Gynecologic Oncology 153 (2019) 452–462

Contents lists available at ScienceDirect

Gynecologic Oncology

journal homepage: www.elsevier.com/locate/ygyno

Review Article

Large-scale meta-analysis of mutations identified in panels of breast/ovarian cancer-related genes — Providing evidence of cancer predisposition genes

Malwina Suszynska [a], Katarzyna Klonowska [a,1], Anna J. Jasinska [a,b], Piotr Kozlowski [a,*]

[a] Department of Molecular Genetics, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznan, Poland
[b] Center for Neurobehavioral Genetics, Semel Institute for Neuroscience & Human Behavior, University of California, Los Angeles, Los Angeles, CA, USA

HIGHLIGHTS

- 37 genes were evaluated in the context of breast and ovarian cancer predisposition.
- Several non-BRCA1/2 genes were attributed to high breast/ovarian cancer risk.
- Mutations in *CDKN2A* contribute to high risk of breast cancer, comparable to *BRCA2*.
- Profiles of genes contributing to breast and ovarian cancer differ substantially.
- *RAD51C*, *RAD51D*, and *BRIP1* are proved to be high-risk ovarian, but not breast cancer genes.

ARTICLE INFO

Article history:
Received 8 October 2018
Received in revised form 23 January 2019
Accepted 30 January 2019
Available online 4 February 2019

Keywords:
Breast cancer
Ovarian cancer
Multi-gene panel testing
Meta-analysis
Cancer predisposition genes

ABSTRACT

Objective. Germline mutations occurring in the highly penetrant genes *BRCA1* and *BRCA2* are responsible for only certain cases of familial breast cancer (BC) and ovarian cancer (OC). Thus, the use of NGS multi-gene panel (MGP) testing has recently become very popular.

Methods. To estimate a reliable BC and OC risk associated with pathogenic variants in the selected candidate BC/OC predisposition genes, a comprehensive meta-analysis of 48 MGP-based studies analyzing BC/OC patients was conducted. The role of 37 genes was evaluated, comparing, in total, the mutation frequency in ~120,000 BC/OC cases and ~120,000 controls, which guaranteed strong statistical support with high confidence for most analyzed genes.

Results. We characterized the strategies of MGP analyses and the types and localizations of the identified mutations and showed that 13 and 11 of the analyzed genes were significantly associated with an increased BC and OC risk, respectively. The risk attributed to some of these genes (e.g., *CDKN2A* and *PALB2* for BC) was similar to that observed for *BRCA2*. The analysis also showed a substantial difference in the profile of genes contributing to either BC or OC risk, including genes specifically associated with a high risk of OC but not BC (e.g., *RAD51C*, and *RAD51D*).

Conclusions. Our study provides strong statistical proof, defines the risk for many genes often considered candidates for BC/OC predisposition and excludes the role of other genes frequently analyzed in the MGPs. In the context of clinical diagnostics, the results support the knowledge-based interpretation of identified mutations.

Contents

* Corresponding author at: Department of Molecular Genetics, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14 Street, 61-704 Poznan, Poland.
E-mail address: kozlowp@ibch.poznan.pl (P. Kozlowski).
[1] Present address: Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA.

https://doi.org/10.1016/j.ygyno.2019.01.027

1. Introduction

Germline mutations occurring in highly penetrant genes *BRCA1* and *BRCA2* are responsible for 16–40% of familial breast cancer (BC) and ovarian cancer (OC) cases. Therefore, genetic background of cancer predisposition remains unrecognized in substantial fraction of BC/OC patients.

Recent advances in next-generation sequencing (NGS) technologies, have led to the development of a broad range of NGS multi-gene panels (MGPs) for the genetic testing of various human hereditary diseases, including BC and OC. The results of several MGP-based analyses focused on exon sequences of the additional risk candidate genes performed in groups of BC/OC patients have already been published. The analysis has led to the identification of various deleterious or potentially deleterious mutations causing premature termination of translation, alternative splicing or disruption of protein structure/function. However, unequivocal evidence for the role of a significant fraction of BC/OC candidate genes routinely analyzed within diagnostic MGPs is still lacking. Consequently, the practical understanding and the clinical/diagnostic utility of the substantial fraction of the identified mutations are still very limited. The lack of statistically significant and convincing proof for the role of specific candidate genes in BC predisposition may result from (i) the low frequency of mutations occurring in candidate genes (predominantly ≪ 1%), (ii) the limited number of analyzed subjects within particular studies (most studies comprise groups of 100–1000 subjects) and/or (iii) the absence of data for matched population controls to inform the interpretation of the results obtained in groups of BC/OC subjects. It has to be noted, however, that two studies published very recently encompassed very large groups of BC subjects [1,2].

In this study, we conducted a comprehensive meta-analysis, considering 48 published MGP-based analyses performed in groups of subjects with BC/OC. Our meta-analysis comprised 37 genes most frequently included in the MGPs, developed for either clinical or research applications. We characterized the type of utilized approaches, types and locations of identified mutations, and cumulative frequency of mutations in particular genes in BC and OC patients. Most importantly, taking advantage of a large collection of data that cumulatively included ~120,000 cases (either BC or OC), we performed a large-scale association analysis, comparing the frequencies of mutations in the cases with the frequencies of mutations in a group of ~120,000 controls (derived from the gnomAD database) [3]. Due to the unprecedented large scale, the analysis allowed a high statistical power to estimate with high confidence the risk attributed to most analyzed genes in the general (mixed) population. The overall results held when the analysis was performed in separate ethnic groups (e.g., European/White). Among the most striking results is the identification of several genes associated with a high risk of BC at a level similar to the *BRCA2* risk. The other striking observation of our study is a substantial difference in the profile of genes contributing to either BC or OC risk. The results allow a better understanding of the genetic predisposition to BC/OC but also have practical implications for BC/OC genetic testing and the informed interpretation of detected mutations.

2. Methods

We used the PubMed database to search articles published until July 2017, which reported the genetic test results of BC and OC patients referred for evaluation by an MGP. To achieve this aim, we used the following search string: “breast cancer AND multi-gene panel test/NGS” or “ovarian cancer AND multi-gene panel test/NGS”. The two key inclusion criteria were as follows: (i) BC/OC patients and (ii) MGPs as a mutation detection method. We excluded studies concerning only the *BRCA1/2* genes. The review included unselected BC and OC studies that were not restricted to only high-risk individuals (familial, bilateral, or early-onset BC). Screening of eligible studies and selection of articles to be included in our meta-analysis were independently performed by 2 reviewers (MS and KK). In total, we selected 48 published studies for analysis. The studies reporting results from a single cohort, as well as cumulative studies combining the data from several cohorts, were taken into account. In cases of patients analyzed in more than one study (when it was clearly indicated), repeatable data were not used.

The genetic test results of BC/OC patients referred for evaluation by an MGP were used to calculate the mutation frequencies in particular genes. We focused on mutations that were defined by authors as loss-of-function variants (frameshift, nonsense, ±1/±2 position splicing mutations), large rearrangements (if analyzed in studies), as well as missense and other intron variants described as pathogenic in the ClinVar database (https://www.ncbi.nlm.nih.gov/clinvar/, as of December 2017) [4], which contains data regarding the clinical significance of variants. Variants classified in several studies as likely pathogenic, variants that were missense mutations for which there was a lack of pathogenicity information in the ClinVar database, or variants that were described as being of uncertain significance/likely benign/benign were excluded from the analysis.

As a control, population mutation frequencies were calculated using data extracted from the online Genome Aggregation Database [gnomAD (http://gnomad.broadinstitute.org/)], which contains sequencing data from >120,000 exomes from unrelated individuals sequenced as part of various disease-specific and population genetic studies [3]. Pathogenic mutations in the control group were classified in a similar way as in the BC and OC groups. Only pathogenic variants obtained from exome sequencing were used, including all loss-of-function variants (frameshift, nonsense, ±1/±2 position splicing mutations) and carefully selected missense variants, classified as Pathogenic/Likely Pathogenic in the ClinVar database. Genetic variants with an allelic population frequency >0.03 (mostly ≫0.03), which was too high to be considered pathogenic, were excluded from the analysis (except for a few well-documented founder mutations). Large rearrangements, such as whole-exon deletions or duplications, were not taken into account in association studies (removed from the BC and OC groups). Double and multiple mutation carriers were not excluded from the case groups because they were not specified in the control group. Nevertheless, the fraction of multiple carriers was negligible and should not affect the risk assessment.

Association analyses were performed for BC and OC patients, using the pooled data from multiple ethnic groups in both cases and control groups. No individual patient data were available for the cases and

controls, as data were extracted from studies and public database. As mutation frequencies may vary in different populations, the calculations were also performed for the European population only. Data from all studies conducted on a European/White population were selected and compared to non-Finnish European populations from the control group. The ethnicity was determined based on the country/origin of publication or based on the prevailing population fraction. If possible, data for different populations were extracted and calculated separately. Several studies conducted in multinational countries (e.g., USA), which did not report a detailed ethnicity, or with the ethnicity of the prevailing population below 75%, were excluded from the association analysis conducted on the European population. Additional meta-analyses were performed for Asian, African, and Latino populations, but they were limited to several genes for the *BC most* group, due to the small number of studies available to include in our analysis.

Associations between mutations in the selected genes and BC and OC risk were assessed using odds ratios (ORs) and 95% confidence intervals (95% CIs) based on a chi-squared test, conducted with MedCalc Statistical Software version 14.8.1 (MedCalc Software bvba, Ostend, Belgium; http://www.medcalc.org). All statistical tests were two-sided, and due to performed multiple comparisons, a p-value of <0.0001 (indicated by **) was considered statistically significant. The nominally significant p-values (0.05–0.0001) are indicated by *. To test for heterogeneity between ethnicities, a meta-analysis was performed in groups of different ethnicities with the use of the Mantel-Haenszel method, calculating the ORs under fixed and random effects models depending on the results of heterogeneity tests ($p > 0.1$ for the Q test and $I^2 < 50\%$ were considered to indicate a lack of significant heterogeneity). To assess study heterogeneity (I^2) and publication bias (funnel plots), we performed additional analysis, comparing independently each study with the gnomAD control.

3. Results

3.1. Characteristics of the selected studies

A retrospective review was performed on the genetic test results of BC and OC patients referred for evaluation by an MGP. A total of 48 systematic studies that passed the selection criteria described in the Methods, were selected by electronically searching the PubMed database. Characteristics of the identified studies are reported in Fig. 1 (A, B, C).

We obtained data for BC and OC from 43 and 15 studies, respectively. In case of studies including patients with different cancers, BC and OC patient data were separated and classified into BC (*BC only*) or OC (*OC only*) groups, respectively. However, in 12 studies, BC could not be separated from other cancers but constituted the majority of samples (over 75%, most frequently, ~90% of patients had BC). Thus, these data were combined with *BC only*, as the *BC most* group.

The number of genes analyzed in the selected studies ranged from a few to several hundred (Fig. 1A). However, the most studies analyzed 10–29 genes. Additionally, the number of patients participating in individual studies varied substantially and ranged from several to several thousand individuals (Fig. 1B). Additionally, two studies comprised >30,000 patients each [1,2]. Usually, one specific MGP per study was used, but several studies gathered data summarizing the results from multiple (2–9) commercially available MGPs comprising combinations of genes (Fig. 1C). In most studies (52%), the MGPs were custom designed. Other studies used commercially available MGPs, such as BROCA, a test developed at the University of Washington; MGPs designed by Ambry Genetics; TruSight Cancer, designed by Illumina; myRisk, designed by Myriad; an MGP designed by GeneDx; and an MGP designed by Color. An additional two studies performed whole-exome sequencing focusing only on several genes associated with BC/OC.

3.2. List of next-generation sequencing multi-gene panels

The full set of genes included in the MGPs, containing >500 genes, is summarized in Table S1. Of those genes, 37 (Fig. 2) were selected for further analysis based on the frequency of recurrence in the studies (10 and more). A few additional genes were chosen on the basis of the total number of analyzed patients (~ at least 1000). In most cases, the selected genes overlap with genes commonly considered BC/OC risk candidates.

3.3. Types of mutations found in the selected studies

The unified list of detected mutations is shown in Table S2. Table S2 provides the description of mutations at both the DNA level and the protein level, as well as information about the type, number of particular mutations and ClinVar pathogenicity status [4]. Table S2 summarizes the data from 39 selected studies (87%), which included detailed information about the specific genomic location of particular mutations and provided sufficient information to be used in our mutation type compilation (i.e., ID of the identified mutations along with the matched type of cancer if many types were tested).

As shown in Fig. 3A, the largest fraction of mutations were frameshift mutations, constituting 45% of all variants, followed by nonsense mutations (27%), missense mutations (15%), and splicing mutations (10%). Large mutations such as large deletions and duplications accounted for ~3% of all variants; however, not all investigators have analyzed this type of mutation.

The analysis of such a large number of mutations led to the identification of a series of peculiar mutations that cannot be unambiguously

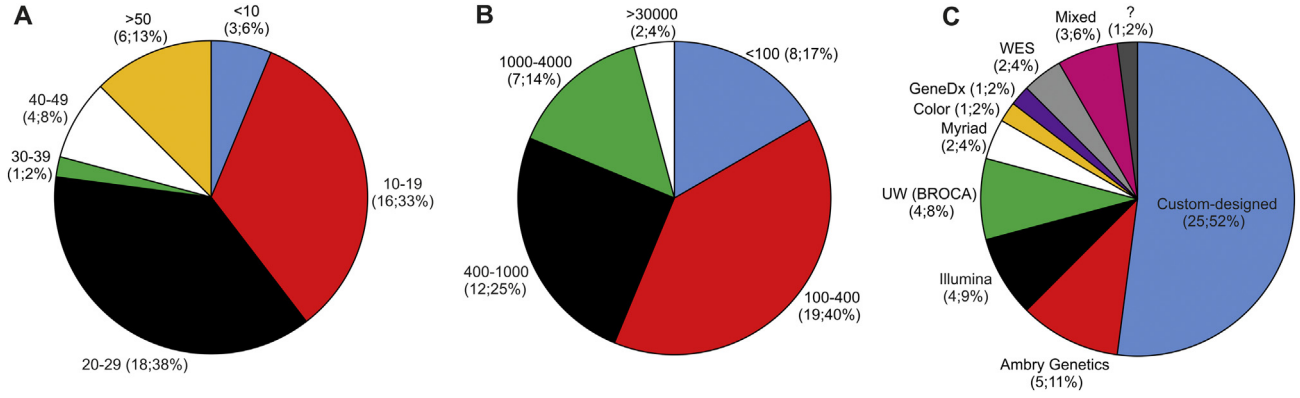


Fig. 1. Characteristics of the BC/OC MGP studies. (A) The number of genes analyzed in the studies. (B) The number of patients analyzed in the studies. (C) Types/brands of the MGPs used in the studies. The values in parentheses reflect the number of studies and the percentage of studies, respectively.

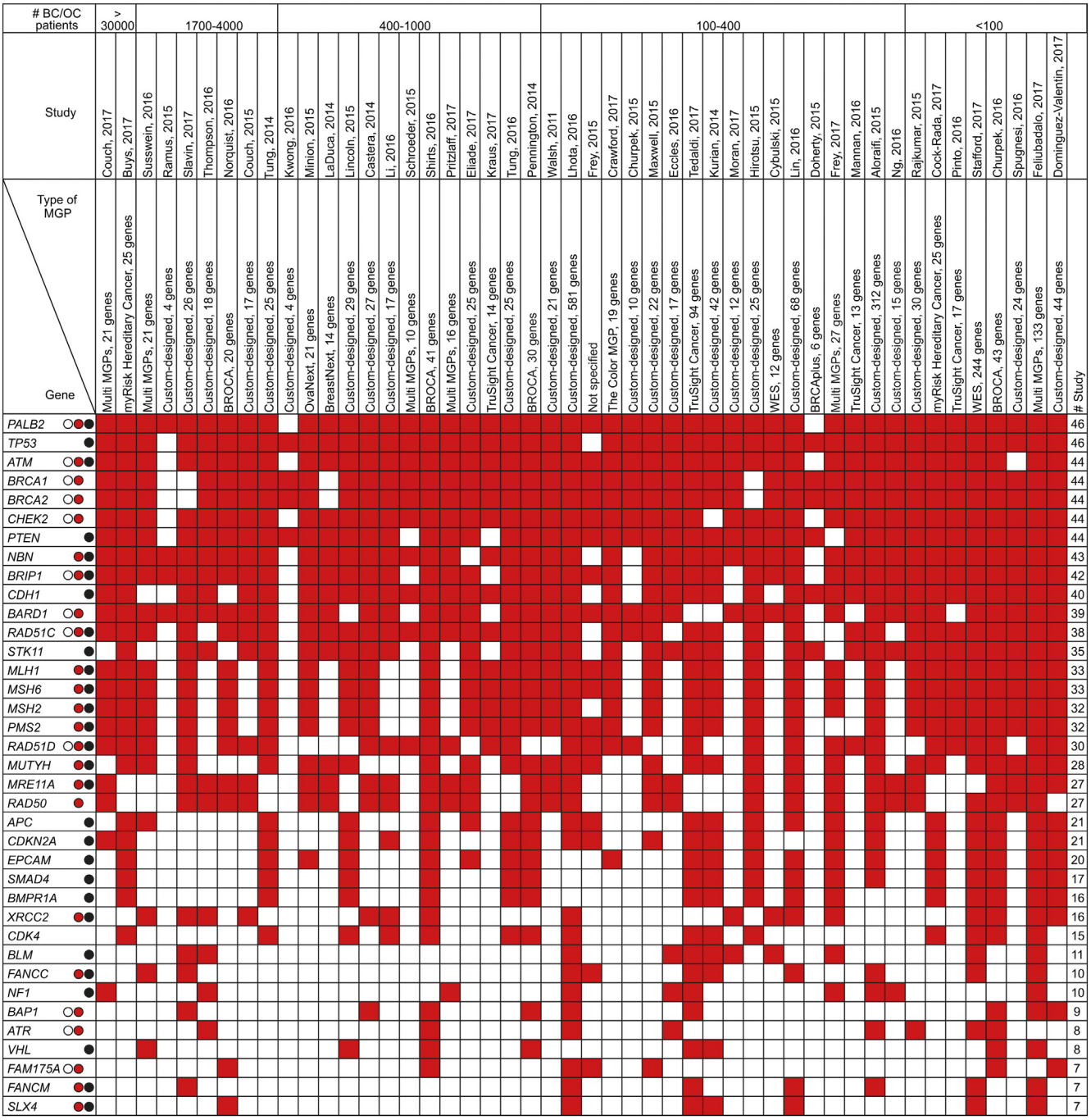


Fig. 2. The genes (rows) most frequently analyzed in BC/OC MGP studies (columns). The table includes 37 (out of ~500) genes selected for the meta-analysis, based on the recurrence in the studies >10, or based on the total number of analyzed patients (>1000). Dots in the first column indicate: *BRCA1* and/or *BRCA2* genes or genes functionally related to the *BRCA1*/*BRCA2* (white dots), genes which primary function related to DNA damage repair (red dots), and genes associated with other hereditary cancer syndromes (black dots). The last column indicates the number of studies that analyzed a particular gene. Please, note that some studies did not analyze all genes from MGP or MGPs (number of analyzed genes is indicated). List of all genes and details are shown in Table S1.

classified to canonical types of mutations. The largest group of atypical mutations are deletions (here designated $FS_{DEL\&N}$, 2%) or duplications ($FS_{DUP\&N}$, 1%) immediately leading to the termination of translation. Another types of peculiar mutation are variants disrupting the start codon (MS_{START}) and "apparent silent mutations" (here designated as =), which do not change AA, but their pathogenic effect most likely results from the effect on the transcript. Because atypical mutations constitute a small portion of all mutations, for simplification, they were classified as either frameshift or missense mutations (see Fig. 3).

The observed types of mutations varied by the analyzed genes (Fig. 3B). In most genes, frameshift and nonsense mutations were predominant. The exception was *TP53*, in which the majority of the mutations (81% of all variants) were missense mutations. A relatively large number of missense mutations have been shown also in the following genes: *CHEK2* (28%), *PMS2* (18%) and *ATM* (13%). Moreover, it is worth mentioning that nonsense mutations were relatively rare in *CHEK2* (only 4%), *NBN* (6%) and *TP53* (9%). Due to the very low number of mutations detected in *MRE11A*, *MSH6*, *PMS2* and *RAD50*, the actual fractions of mutation types in these genes may differ from those shown in Fig. 3B. The exact location of the identified mutations in the most frequently mutated genes is presented in Figs. 4 and S1. As shown, although most mutations are evenly

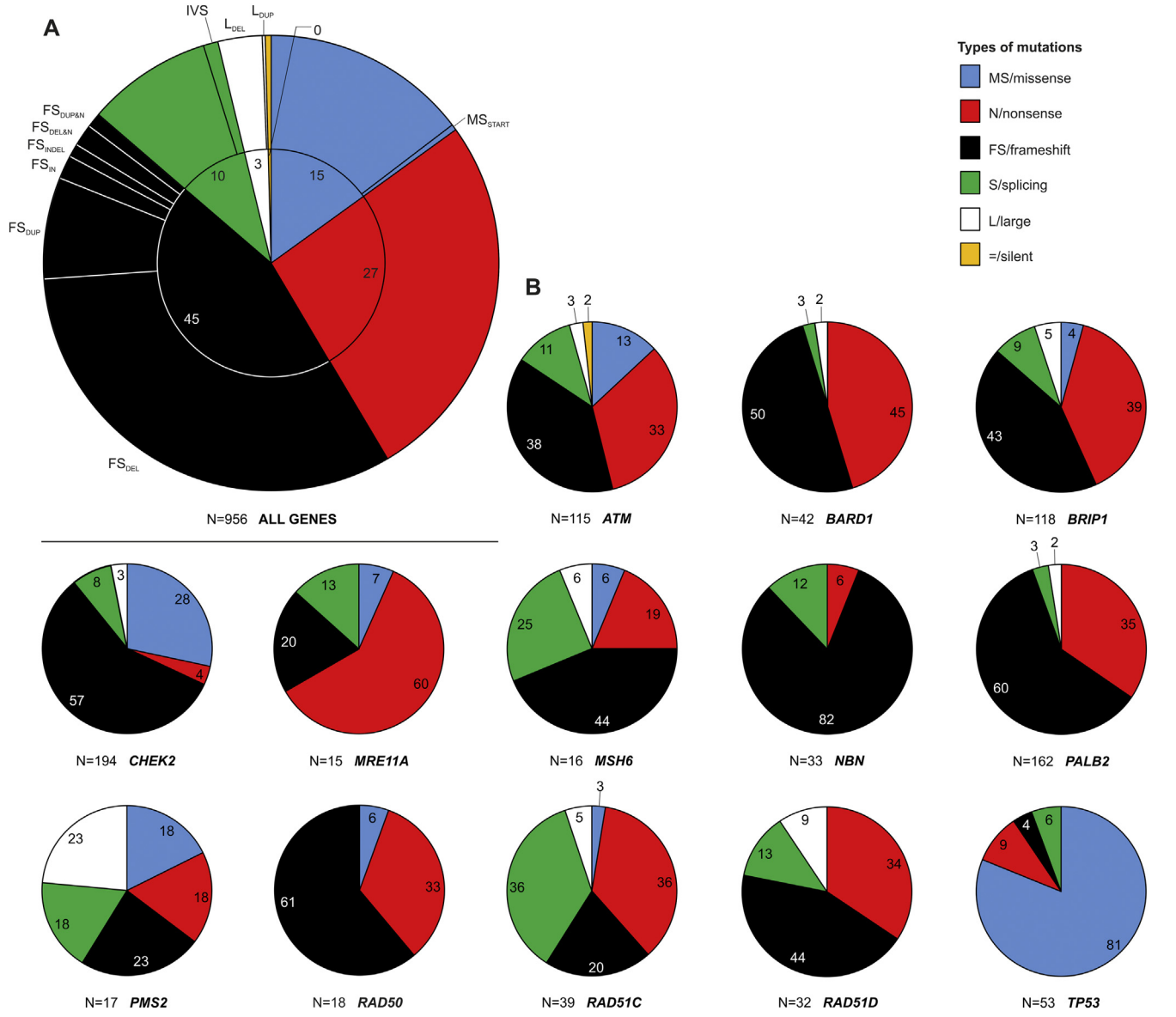


Fig. 3. Mutation types identified in BC and OC patients. (A) Mutation types identified in all genes selected for the meta-analysis. As shown in the legend, the colors are indicated as follows: blue – MS (missense mutations); red – N (nonsense mutations); black – FS (frameshift mutations); green – S (splicing mutations); white – L (large mutations); yellow – =(pathogenic silent mutations). MS include MS_{START} (mutations disrupting the start codon), and S include IVS (pathogenic intronic mutations). L are divided into L_{DEL}/large deletions and L_{DUP}/large duplications; FS are divided into FS_{DEL} (deletions), FS_{DUP} (duplications), FS_{IN} (insertions), FS_{INDEL} (insertions and deletions), $FS_{DEL\&N}$/deletion immediately leading to premature termination of translation, and $FS_{DUP\&N}$/duplication immediately leading to premature termination of translation. (B) Mutation types identified in particular genes (shown for genes with at least 10 identified mutations). Colors reflect the types of mutations, as described in (A). Values in pie-charts indicate the % of mutations. "N" under the pie-charts indicates the total number of mutations. The values are based on Table S2.

distributed along the coding sequences, recurrent mutations ("hypothetical founder mutations") may also be identified. Furthermore, missense mutations in *ATM*, *CHEK2*, *PMS2* and *TP53* clustered predominantly in functionally important domains in the corresponding proteins at positions that showed a high degree of conservation across species.

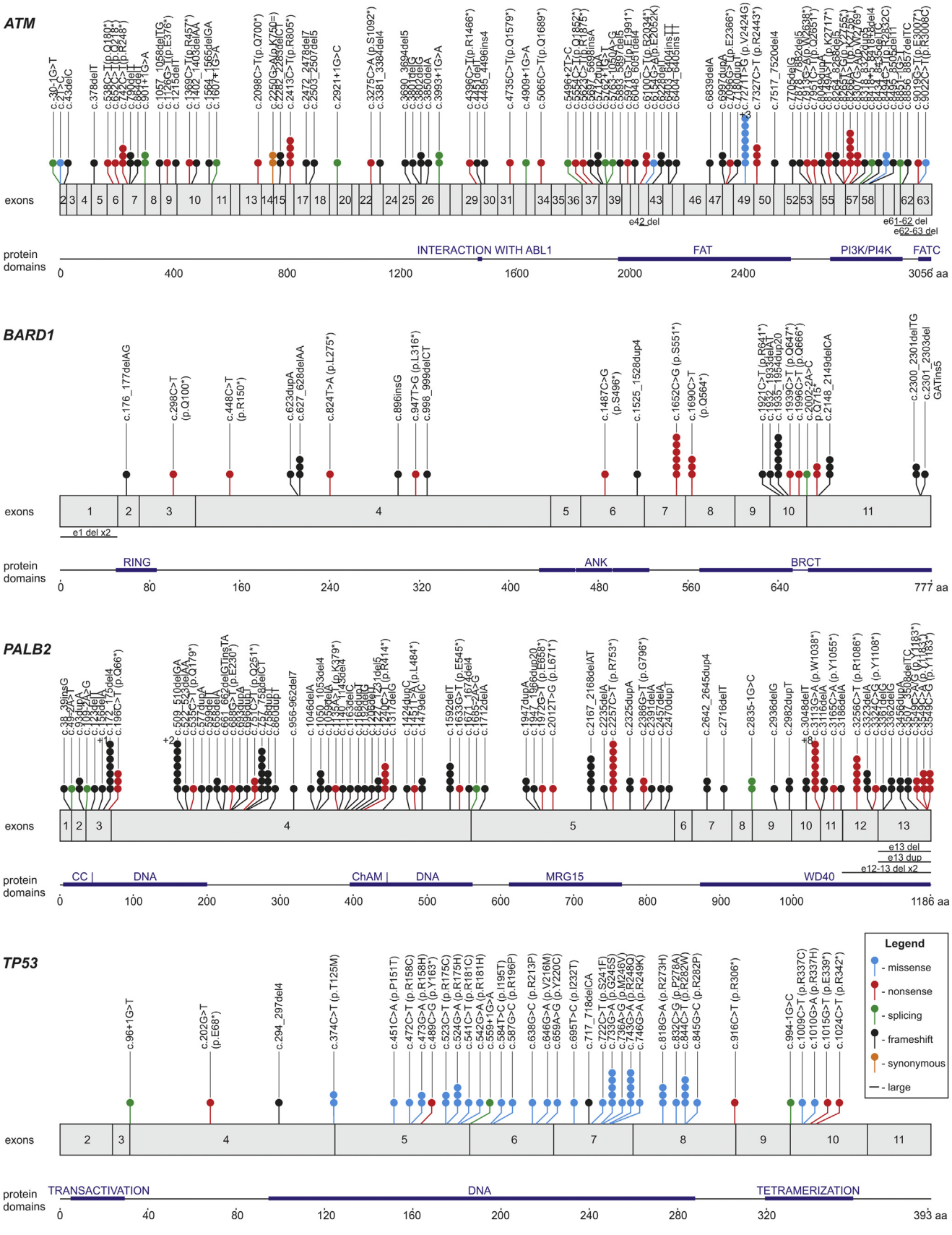
3.4. The frequency of germline mutations in the selected genes among BC and OC patients

The summary and detailed information on the mutation frequency are shown in Fig. 5 and Table S3, respectively.

Overall, ~11,250 mutations were identified in ~120,000 patients. The combined frequency of the identified mutations in the BC (*BC most*) and OC (*OC only*) patients was 12.2% and 21.1%, respectively. After exclusion of the *BRCA1* and *BRCA2* genes, the mutation frequency decreased to 7.7% and 8.0%, respectively. Among 106,529 BC patients (*BC most*) and 3230 OC patients, 4774 (4.48%) and 424 (13.1%) *BRCA1*/*BRCA2* mutations were found, respectively.

Except for *BRCA1* and *BRCA2*, the most frequently mutated genes in BC were *CHEK2* (1.7%), *ATM* (1.0%), and *PALB2* (0.9%). There were no mutations found in *BMPR1A* (39,418 patients analyzed), *CDK4* (40,179 patients), *BAP1* (3622 patients), and *FAM175A* (1152 patients). The mutation frequencies in the *BC most* group was almost the same as that in the *BC only* group, what indicated little/no effect of admixtures of other cancer types in *BC most*. In OC, the most frequently mutated genes were *BRIP1* (1%), *CHEK2* (0.7%), *ATM* (0.7%), *RAD51C* (0.6%), and *RAD51D* (0.6%).

At first sight, many genes showed differences in the mutation frequencies between the BC and OC groups, which suggests that mutations



in particular genes can be preferential or specific for BC or OC. Mutations in *BRCA1*, *BRCA2*, *BRIP1*, *MSH2*, *MSH6*, *NBN*, *RAD51C*, and *RAD51D*, were significantly more frequent in the OC group. In contrast, mutations in *CHEK2*, *PALB2*, and *SLX4* were more frequent in the BC groups. As shown, the proportion of *BRCA1*:*BRCA2* mutations was much higher in OC (2:1) than in BC (1:1) (Fig. 5).

3.5. Estimation of the BC and OC risk associated with mutations within BC and OC susceptibility genes

The summary of the association analyses regarding the odds ratios (ORs; a measure of the disease risk associated with mutational/genetic variants) and 95% confidence interval (95% CI) calculated for BC and OC are shown next to each other in Fig. 6, to facilitate the direct comparison of the results. The results of association analyses, indicating the OR, 95% CI and the p-value for the particular genes are provided in Table S4.

As expected *BRCA1* (OR 9.251) and *BRCA2* (OR 5.672) are at the forefront of high-risk BC genes. However, less expectedly, four other genes, *CDKN2A* (OR 6.205), *PTEN* (OR 5.396), *PALB2* (OR 4.873) and *TP53* (OR 4.362), turned out to be associated with a high risk of BC, on a level similar to that of *BRCA2*. Additionally, other analyzed genes, such as *CDH1* (OR 2.656), *ATM* (OR 2.420) and *BARD1* (OR 2.334), showed a significant association with BC and were classified as moderate-BC-risk genes. Mutations in several other genes, such as *MSH6*, *CHEK2** (after the exclusion of two common variants c.470T>C and c.1283C>T), *MSH2* and *BRIP1*, slightly increased the risk of BC, reaching an OR value of ~1.5. Mutations in these genes may be classified as low-risk variants. Mutations in genes *XRCC2*, *RAD51D*, *NBN*, *RAD51C*, *APC*, *BLM*, *NF1*, *PMS2*, *CHEK2*, *MRE11A*, *FANCC*, *FANCM*, *MLH1*, and *RAD50* were not associated with increased BC risk, whereas the risk attributed to mutations in (i) *BAP1*, *BMPR1A*, *CDK4*, *FAM175A* and (ii) *ATR*, *EPCAM*, *SLX4*, *SMAD4*, *STK11*, and *VHL* could not be determined due to (i) no or (ii) a very low (<5) number of mutations in BC patients.

To check whether the admixture of other cancer types does not affect the associations in the *BC most* group, we compared these results with the results calculated for the *BC only* group. As shown in Fig. S2 and Table S4, the results from the *BC only* and *BC most* groups were almost identical, and therefore, the *BC most* group could be used as a representative group of BC.

As shown in Fig. 6, mutations in *BRCA1* (OR 35.260), *BRCA2* (OR 11.910), *RAD51D* (OR 7.276), *PTEN* (OR 5.473), *TP53* (OR 5.046), *BRIP1* (OR 4.878), *RAD51C* (OR 4.241), *MSH6* (OR 4.077), and *MSH2* (OR 3.976) can be classified as high-OC-risk variants. Mutations in genes *NBN* (OR 2.166), *PALB2* (OR 2.134) and *ATM* (OR 1.977) were associated with a moderate risk of OC. *FANCC*, *FANCM*, *FAM175A*, *MLH1*, and *BARD1* could be classified as low-moderate OC risk genes, with ORs ranging from 1.4 to 3.4, but these values were not statistically significant, due to the small number of pathogenic variants observed in these genes in OC patients. Genes *CHEK2*, *PMS2*, *MRE11A*, and *RAD50* were not associated with OC risk; however, a very low number of pathogenic variants were found in those genes. As no pathogenic variants were found in *APC*, *ATR*, *BAP1*, *BLM*, *BMPR1A*, *CDH1*, *CDK4*, *CDKN2A*, *EPCAM*, *NF1*, *SLX4*, *SMAD4*, *STK11*, *VHL*, *XRCC2* in OC patients, the OC risk associated with a mutation in these genes could not be estimated.

Another striking observation from this meta-analysis is the substantial difference between BC and OC risk (usually a much higher OC risk) attributed to particular genes. Risk estimates for mutations in *BRCA1*, *BRCA2*, and *MSH6* were much higher for OC. The critical difference was also observed for *BRIP1* and *MSH2* genes, which were classified as low-BC-risk variants, but attributed to a high OC risk. Mutations in *RAD51D*, *RAD51C*, and *NBN* seem to be OC-specific, associated only with OC risk, but not with BC. In contrast, mutations in *PALB2* were attributed to a higher BC risk than OC risk, while pathogenic variants in *CHEK2** and *BARD1* were specific only to the BC risk.

As *BC most* and *OC only* groups represent subjects of mixed ethnicity, which may affect the association analysis, we compared the obtained results with the results of the similar analysis performed for extracted European/White population cases. As shown in Fig. S3A for BC and Fig. S3B for OC (as well as in Table S5), there were no substantial differences between the two analyses, and all the observed differences were within CI of each other. The highest differences in the association of BC risk were observed for *CDKN2A*, *PTEN*, and *RAD51D*, while in the association of OC risk was observed for *RAD51D* and *MSH2*. Nevertheless, it should be mentioned that the European/White population group constituted the majority of studies. Furthermore, for the most frequently mutated genes, we calculated the OR independently for 4 major ethnic groups (European, Asian, African, Latino), as presented in Fig. S4.

As two of the analyzed studies were much larger than the others ($n > 30{,}000$ vs. $n < 4000$), they could have been the main drivers of the final results. Therefore, to estimate the potential bias, we repeated the analysis without these two studies [1,2]. As shown in Fig. S5 (and Table S6), the exclusion of the large studies did not change substantially most of the results. The increased BC risks attributed to mutations in *BRCA1* and *BRCA2*, most likely result from prescreening and exclusion of some of the *BRCA1/2* positive patients from the studies, and the decreased BC risk attributed to mutations in *CDKN2A* results from the fact that *CDKN2A* was rarely tested in other than the two large studies. In consequence, the number of tested samples and the number of detected mutations ($n = 2$) were too small to reliably evaluate the *CDKN2A*-attributed risk.

To asses for study heterogeneity (I^2 value) and potential publication bias, for most candidate genes (with at least 10 detected mutations), we independently compared each individual study with common controls and visualized them on funnel plots. As is shown in Fig. S6 (for BC studies) and Fig. S7 (for OC studies) most of the individual study-specific ORs were within the area representing (pseudo) 95% confidence limits, and I^2 values for most genes were close or below 50%, considered borderline value for heterogeneous studies. It has to be noted however that the study-specific ORs are not truly independent values, therefore this analysis has to be interpreted cautiously. Also, due to the low frequency of mutations occurring in individual genes, most of the studies are too small to reliably estimate the OR value. For genes where the total number of mutations is smaller than ~50, a substantial number of individual studies have 0 or few mutations.

4. Discussion

We accomplished a wide-scale meta-analysis of genes that were repeatedly adopted in MGPs to estimate a reliable BC and OC risk associated with mutations in many tested genes. The association of the majority of these genes with BC/OC has not been clearly determined; some were included in MGPs due to their association with other hereditary cancer syndromes and others due to their functional relation to *BRCA1*/*BRCA2*.

Most of the individual studies (included in the meta-analysis as well as others) are too small to detect a sufficient number of mutations that would allow a reliable evaluation of risk in particular genes, especially with mutations that are frequently incidental and identified within genes related to other cancers. However, it must be mentioned that few cumulative studies that combined results (often from different clinical laboratories and mostly generated for diagnostic purposes) were large enough to detect a substantial number of mutations and calculate the risk attributed to particular genes [1,2,5–7]. Additionally, previously reported risk estimates for some genes differed substantially in various studies. Those conflicting results suggest that genes should not be

Fig. 4. Mutation-gene "maps" showing the location, type, and number of particular mutations (limited to studies with a full list of pathogenic mutations) in the *ATM*, *BARD1*, *PALB2* and *TP53* genes. The structure of particular genes/coding exons and protein domains were created on the basis of data from the Ensembl genome browser and UniProt database. Mutations are marked with dots. The color of the dots reflects the type of mutation, as shown in the legend. The number of dots reflects the identified number of particular mutations (up to 7). The number of mutations above the 7 is indicated above the symbol as +n. Large deletions and duplications are indicated as horizontal lines under the gene maps spanning affected exons. The abbreviations are as follows: FAT (named after representatives of the three main groups sharing the domain FRAP, ATM and TRRAP); PI3K/PI4K (phosphatidylinositol3-/4-kinase, catalytic domain); FATC (FRAP-*ATM*-TRRAP-C-terminal *domain*); RING (zinc finger, RING-type); ANK (ankyrin repeat); BRCT [BRCA1 C-terminal (BRCT) domains]; CC (coiled coil); DNA (DNA-binding); ChAM (chromatin-association motif); MRG15 (interaction with MRG15); WD40 (WD40-repeat-containing domain).

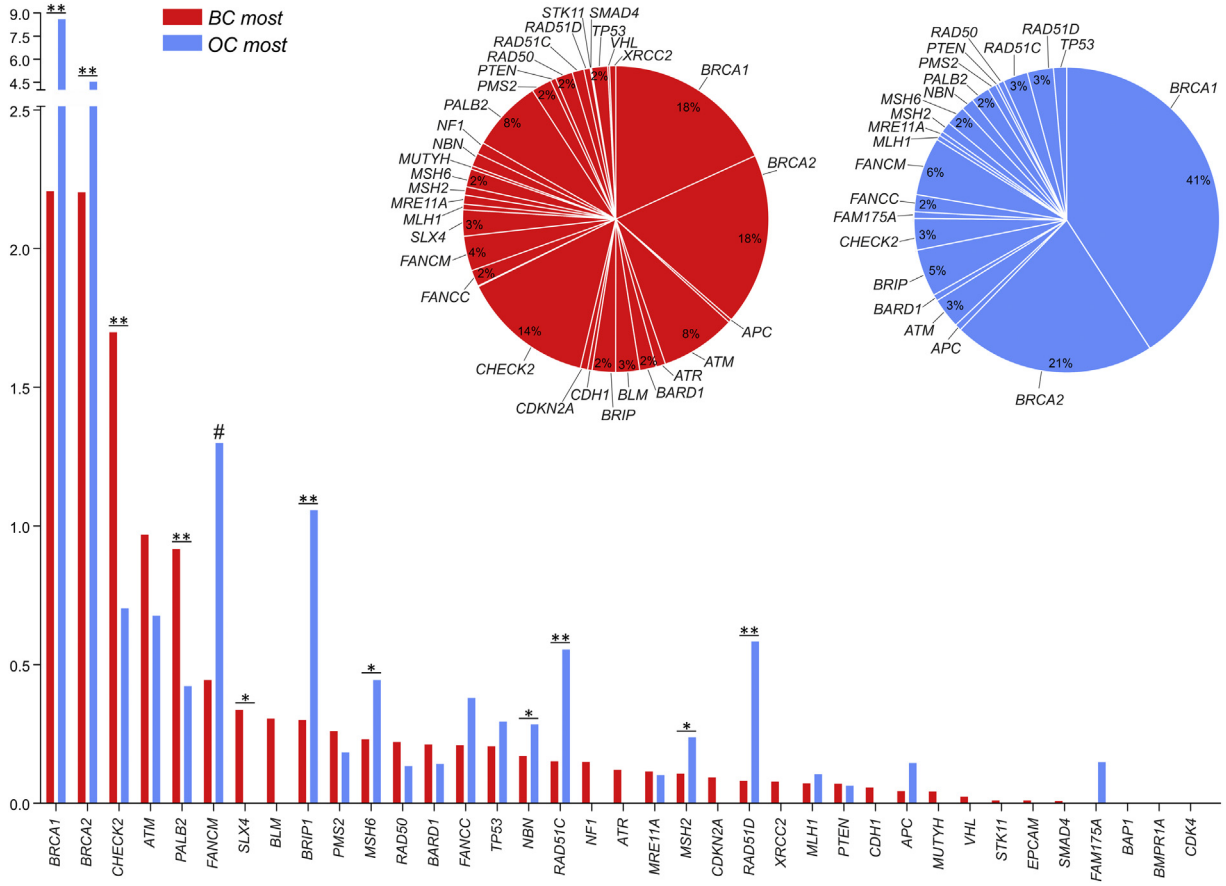


Fig. 5. The frequency of mutations [%] in particular genes in BC and OC patients, calculated as the number of identified mutations per number of analyzed patients. A chi-square test was used to indicate statistically significant differences between BC and OC cases (* and ** reflect $p < 0.05$ and $p < 0.0001$, respectively). Pie-charts indicate the contribution of particular genes to the total number of mutations detected in BC and OC patients. #Due to the very low number of patients analyzed for mutations in *FANCM* in the *OC most* group, the actual frequency of mutation may differ from that reported. The detailed frequency of mutation in particular gene is provided in Table S3.

classified as low-moderate-high BC and OC risk genes in the face of insufficient data, especially with regard to the rarely mutated genes. The above facts substantially hamper the interpretation of mutations being identified in non-*BRCA1/BRCA2* genes and may cause problems with patients consulting and the follow-up treatment.

The summary of the enrolled studies showed great heterogeneity (as demonstrated in Fig. 1). The studies differed substantially in terms of the sequenced genes, the structure of the study and the analytical strategies. Additionally, only a few studies had an appropriate, matched control group, typical for a case-control study, but usually, those studies were too small to estimate the BC and OC risk correctly. Differences in the selection of particular genes emphasize the lack of sufficient information about the association of many genes with a particular type of cancer. As MGP sequencing is becoming more popular, the results of few new studies [8–14], including one cumulative, large-scale analysis [5], were published during the preparation of the manuscript. The results of these studies are not included in the meta-analysis, but we refer to them in the discussion.

The aggregation of results from different studies in our meta-analysis lead to a higher statistical power and a more robust risk estimate (ORs) with substantially tighter confidence intervals than previously reported in single studies. In total, 13 and 11 of the analyzed genes were significantly associated with an increased BC and OC risk, respectively.

In the meta-analysis, 4.5% of BC patients and 13.1% of OC patients carried *BRCA1* or *BRCA2* mutations. The frequency of mutations and the risk attributed to the mutations in these genes were higher for OC than for BC. The prevalence is especially high for *BRCA1* (OR = 9.3 and OR = 35.3 for BC and OC, respectively), what indicates observed also before higher attribution of *BRCA1* than *BRCA2* mutations to OC. It has to be noted, however, that although risk (OR values) attributed to the *BRCA1/2* mutations is higher for OC than BC, due to much higher overall frequency of BC, asymptomatic *BRCA1/2* mutations carriers are more likely to develop BC than OC.

The analysis revealed that four genes, i.e., *CDKN2A*, *PTEN*, *PALB2*, and *TP53*, were associated with a BC risk (ORs > 4) similar to that of *BRCA2*. Except the very recent study of Couch et al., that also showed elevated BC risk attributed to *CDKN2A* mutations [1], *CDKN2A* was not considered before as conferring a strong BC risk. Additionally, several independent association studies revealed an association of single SNPs in the *CDKN2A* gene with a slightly increased risk of esophageal cancer, endometrial cancer, as well as BC (OR ~ 1.2–1.5) [15–18]. Lack of a previous association of the *CDKN2A* mutations with BC may result from a low frequency of mutations in that gene (0.1% of all BC), which results in a relatively high range of confidence intervals (still mostly overlapping with high-risk values).

Mutations in *BRCA1*, *BRCA2*, *RAD51D*, *TP53*, *BRIP1*, *RAD51C*, *MSH6*, *MSH2*, and *PTEN* were associated with a high OC risk (OR > 4). The elevated OC risk for *MSH6* (OR = 4.1) and *MSH2* (OR = 4.0) has also been observed in other studies [1,5,6]; however, in the context of our results, *MSH2* was attributed to a higher BC/OC risk than *MSH6*. *RAD51C* and *RAD51D* mutations were identified in families with BC and OC [19,20] and implicated in BC and OC risk. However, our results clearly showed an association of both *RAD51C* and *RAD51D* with a high risk of OC, but not BC. A higher OC vs. BC risk attributed to *RAD51C* and *RAD51D* has also been suggested in several smaller studies [1,2,5,7,21–23]. Mutations in *BRIP1* appear to also be important in the development of OC (OR = 4.9). However, it must be noted that in a series of previous

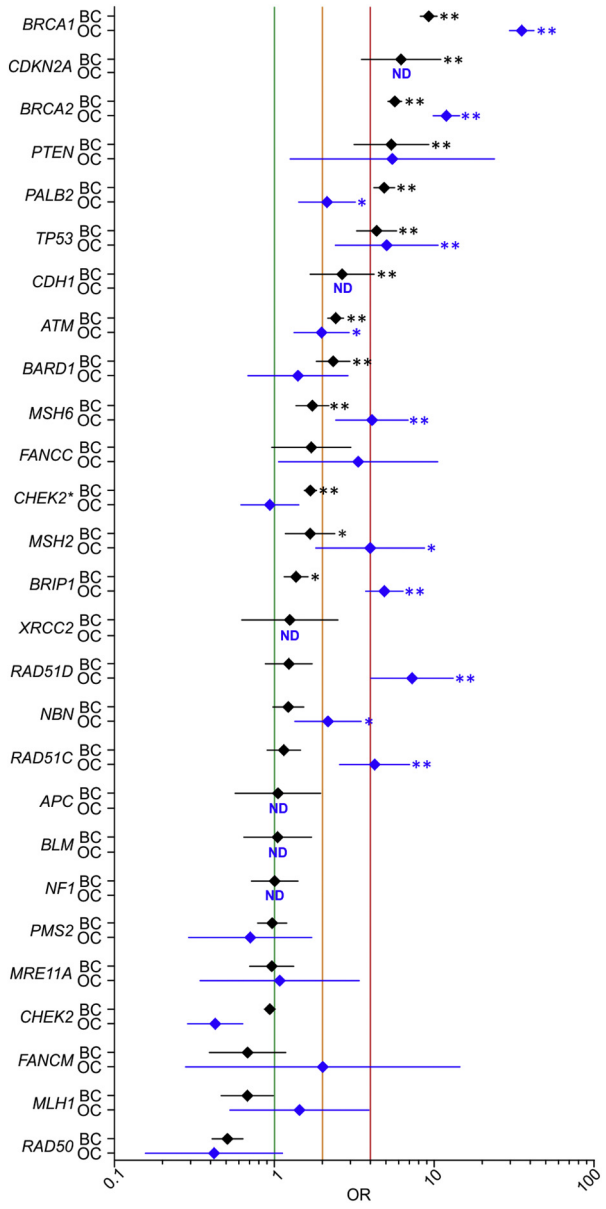


Fig. 6. Summary of the association analysis and risk (ORs) estimated for the particular analyzed genes. Diamonds and horizontal lines indicate the OR values and 95% CI, respectively. Black and blue symbols indicate BC and OC values, respectively. The genes are sorted according to ORs estimated for BC (from highest to lowest). Green, orange, and red vertical lines highlight 1 (no risk), 2 (the threshold for moderate risk), and 4 (the threshold for high risk) OR values, respectively. Low-risk genes are located under the orange line. ND indicates that no mutation was detected in the genes in a particular cancer. An asterisk next to the symbols indicates p-value associated with an increased risk (* and ** reflect $p < 0.05$ and $p < 0.0001$, respectively). Note, that not all genes with <5 mutations identified are shown. The detailed values of OR for particular genes are provided in Table S4.

association studies, the risk estimated was much higher (OR in a range between 5 and 29, in different populations) [5,7,24–26]. Although *BRIP1* mutations were not associated with BC risk in our study, confirming other results [6,26,27], some reports suggest that these mutations may be associated with a high BC risk in patients with an early-onset BC or a strong family history of BC [28,29].

Another striking conclusion of our study is the substantially different pattern of mutation frequency and the consequent different risk attributed to particular genes in BC and OC. Although, a higher risk of either BC or OC associated with individual genes has been often reported before; here, for the first time, we present a direct, comprehensive comparison of many of the most important BC and OC risk candidates, additionally supported by the strong statistical power of estimated risk. Significant differences were shown for the *BRCA1*, and *BRCA2* genes, mutations of which were attributed to a much higher OC than BC risk. Additionally, mutations in *BRIP1*, *MSH2*, and *MSH6* genes were associated with a low BC risk and a high OC risk. More importantly, mutations in *RAD51D*, *RAD51C*, and *NBN* were specific for OC (associated with the OC but not BC risk), indicating their selective role in the genetic predisposition of OC development. On the other hand, a higher risk of BC than OC was established for *PALB2* mutations, and it appears that mutations in *CHEK2** and *BARD1* are specific for BC (associated with a low BC risk but not with an OC risk). Nevertheless, it has to be noted, that increased BC risk estimates for *RAD51D* were obtained in the sensitivity analysis (in extracted European/White population, and after exclusion of the two largest studies), suggesting a potential role of *RAD51D* in BC susceptibility.

Of equal importance, however, is the finding that mutations in more than a dozen analyzed candidate genes, i.e., *XRCC2*, *RAD51D*, *NBN*, *RAD51C*, *APC*, *BLM*, *NF1*, *PMS2*, *CHEK2*, *MRE11A*, *FANCC*, *FANCM*, *MLH1*, and *RAD50*, are not associated with any BC risk. Most of these genes are also either not associated with OC or mutations in these genes are absent or too scarce in OC. These findings indicate that the inclusion of these genes in BC/OC MGPs must be reconsidered. However, the lack of association <lack of mutations> in genes such as *STK11*, *NF1*, or *VHL* associated with characteristic hereditary cancer syndromes (Peutz-Jeghers syndrome, neurofibromatosis type 1, and von Hippel–Lindau disease, respectively), may result from the fact that BC/OC cases with these syndromes are tested for genes specific for particular syndromes, which still raises questions regarding the need to include these genes in BC/OC MGP testing. Some opposition to the above statement is observed in our study on the association of *TP53* mutations with a high risk for both BC and OC. The question remains whether mutations identified in MGPs and associated with BC/OC that are mostly missense mutations located in specific TP53 regions (protein domains) are also associated with the typical Li-Fraumeni syndrome or are preferentially associated with particular cancers.

Although the accuracy of our results is enhanced by the size of the meta-analysis, they should be interpreted within the context of several limitations. First, despite our efforts to standardize case and control groups, possible bias could have been introduced, due to using control data from the public database that were not perfectly matched in terms of sex, age, ethnicity, geographical area or sequencing platforms to the case groups. Moreover, technical differences in the selected studies such as the (i) DNA sample preparation and quality; (ii) NGS library preparation and enrichment strategy; (iii) NGS sequencing strategy, including the depth of coverage of analyzed sequences; and (iv) algorithm used for mutation detection and quality (inclusion thresholds) may affect the mutation detection and therefore the bias comparison between groups. Second, as not all studies included a full list of identified mutations, we classified mutations as defined by authors, and possible differences in the meaning of the pathogenicity of mutation cannot be ruled out. Third, the European/White population group constituted the majority of studies fulfilling our criteria; therefore, general ORs can reflect the true trends of the European/White population. Forth, the population of mixed ethnicity used in our meta-analysis had little power to detect founder mutations that may be specific and account for a substantial fraction of mutations in particular gene in particular population. Additionally, our analysis was limited almost exclusively to mutations located in protein-coding sequences and did not include large mutations or mutations in noncoding sequences. However, we may expect from other intensively studied genes that large mutations may account for 5–10% of all mutations in genes included in BC/OC MGPs [30–32]. The mutations identified recently in the promoter of *TERT* [33,34] and examples of deep-intronic mutations associated with Mendelian disorders (e.g., tuberous sclerosis complex [35], cystic fibrosis [36], Duchenne muscular dystrophy [37], polycystic kidney and hepatic disease [38]

and hemophilia A [39]), reviewed recently in Human Genetics [40], indicate that deleterious DNA variants may also occur in noncoding regions. However, the identification of mutations in noncoding sequences is not an easy task, although currently, it is not limited anymore by sequencing capabilities but by problems with their interpretations as deleterious variants. Finally, our analysis, as any meta-analysis, may be affected by the heterogeneity of the included studies. The heterogeneity of our analysis may results from (i) different ethnic background of the studies, (ii) different contribution of founder mutations, and (iii) different criteria of samples selection.

These limitations create the need for further association analyses, with case-control groups matched by gender, age, ethnicity, and sequencing platform, to confirm the results obtained in our meta-analysis. As mutation frequencies may vary in different ethnicities, association analyses should be performed for separate geographical region/ethnicities, which would help define founder mutations characteristic for each population, which in turn could be used in screening tests. Of equal importance is the estimation of BC and OC risk for particular recurrent variants/mutations, especially for founder mutations, as the risk presented for a particular gene cannot be generalizable to all identified mutations. Finally, there is still a need to collect data derived from a larger group of patients in order to calculate the correct BC and OC risk for mutations in several genes.

Acknowledgments

This work was supported by grants from the Polish National Science Centre (NCN 2015/17/B/NZ2/01182 and 2016/22/A/NZ2/00184 awarded to Piotr Kozlowski, and 2015/19/N/NZ5/02247 awarded to Katarzyna Klonowska).

The authors would like to thank the Genome Aggregation Database (gnomAD) and the groups that provided exome and genome variant data to this resource. A full list of contributing groups can be found at http://gnomad.broadinstitute.org/about.

Conflict of interest statement

This work was supported by grants from the Polish National Science Centre (NCN 2015/17/B/NZ2/01182, 2016/22/A/NZ2/00184, and 2015/19/N/NZ5/02247). The funders reviewed the design of grants, but otherwise had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The authors have no other conflicts of interest to declare.

Author contributions

MS – collected articles, performed statistical analyses, and prepared the manuscript, including the figures, tables, and supporting materials; KK – collected articles, performed statistical analyses, and participated in manuscript preparation, including the figures, tables and supporting materials; AJ – discussed the concept of the manuscript; PK – conceived and coordinated the study and supervised the manuscript preparation.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ygyno.2019.01.027.

References

[1] F.J. Couch, H. Shimelis, C. Hu, S.N. Hart, E.C. Polley, J. Na, et al., Associations between cancer predisposition testing panel genes and breast cancer, JAMA Oncol. 3 (2017) 1190–1196.
[2] S.S. Buys, J.F. Sandbach, A. Gammon, G. Patel, J. Kidd, K.L. Brown, et al., A study of over 35,000 women with breast cancer tested with a 25-gene panel of hereditary cancer genes, Cancer 123 (2017) 1721–1730.
[3] M. Lek, K.J. Karczewski, E.V. Minikel, K.E. Samocha, E. Banks, T. Fennell, et al., Analysis of protein-coding genetic variation in 60,706 humans, Nature 536 (2016) 285–291.
[4] M.J. Landrum, J.M. Lee, M. Benson, G.R. Brown, C. Chao, S. Chitipiralla, et al., ClinVar: improving access to variant interpretations and supporting evidence, Nucleic Acids Res. 46 (2018) D1062–D1067.
[5] J. Lilyquist, H. LaDuca, E. Polley, B.T. Davis, H. Shimelis, C. Hu, et al., Frequency of mutations in a large series of clinically ascertained ovarian cancer cases tested on multigene panels compared to reference controls, Gynecol. Oncol. 147 (2017) 375–380.
[6] T.P. Slavin, K.N. Maxwell, J. Lilyquist, J. Vijai, S.L. Neuhausen, S.N. Hart, et al., The contribution of pathogenic variants in breast cancer susceptibility genes to familial breast cancer risk, NPJ Breast Cancer 3 (2017) 22.
[7] B.M. Norquist, M.I. Harrell, M.F. Brady, T. Walsh, M.K. Lee, S. Gulsuner, et al., Inherited mutations in women with ovarian carcinoma, JAMA Oncol. 2 (2016) 482–490.
[8] P.L. Sung, K.C. Wen, Y.J. Chen, T.C. Chao, Y.F. Tsai, L.M. Tseng, et al., The frequency of cancer predisposition gene mutations in hereditary breast and ovarian cancer patients in Taiwan: from BRCA1/2 to multi-gene panels, PLoS One 12 (2017), e0185615. .
[9] A.K. Siraj, T. Masoodi, R. Bu, S.K. Parvathareddy, I.A. Al-Badawi, N. Al-Sanea, et al., Expanding the spectrum of germline variants in cancer, Hum. Genet. 136 (2017) 1431–1444.
[10] J.S. Park, S.T. Lee, E.J. Nam, J.W. Han, J.Y. Lee, J. Kim, et al., Variants of cancer susceptibility genes in Korean BRCA1/2 mutation-negative patients with high risk for hereditary breast cancer, BMC Cancer 18 (2018) 83.
[11] J. Li, H. Li, I. Makunin, I. kConFab, B.A. Thompson, K. Tao, et al., Panel sequencing of 264 candidate susceptibility genes and segregation analysis in a cohort of non-BRCA1, non-BRCA2 breast cancer families, Breast Cancer Res. Treat. 166 (2017) 937–949.
[12] K.J. Eoh, J.E. Kim, H.S. Park, S.T. Lee, J.S. Park, J.W. Han, et al., Detection of germline mutations in patients with epithelial ovarian cancer using multi-gene panels: beyond BRCA1/2, Cancer Res. Treat. 50 (2018) 917–925.
[13] F. Fostira, E. Saloustros, P. Apostolou, A. Vagena, D. Kalfakakou, D. Mauri, et al., Germline deleterious mutations in genes other than BRCA2 are infrequent in male breast cancer, Breast Cancer Res. Treat. 169 (2018) 105–113.
[14] W. Jian, K. Shao, Q. Qin, X. Wang, S. Song, X. Wang, Clinical and genetic characterization of hereditary breast cancer in a Chinese population, Hereditary Cancer Clin. Pract. 15 (2017) 19.
[15] T. Debniak, B. Gorski, T. Huzarski, T. Byrski, C. Cybulski, A. Mackiewicz, et al., A common variant of CDKN2A (p16) predisposes to breast cancer, J. Med. Genet. 42 (2005) 763–765.
[16] T. Debniak, C. Cybulski, B. Gorski, T. Huzarski, T. Byrski, J. Gronwald, et al., CDKN2A-positive breast cancers in young women from Poland, Breast Cancer Res. Treat. 103 (2007) 355–359.
[17] K.E. Driver, H. Song, F. Lesueur, S. Ahmed, N.L. Barbosa-Morais, J.P. Tyrer, et al., Association of single-nucleotide polymorphisms in the cell cycle genes with breast cancer in the British population, Carcinogenesis 29 (2008) 333–341.
[18] F. Gu, R.M. Pfeiffer, S. Bhattacharjee, S.S. Han, P.R. Taylor, S. Berndt, et al., Common genetic variants in the 9p21 region and their associations with multiple tumours, Br. J. Cancer 108 (2013) 1378–1386.
[19] A. Meindl, H. Hellebrand, C. Wiek, V. Erven, B. Wappenschmidt, D. Niederacher, et al., Germline mutations in breast and ovarian cancer pedigrees establish RAD51C as a human cancer susceptibility gene, Nat. Genet. 42 (2010) 410–414.
[20] C. Loveday, C. Turnbull, E. Ramsay, D. Hughes, E. Ruark, J.R. Frankum, et al., Germline mutations in RAD51D confer susceptibility to ovarian cancer, Nat. Genet. 43 (2011) 879–882.
[21] L.M. Pelttari, T. Heikkinen, D. Thompson, A. Kallioniemi, J. Schleutker, K. Holli, et al., RAD51C is a susceptibility gene for ovarian cancer, Hum. Mol. Genet. 20 (2011) 3278–3288.
[22] L.M. Pelttari, J. Kiiski, R. Nurminen, A. Kallioniemi, J. Schleutker, A. Gylfe, et al., A Finnish founder mutation in RAD51D: analysis in breast, ovarian, prostate, and colorectal cancer, J. Med. Genet. 49 (2012) 429–432.
[23] H. Song, E. Dicks, S.J. Ramus, J.P. Tyrer, M.P. Intermaggio, J. Hayward, et al., Contribution of germline mutations in the RAD51B, RAD51C, and RAD51D genes to ovarian cancer in the population, J. Clin. Oncol. 33 (2015) 2901–2907.
[24] T. Rafnar, D.F. Gudbjartsson, P. Sulem, A. Jonasdottir, A. Sigurdsson, A. Jonasdottir, et al., Mutations in BRIP1 confer high risk of ovarian cancer, Nat. Genet. 43 (2011) 1104–1107.
[25] S.J. Ramus, H. Song, E. Dicks, J.P. Tyrer, A.N. Rosenthal, M.P. Intermaggio, et al., Germline mutations in the BRIP1, BARD1, PALB2, and NBN genes in women with ovarian cancer, J. Natl. Cancer Inst. 107 (2015).
[26] N. Weber-Lassalle, J. Hauke, J. Ramser, L. Richters, E. Gross, B. Blumcke, et al., BRIP1 loss-of-function mutations confer high risk for familial ovarian cancer, but not familial breast cancer, Breast Cancer Res. 20 (2018) 7.
[27] D.F. Easton, F. Lesueur, B. Decker, K. Michailidou, J. Li, J. Allen, et al., No evidence that protein truncating variants in BRIP1 are associated with breast cancer risk: implications for gene panel testing, J. Med. Genet. 53 (2016) 298–309.
[28] S. Seal, D. Thompson, A. Renwick, A. Elliott, P. Kelly, R. Barfoot, et al., Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles, Nat. Genet. 38 (2006) 1239–1241.
[29] G.B. Byrnes, M.C. Southey, J.L. Hopper, Are the so-called low penetrance breast cancer genes, ATM, BRIP1, PALB2 and CHEK2, high risk for women with strong family histories? Breast Cancer Res. 10 (2008) 208.
[30] P. Kozlowski, P. Roberts, S. Dabora, D. Franz, J. Bissler, H. Northrup, et al., Identification of 54 large deletions/duplications in TSC1 and TSC2 using MLPA, and genotype-phenotype correlations, Hum. Genet. 121 (2007) 389–400.
[31] M.J. Smith, J.E. Urquhart, E.F. Harkness, E.K. Miles, N.L. Bowers, H.J. Byers, et al., The contribution of whole gene deletions and large rearrangements to the mutation spectrum in inherited tumor predisposing syndromes, Hum. Mutat. 37 (2016) 250–256.

[32] L.M. Kayes, W. Burke, V.M. Riccardi, R. Bennett, P. Ehrlich, A. Rubenstein, et al., Deletions spanning the neurofibromatosis 1 gene: identification and phenotype of five patients, Am. J. Hum. Genet. 54 (1994) 424–436.
[33] F.W. Huang, E. Hodis, M.J. Xu, G.V. Kryukov, L. Chin, L.A. Garraway, Highly recurrent TERT promoter mutations in human melanoma, Science 339 (2013) 957–959.
[34] S. Horn, A. Figl, P.S. Rachakonda, C. Fischer, A. Sucker, A. Gast, et al., TERT promoter mutations in familial and sporadic melanoma, Science 339 (2013) 959–961.
[35] M.E. Tyburczy, K.A. Dies, J. Glass, S. Camposano, Y. Chekaluk, A.R. Thorner, et al., Mosaic and intronic mutations in TSC1/TSC2 explain the majority of TSC patients with no mutation identified by conventional testing, PLoS Genet. 11 (2015), e1005637. .
[36] W.E. Highsmith, L.H. Burch, Z. Zhou, J.C. Olsen, T.E. Boat, A. Spock, et al., A novel mutation in the cystic fibrosis gene in patients with pulmonary disease but normal sweat chloride concentrations, N. Engl. J. Med. 331 (1994) 974–980.
[37] Y. Takeshima, M. Yagi, Y. Okizuka, H. Awano, Z. Zhang, Y. Yamauchi, et al., Mutation spectrum of the dystrophin gene in 442 Duchenne/Becker muscular dystrophy cases from one Japanese referral center, J. Hum. Genet. 55 (2010) 379–388.
[38] L. Michel-Calemard, F. Dijoud, M. Till, J.C. Lambert, M. Vercherat, V. Tardy, et al., Pseudoexon activation in the PKHD1 gene: a French founder intronic mutation IVS46+653A>G causing severe autosomal recessive polycystic kidney disease, Clin. Genet. 75 (2009) 203–206.
[39] B. Pezeshkpoor, N. Zimmer, N. Marquardt, I. Nanda, T. Haaf, U. Budde, et al., Deep intronic 'mutations' cause hemophilia A: application of next generation sequencing in patients without detectable mutation in F8 cDNA, J. Thromb. Haemost. 11 (2013) 1679–1687.
[40] R. Vaz-Drago, N. Custodio, M. Carmo-Fonseca, Deep intronic mutations and human disease, Hum. Genet. 136 (2017) 1093–1111.