer is a heterogeneous disease with a high proportion of hereditary cases. The lack of efficient screening for early diagnosis emphasizes the importance of identifying carriers of mutations in ovarian cancer-predisposing genes; this is because proper follow-up and prevention strategies can reduce overall ovarian cancer-related mortality.

Key words

ovarian neoplasms – cancer genes – mutation – massively-parallel sequencing – next generation sequencing – gene panel

Úvod

Karcinom ovaria patří mezi závažná nádorová onemocnění žen v ČR. Každým rokem je diagnostikován u přibližně 1 000 žen a zhruba 700 žen tomuto onemocnění podlehne [1]. Incidence ovariálního karcinomu stoupá s věkem a nejvyšší prevalence dosahuje mezi 6. a 7. dekádou. Celoživotní riziko rozvoje ovariálního karcinomu v běžné populaci se u žen pohybuje kolem 1,5 %, avšak je významně zvýšeno u nosiček zárodečných mutací v některých nádorových predispozičních genech.

Dědičná forma karcinomu ovaria představuje přibližně 20 % případů, tedy významně více než u jiných běžných typů nádorů. V současné době jsou proto ke genetickému vyšetření indikovány všechny pacientky s karcinomem ovaria (ale i vejcovodů a primárního peritoneálního karcinomu) bez ohledu na věk. Identifikace mutací predisponujících ke vzniku karcinomu ovaria umožňuje nabídnout nosičkám odpovídající léčebnou strategii a genetické poradenství a je podmínkou cílené preventivní péče o asymptomatické osoby s mutací.

Hlavními genetickými faktory dědičné formy karcinomu ovaria jsou, podobně jako u hereditární formy karcinomu prsu, mutace v genech *BRCA1* a *BRCA2*. Celoživotní riziko rozvoje karcinomu ovaria se pohybuje u nosiček mutací v genu *BRCA1* mezi 35–60 %, u nosiček mutací v *BRCA2* mezi 12–25 % [2,3]. Ačkoliv je klinický význam mutací v *BRCA1/2* vysoký, nevysvětluje všechny případy familiárních forem onemocnění. Zvýšené riziko vzniku karcinomu ovaria bylo prokázáno i u nosiček mutací v dalších genech kódujících proteiny, které se podobně jako *BRCA1* a *BRCA2* podílejí na reparaci dvouřetězcových zlomů DNA cestou homologní rekombinace. V současné době existují klinická doporučení zohledňující zvýšené riziko vzniku karcinomu ovaria u nosiček mutací v 10 genech (*BRCA1, BRCA2, BRIP1, EPCAM, MSH2, MLH1, MSH6, RAD51C, RAD51D, STK11)* a zvýšené riziko vzniku karcinomu prsu pro 12 genů (*ATM, BARD1, BRCA1, BRCA2, CDH1, CHEK2, NBN, NF1, PALB2, PTEN, STK11, TP53*) [4].

V porovnání s četností mutací v genech *BRCA1/2* je frekvence mutací v dalších predispozičních genech výrazně (často řádově) nižší a významně se liší mezi populacemi. Postupná analýza jednotlivých genů je z důvodu značné finanční a časové náročnosti v praxi nepoužitelná. K dramatické změně možností diagnostiky došlo s rozvojem sekvenování nové generace (next generation sequencing – NGS), jež umožnilo vyšetření většího počtu genů v krátké době při příznivých ekonomických nákladech [5].

Pro diagnostické účely identifikace nádorové predispozice jsme zkonstruovali a validovali univerzální panel CZECANCA (CZEch CAncer paNel for Clinical Application) [6] pro komplexní, rentabilní a rychlou analýzu germinálních mutací vč. velkých přestaveb v hlavních predispozičních genech, ale i v kandidátních genech asociovaných se zvýšeným rizikem vzniku nejčastějších solidních nádorů v naší populaci [7].

Metody

Pomocí panelu CZECANCA v1.0, který obsahuje sondy cílící na kódující sekvence 219 genů, jsme analyzovali 1 057 pacientek s karcinomem ovaria. Analýza probíhala v laboratořích Ústavu biochemie a experimentální onkologie 1. LF UK v Praze, Masarykova onkologického ústavu v Brně, v laboratořích Gennet v Praze a Agel v Novém Jičíně. Medián věku v době diagnózy byl 52,6 roku (14,8–86,2). Klinické charakteristiky pacientek jsou uvedeny v tab. 1.

Abychom mohli odlišit vzácné patogenní a populačně specifické mutace, vyšetřili jsme pomocí stejného panelu genů skupinu 617 zdravých kontrol (jedinci starší 60 let bez osobní a rodinné nádorové anamnézy u přímých příbuzných).

Vzorky genomové DNA izolované z leukocytů periferní krve od pacientů indikovaných ke genetickému vyšetření byly analyzovány podle jednotného protokolu, který zahrnoval i následné bioinformatické zpracování dat [6].

Identifikované zárodečné varianty v genech spojených se syndromem dědičného karcinomu prsu a ovaria byly prioritizovány na základě frekvence výskytu varianty u méně než 1 % vzorků v projektu 1 000 Genomes [8] a současně u méně než 1 % vzorků našich populačních kontrol. Zbývající raritní varianty byly klasifikovány do 5 tříd dle IARC (International Agency for Research on Cancer): 1 – benigní, 2 –

proLékaře.cz | 28.9.2020

pravděpodobně benigní, 3 – varianta nejasného významu (variant of unknown signifikance – VUS), 4 – pravděpodobně patogenní, 5 – patogenní. Jako patogenní nebo pravděpodobně patogenní byly dle doporučení ENIGMA (Evidence-based Network for the Interpretation of Germline Mutant Alleles) konsorcia [9] označeny varianty vedoucí k předčasnému zkrácení proteinu (nonsense a posunové mutace, velké genomové přestavby, pokud nejsou klasifikovány jinak), mutace postihující konzervativní sestřihová místa a missense mutace klasifikované jako patogenní v databázi ClinVar [10].

V analyzovaném souboru jsme se zaměřili na hodnocení prokazatelně patogenních alterací v genech, ke kterým v současnosti existují klinická doporučení péče o nosičky mutací zohledňující riziko vzniku karcinomu ovaria (*BRCA1, BRCA2, BRIP1, EPCAM, MSH2, MLH1, MSH6, RAD51C, RAD51D, STK11*) a prsu (*ATM, BARD1, CHEK2, NBN, PALB2*) [4].

Tab. 1. Charakteristika souboru 1 057 pacientek s karcinomem ovaria.

	Počet pacientek	% ze známých
Věk v době diagnózy		
do 29 let	75	7,2
30–49 let	365	35,2
50–69 let	529	51
> 70 let	68	6,6
není k dispozici	20	
Histologie		
high-grade serózní	375	41,3
serózní, bez určení grade	138	15,2
low-grade serózní	75	8,3
endometrioidní	77	8,5
mucinózní	35	3,9
ze světlých buněk	12	1,3
jiný maligní histologický typ	81	8,9
border-line tumors	121	13,3
není k dispozici	143	
Osobní anamnéza		
pouze karcinom ovaria	817	77,3
karcinom ovaria a prsu	180	17
karcinom ovaria a jiný nádor (mimo karcinom prsu)	60	5,7
Rodinná anamnéza		
bez onkologického onemocnění	495	48,2
karcinom prsu a ovaria v rodině	288	28
karcinom ovaria v rodině	102	9,9
mnohočetný výskyt nádorových onemocnění v rodině	143	13,9
není k dispozici	29	

Výsledky

Mutace v genech predisponujících ke vzniku dědičné formy karcinomu ovaria (*BRCA1, BRCA2, BRIP1, EPCAM, MSH2, MLH1, MSH6, RAD51C, RAD51D, STK11*)

Patogenní dědičnou mutaci v některém z genů predisponujících ke karcinomu ovaria jsme identifikovali celkem u 323 z 1 057 (30,6 %) analyzovaných pacientek s karcinomem ovaria (tab. 2). Nejvíce patogenních mutací jsme zachytili v hlavních predispozičních genech *BRCA1* (210/1 057; 19,9 %) a *BRCA2* (75/1 057; 7,1 %). Mutace v genech způsobujících Lynchův syndrom byly nalezeny u 9 z 1 057 (0,9 %) pacientek. Zbývající 3 % nalezených mutací se rovnoměrně rozdělila mezi 29 nosiček mutací v genech *RAD51C, RAD51D, BRIP1*. Ve skupině kontrol byly v uvedených genech s prokázanou asociací s karcinomem ovaria zachyceny patogenní mutace pouze u 4 z 617 osob (0,7 %; tab. 2).

Klinické a histopatologické charakteristiky ovlivňující pravděpodobnost výskytu germinálních mutací

S ohledem na osobní onkologickou anamnézu (tab. 3) je nejvyšší pravděpodobnost výskytu mutací u pacientek s duplicitou karcinomu prsu a ovaria, kde mutaci nacházíme u téměř dvou třetin vyšetřovaných (107/180; 59,4 %). U čtvrtiny vyšetřovaných se dědičné mutace vyskytovaly u pacientek se solitární diagnózou karcinomu ovaria (203/817; 24,8 %) nebo u pacientek, které kromě karcinomu ovaria vyvinuly i jiné nádorové onemocnění (15/60; 25,0 %). Mutace v genech *RAD51C, RAD51D a BRIP1* převažovaly u pacientek s diagnózou solitárního ovariálního karcinomu (25/29; 86 %).

Vzhledem k rodinné anamnéze (tab. 3) jsme nalezli nejvyšší procento mutací u pacientek z rodin, ve kterých se vyskytoval pouze karcinom ovaria (55/102; 53,9 %), přičemž mutace téměř výhradně postihovaly geny *BRCA1* (43/55; 78 %) a *BRCA2* (10/55; 18 %). Vysoký výskyt mutací v predispozičních genech jsme zaznamenali i u pacientek s rodinným výskytem karcinomu ovaria a prsu

proLékaře.cz | 28.9.2020

Tab. 2. Frekvence patogenních mutací v genech jasně predisponujících ke vzniku dědičné formy karcinomu ovaria.

	Pacientky; n = 1 057 n (%)	Kontroly; n = 617 n (%)	p
BRCA1	210 (19,9 %)	0	$2,2 \times 10^{-16}$
BRCA2	75 (7,1 %)	4 (0,6 %)	$2,8 \times 10^{-11}$
RAD51C	11 (1,0 %)	0	0,009
BRIP1	9 (0,9 %)	0	0,031
RAD51D	9 (0,9 %)	0	0,031
MLH1	5 (0,5 %)	0	0,031
MSH2	2 (0,2 %)	0	
MSH6	2 (0,2 %)	0	
celkem	323 (30,6 %)	4 (0,6 %)	$2,2 \times 10^{-16}$

(129/288; 44,8 %) a u pacientek s rodinným výskytem karcinomu ovaria a dalších typů nádorů (45/143; 31,5 %). Mutace však byla zachycena i u 18,2 % pacientek s negativní rodinnou anamnézou (90/495), což ve výsledku znamená, že tato skupina zahrnovala více než čtvrtinu nosiček patogenních mutací (90/323; 27,9 %).

Zastoupení histologických typů ovariálních tumorů u nosiček mutací v predispozičních genech ukazuje, že serózní nádory tvoří přibližně dvě třetiny ovariálních karcinomů ve skupinách nosiček s mutacemi v genech predisponujících ke vzniku hereditárního karcinomu ovaria (graf 1). Výjimkou byla malá skupina devíti nosiček mutací v genech Lynchova syndromu, ve které jsme za-

Tab. 3. Výskyt germinálních mutací v genech predisponujících ke vzniku karcinomu ovaria v závislosti na klinických a histopatologických charakteristikách.

	BRCA1 (%)	*BRCA2* (%)	*BRIP1, RAD51C, RAD51D* (%)	*MLH1, MSH2, MSH6* (%)	Celkem (%)
Dle osobní anamnézy					
pouze karcinom ovaria (n = 817)	132 (16,2)	43 (5,3)	25 (3,1)	3(0,4)	203 (24,8)
karcinom prsu a ovaria (n = 180)	72 (40,0)	29 (16,1)	3 (1,7)	3 (0,6)	107 (59,4)
karcinom ovaria a jiný nádor mimo karcinom prsu (n = 60)	6 (10,0)	3 (5,0)	1 (1,7)	5 (8,3)	15 (25,0)
Dle rodinné anamnézy					
pouze karcinom ovaria v rodině (n = 102)	43 (42,2)	10 (9,8)	0	2 (2,0)	55 (53,9)
karcinom prsu a ovaria v rodině (n = 288)	86 (29,9)	33 (11,5)	8 (2,8)	2 (0,7)	129 (44,8)
karcinom ovaria a jiný nádor mimo karcinom prsu v rodině (n = 143)	31 (21,7)	11 (7,8)	1 (0,7)	2 (1,4)	45 (31,5)
bez rodinné anamnézy (n = 495)	48 (9,7)	21 (4,2)	18 (3,6)	3 (0,6)	90 (18,2)
není k dispozici (n = 29)	2	0	2	0	4
Dle histologického typu					
high-grade serózní (n = 357)	88 (23,5)	35 (9,3)	12 (3,2)	3 (0,8)	150 (40,0)
serózní bez specifikovaného grade (n = 138)	35 (25,4)	8 (5,8)	5 (3,6)	0	50 (36,2)
low-grade serózní (n = 75)	9 (12,0)	3 (4,0)	1 (1,3)	0	13 (17,3)
endometrioidní (n = 77)	14 (18,2)	2 (2,6)	4 (5,2)	2 (2,6)	26 (36,4)
mucinózní (n = 38)	3 (8,6)	2 (5,7)	0	0	7 (20,0)
ze světlých buněk (n = 12)	1 (8,3)	0	0	1 (8,3)	2 (16,7)
jiný maligní histologický typ (n = 81)	2 (2,5)	1 (1,2)	1 (1,2)	1 (1,2)	9 (11,1)
border-line tumors (n = 121)	34 (28,1)	7 (5,8)	3 (2,5)	0	47 (38,8)
není k dispozici (n = 143)	24 (16,8)	17 (14,0)	3 (2,5)	2 (1,4)	53 (37,1)

proLékaře.cz | 28.9.2020

znamenali vyšší zastoupení endometroidních nádorů a nádorů ze světlých buněk.

Nejnižší průměrný věk v době diagnózy karcinomu ovaria byl u nosiček vzácných mutací v genu *MSH2* (42,2 roku), následovaly geny *RAD51C* (49,2 roku), *MLH1* (49,9 roku), *BRCA1* (50,3 roku), *RAD51D* (53,9 roku), *BRIP1* (54,4 roku), *BRCA2* (57,6 roku) a *MSH6* (62,0 roku). Výskyt patogenních mutací v uvedených genech byl překvapivě více než trojnásobný u pacientek starších 70 let (14/68; 21,0 %) v porovnání s pacientkami mladšími 30 let (4/75; 6,7 %; p = 0,01).

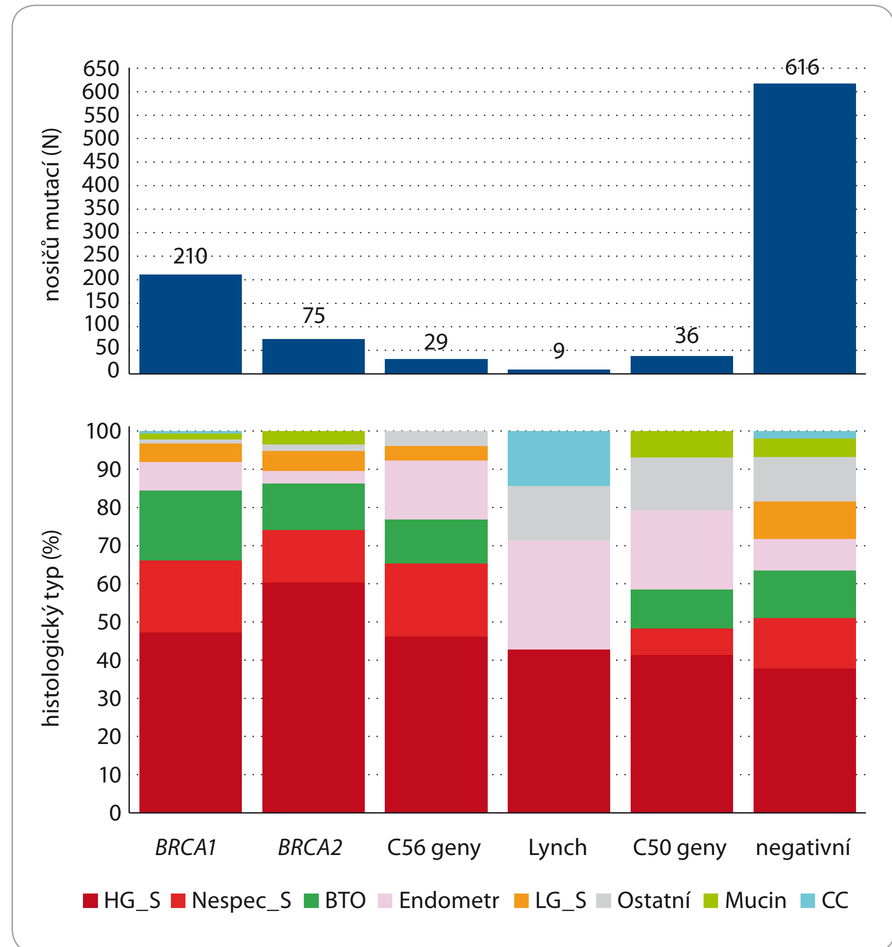
Mutace v genech predisponujících ke vzniku dědičné formy karcinomu prsu (*ATM, BARD1, CHEK2, NBN, PALB2*)

V dalších genech, jejichž mutace jsou spojeny se zvýšeným rizikem vzniku karcinomu prsu a kde bychom mohli předpokládat rovněž zvýšení rizika rozvoje karcinomu ovaria, jsme zachytili patogenní mutace celkem u 36/1 057 (3,4 %) pacientek a 7/617 (1,1 %) kontrol (celkem p = 0,0037; tab. 4).

Nejvyšší frekvence mutací v těchto genech jsme nalezli ve skupině pacientek s diagnózou karcinomu prsu a ovaria v osobní anamnéze (11/180; 6,1 %) a u pacientek z rodin s výskytem pouze karcinomu ovaria (6/102; 5,9 %) (tab. 3). Nejnižší věk v době diagnózy karcinomu ovaria byl u pacientek s mutacemi v genu *CHEK2* (41,8 roku), následován geny *ATM* (49,6 roku), *NBN* (51,1 roku), *BARD1* (54,7 roku) a *PALB2* (61,4 roku).

V dalších genech spojovaných se syndromem dědičného karcinomu prsu a ovaria – *NF1, PTEN, STK11, TP53, CDH1* – jsme v našem souboru nezachytili žádnou patogenní alteraci.

Více než jedna patogenní mutace v genech spojených se syndromem dědičného karcinomu prsu a ovaria byla nalezena u 7 pacientek (0,66 %, což zahrnovalo heterozygotní nosičky mutací v *BRCA1* a *BRCA2* (2×), *BRCA1* a *NBN* (2×), *BRCA1* a *CHEK2*, *BRCA1* a *ATM*, *BRCA1* a *MLH1*).



Graf. 1. Zastoupení histologických typů u nosiček mutací v predispozičních genech.

Tab. 4. Záchyt mutací v genech predisponujících ke karcinomu prsu.

	Pacientky n = 1 057 (%)	Kontroly n = 617 (%)	
NBN	13 (1,2)	2 (0,3)	
CHEK2	8 (0,8)	1 (0,2)	
PALB2	6 (0,6)	2 (0,3)	
BARD1	5 (0,5)	0	
ATM	4 (0,4)	2 (0,3)	
celkem	36 (3,4)	7 (1,1)	p = 0,0037

Diskuze a závěr

Rozvoj NGS umožňuje paralelní analýzu germinálních mutací řady nádorových predispozičních genů. U karcinomu ovaria jsou s prokázaným vysokým rizikem onemocnění (relativní riziko (RR) > 5) v současné době spojeny mutace v genech *BRCA1, BRCA2, BRIP1, RAD51C, RAD51D, STK11, MSH2* a *MLH1*. Analýza může být provedena zároveň u několika desítek pacientů, což vede ke zrychlení diagnostického procesu. Vyšetření rozsáhlých populací indikovaných pacientů umožňuje získat dostatečné informace o frekvenci výskytu příčinných mutací v populaci, které jsou nezbytným předpokladem k postupnému zpřesňování rizik spojených se vznikem dědičných nádorů u nosičů mutací v nádorových predispozičních genech. Výskyt mutací v některých nově identifikovaných pre-

proLékaře.cz | 28.9.2020

dispozičních genech (*BRIP1, RAD51C, RAD51D* a geny Lynchova syndromu) je však velmi vzácný a k vyhodnocení jejich podílu na vzniku karcinomu ovaria je nezbytná spolupráce diagnostických týmů v ČR i zahraničí. Na základě takto získaných poznatků se dynamicky vyvíjejí nejen indikační kritéria ke genetickému vyšetření, ale i klinická doporučení péče o nosiče patogenních mutací [11].

Z provedených analýz u pacientek s ovariálním karcinomem v ČR vyplývá, že mezi nosiči příčinných mutací v genech s jasně prokázanou asociací se zvýšeným rizikem vzniku ovariálního karcinomu dominují nosičky mutací *BRCA1* a *BRCA2* (88 %). Zbývajících nezanedbatelných 12 % případů v naší populaci představují nosičky mutací v dalších ovariálních predispozičních genech (*BRIP1, RAD51C, RAD51D* a geny Lynchova syndromu). Z výsledků naší analýzy vyplývají určitá populační specifika v zastoupení mutací v těchto nových predispozičních genech v ČR. Zatímco v naší populaci bylo zastoupení mutací v těchto genech rovnoměrné, práce Norquista et al (USA) popisuje dvojnásobnou frekvenci mutací v genu *BRIP1* (26/1 915; 1,3 %) v porovnání s četností mutací v genech *RAD51C* a *RAD51D* (každý 11/1 915; 0,6 %) [12]. V německé populaci byla popsána frekvence mutací v genu *BRIP1* u pacientek s karcinomem ovaria dokonce 2,6 %, přestože v této studii nebyly hodnoceny velké genomové přestavby [13]. Ve skupině genů spojených s Lynchovým syndromem jsme nejvíce mutací nalezli v genu *MLH1* (gen *PMS2* nebyl vyšetřován), podobně jako je tomu u pacientů s kolorektálním karcinomem, zatímco v některých pracích jsou u pacientek s karcinomem ovaria popisovány mutace především v genu *MSH6* a *PMS2* [12]. Porovnání výsledků našich analýz s výsledky zahraničních studií však komplikuje skutečnost, že některé studie neanalyzují přítomnost velkých genomových přestaveb v zárodečném genomu, protože jejich identifikace pomocí NGS může být obtížná. V naší práci jsme tyto důležité (a v případě delecí v naprosté většině i patogenní) genetické aberace potvrdili z vyšetření CZECANCA panelem, který byl optimalizován na detekci velkých přestaveb [6], u 3,1 % (33/1 057) pacientek, a to v genech *BRCA1* (22), *BRIP1* (4), *MLH1* (2), *CHEK2* (4) a *PALB2* (1). V genu *BRIP1* představovaly rozsáhlé delece dokonce polovinu všech nalezených mutací. Četnost velkých genomových přestaveb ukazuje, že jejich analýza musí být nepostradatelnou součástí genetického testování nádorových predispozičních genů.

Podíl alterací v dalších kandidátních predispozičních genech s možným podílem na vzniku karcinomu ovaria (*ATM, BARD1, CHEK2, NBN, PALB2*) je srovnatelný s celkovou frekvencí mutací v genech *BRIP1*, *RAD51C*, *RAD51D* a genů Lynchova syndromu.

Záchyt sedmi vícenásobných nosičů patogenních variant v našem souboru ukazuje důležitost genetické konzultace a v indikovaných případech opodstatňuje nové testování pacientek pomocí NGS (u probandů z rodin se závažnou rodinnou onkologickou anamnézou či u pacientek s nádorovými multiplicitami negativně testovaných dříve použitými metodami). Genetická konzultace a následné vyšetření hlavních predispozičních genů jsou indikovány dle současných kritérií pro každou pacientku s karcinomem ovaria, bez ohledu na věk, osobní či rodinnou anamnézu či histologický typ nádoru. Toto doporučení podporují i výsledky naší práce, která analyzuje 1 057 pacientek s karcinomem ovaria, což přibližně odpovídá roční incidenci tohoto onemocnění v ČR (998 případů v roce 2016). Patogenní mutace v genech s jasným klinickým významem pro karcinom ovaria jsme zachytili u 323/1 057 (30,6 %) vyšetřovaných pacientek. Srovnatelnou frekvenci mutací (24 %) popsal Walsh et al [14], avšak v některých studiích byla nalezená frekvence mutací v ovariálních predispozičních genech značně nižší (např. Carter et al 13 %) [15]. Rozdíly ve frekvenci mutací u pacientek s karcinomem ovaria mezi jednotlivými pracemi ilustrují nejen odlišnosti ve výskytu patogenních mutací mezi populacemi, ale i ve výběru analyzovaných pacientek, vyšetřovaných predispozičních genů či typu hodnocených mutací. Náš soubor byl obohacen o mladé pacientky a pacientky s pozitivní rodinnou anamnézou. Zatímco frekvence mutací u nemocných před rokem 2010 přesahovala 40 %, po rozšíření indikačních kritérií v roce 2015 [16–19] bylo u 334 pacientek analyzovaných na 1. LF UK zachyceno 83 nosiček mutací (25 %). U neselektovaných, prospektivně testovaných vzorků analyzovaných ve spolupracujících laboratořích zahrnutých v našem souboru pacientek byly zachyceny patogenní mutace u 33/156 (21,2 %) nemocných. Lze tedy obecně konstatovat, že přibližně každá čtvrtá pacientka s karcinomem ovaria v naší populaci je nosičkou patogenní mutace v klinicky významném genu s prokázaným vysokým rizikem vzniku ovariálního karcinomu. U pacientek s negativní rodinnou anamnézou je to přibližně každá pátá. Vysoký podíl pacientek s karcinomem ovaria s hereditární formou onemocnění se do budoucna díky identifikaci mutací v dalších genech s predispozicí ke karcinomu ovaria patrně ještě mírně zvýší. Frekvence mutací je u pacientek starších 70 let více než třikrát vyšší v porovnání s velmi mladými pacientkami (s diagnózou ve věku pod 30 let). Nižší frekvenci patogenních mutací u velmi mladých pacientek lze částečně vysvětlit vyšším zastoupením histologických typů, především low-grade serózních a mucinózních karcinomů, pro které je typická nižší frekvence mutací v predisponujících genech, zatímco u pacientek starších 70 let převažovaly high-grade serózní karcinomy. Přesto byla nízká frekvence mutací ve sledovaných predispozičních genech u velmi mladých pacientek překvapivá a ukazuje na možnou úlohu dalších genů při vzniku onemocnění.

Vzhledem k omezeným možnostem časné diagnostiky karcinomu ovaria je pro nosičky mutací v genech spojených s vysokým rizikem onemocnění (RR > 5) doporučena preventivní salpingooforektomie (risk-reducing salpingo-ophorectomy – RRSO). RRSO je spojena s řadou vedlejších účinků, proto je nutné její vhodné načasování, které se odvíjí od průměrného věku v době diagnózy u nosiček mutací či podle věku onemocnění v rodině (tab. 5). S ohledem na frekvenci mutací v genech vysokého rizika u pacientek s ovariálním karcinomem v ČR může racionální indikace RRSO přispět ke snížení mortality z důvodu karcinomu ovaria v ČR.

proLékaře.cz | 28.9.2020

Serózní adenokarcinomy představují 60–80 % ze všech histopatologických typů ovariálních tumorů [20]. V našem souboru jsme high-grade serózní karcinom ovaria nalezli u dvou třetin pacientek s identifikovanou zárodečnou mutací v některém z ovariálních predispozičních genů, s výjimkou genů Lynchova syndromu. Frekvence mutací u žen s nádory jiného histologického typu však byla také významná a opodstatňuje genetické testování bez ohledu na histologický typ nádoru.

Se vzrůstajícím množstvím poznatků o nádorové predispozici se rozšiřuje spektrum klinicky relevantních genů. Použití větších panelů genů v rutinní diagnostické praxi umožňuje nejen pružně reagovat na rostoucí požadavky onkologů a klinických genetiků, ale v případě rozšíření počtu genů s klinickou utilitou umožňuje vyhodnotit tyto geny zpětně bez nutnosti nového sekvenování, a tedy dodatečných finančních nákladů. Identifikace hereditárních alterací v genech *BRCA1* a *BRCA2* (a pravděpodobně i dalších predispozičních genech, jejichž proteinové produkty se spolupodílejí na reparaci genomové DNA) umožňuje využít genetické analýzy jako prognostického ukazatele pro konvenční chemoterapii (genotoxickými chemoterapeutiky) i cílenou a specifickou léčbu (PARP inhibitory) [21].

Tab. 5. Existující doporučení péče o nosičky patogenních mutací v genech predisponujících ke karcinomu ovaria [4].

Gen	Riziko vzniku karcinomu ovaria**	Doporučení RRSO
*ATM**	?	dle RA
BRCA1	35–60 %	RRSO (35–40let)
*BRCA2**	12–25 %	RRSO (40–45let)
*BRIP1**	10–15 %	RRSO (45–50let)
BARD1	?	dle RA
CHEK2	?	dle RA
MSH2	15–24 % do 70 let	dle RA
MLH1	11–20 % do 70 let	dle RA
MSH6	?	dle RA
*PMS2**	?	dle RA
EPCAM	?	dle RA
*NBN**	?	dle RA
*PALB2**	?	dle RA
*RAD51C**	10–15 %	RRSO (45–50let)
RAD51D	10–15 %	RRSO (45–50let)
STK11	18–21 %	dle RA

*AR sy, **populační riziko 1,3 %
RRSO – riziko redukující salpingooforektomie, RA – rodinná anamnéza

Literatura

1. uzis.cz. Ústav zdravotnických informací a statistiky ČR. [online]. Dostupný z http://www.uzis.cz.
2. Pennington KP, Swisher EM. Hereditary ovarian cancer: beyond the usual suspects. Gynecol Oncol 2012; 124(2): 347–353. doi: 10.1016/j.ygyno.2011.12.415.
3. Antoniou A, Pharoah PD, Narod S et al. Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case Series unselected for family history: a combined analysis of 22 studies. Am J Hum Genet 2003; 72(5): 1117–1130. doi: 10.1086/375033.
4. nccn.org. National Comprehensive Cancer Network; version 3.2019. [online]. Available from: https://www.nccn.org/professionals/physician_gls/pdf/genetics_screening.pdf.
5. Soukupová J. Úskalí interpretace sekvenačních dat v diagnostice dědičných nádorových syndromů. Labor Aktuell 2016; 20(4): 4.
6. Soukupová J, Zemánková P, Lhotová K et al. Validation of CZECANCA (CZEch CAncer paNel for Clinical Application) for targeted NGS-based analysis of hereditary cancer syndromes. PLoS One 2018; 13(4): e0195761. doi: 10.1371/journal.pone.0195761.
7. Soukupová J, Zemánková P, Kleiblová P et al. CZECANCA: CZEch CAncer paNel for Clinical Application – návrh a příprava cíleného sekvenačního panelu pro identifikaci nádorové predispozice u rizikových osob v České republice. Klin Onkol 2016; 29 (Suppl 1): S46–S54. doi: 10.14735/amko2016S46.
8. internationalgenome.org. The 1000 Genomes Browser. [online]. Available from: http://www.internationalgenome.org/1000-genomes-browsers.
9. enigmaconsortium.org. ENIGMA (Evidence-based Network for the Interpretation of Germline Mutant Alleles). [online]. Available from: https://enigmaconsortium.org/.
10. ncbi.nlm.nih.gov. ClinVar. [online]. Available from: https://www.ncbi.nlm.nih.gov/clinvar/.
11. Nielsen SM, Eccles DM, Romero IL et al. Genetic testing and clinical management practices for variants in non-BRCA1/2 breast (and breast/ovarian) cancer susceptibility genes: an international survey by the Evidence-Based Network for the Interpretation of Germline Mutant Alleles (ENIGMA) clinical working group. J Clin Oncol 2018; 36(2): 1–42.
12. Norquist BM, Harrel MI, Brady MF et al. Inherited mutations in women with ovarian carcinoma. JAMA Oncol 2016; 2(4): 482–90. doi: 10.1001/jamaoncol.2015.5495.
13. Weber-Lassalle N, Hauke J, Ramser J et al. BRIP1 loss-of-function mutations confer high risk for familial ovarian cancer, but not familial breast cancer. Breast Cancer Res 2018; 20(1): 7. doi: 10.1186/s13058-018-0935-9.
14. Walsh T, Casadei S, Lee MK et al. Mutations in 12 genes for inherited ovarian, fallopian tube, and peritoneal carcinoma identified by massively parallel sequencing. Proc Natl Acad Sci USA 2011; 108(44): 18032–18037. doi: 10.1073/pnas.1115052108.
15. Carter NJ, Marshall ML, Susswein LR et al. Germline pathogenic variants identified in women with ovarian tumors. Gynecol Oncol 2018; 151(3): 481–488. doi: 10.1016/j.ygyno.2018.09.030.
16. Foretová L, Macháčková E, Palácová M et al. Doporučení rozšíření indikačních kriterií ke genetickému testování mutací v genech BRCA1 a BRCA2 u hereditárního syndromu nádorů prsu a ovarií. Klin Onkol 2016; 29 (Suppl 1): S9–S13. doi: 10.14735/amko2016S9.
17. Plevová P, Novotný J, Petráková K et al. Syndrom hereditárního karcinomu prsu a ovarií. Klin onkol 2009; 22 (Suppl 1): S8–S11.
18. Petráková K, Palácová M, Schneiderová M et al. Syndrom hereditárního karcinomu prsu a ovarií. Klin Onkol 2016; 29 (Suppl 1): S14–S21. doi: 10.14735/amko2016S14.
19. Zikán M. Gynekologická prevence a gynekologické aspekty péče u nosiček mutací genů BRCA1 a BRCA2. Klin Onkol 2016; 29 (Suppl 1): S22–S30. doi: 10.14735/amko2016S22.
20. Levanon K, Crum C, Drapkin R. New insights into the pathogenesis of serous ovarian cancer and its clinical impact. J Clin Oncol 2008; 26(32): 5284–93. doi: 10.1200/JCO.2008.18.1107.
21. Morgan RD, Clamp AR, Evans DG et al. PARP inhibitors in platinum-sensitive high-grade serous ovarian cancer. Cancer Chemother Pharmacol 2018; 81(4): 647–658. doi: 10.1007/s00280-018-3532-9.