UNIVERSITÀ
DI PAVIA

Dipartimento di Biologia e Biotecnologie

Identification and validation of DNA sequence variants in cancer predisposition genes by Next Generation Sequencing approaches

Matteo Dal Molin

Dottorato di Ricerca in
Genetica, Biologia Molecolare e Cellulare
Ciclo XXXV - A.A. 2019-2022

To my parents Bruno and Ivonne

Table of contents

Table of contents

1. Abstract

In our lab (Clinical Genomics Unit at the European Institute of Oncology, Milan) we perform germline genetic tests in patients with suspected predisposition to cancer, specifically focusing on the two most common cancer syndromes, *i.e.* Hereditary Breast and Ovarian Cancer (HBOC) and Lynch Syndrome (LS) which is characterized by the predisposition to colorectal, endometrial, gastric and ovarian cancers. In addition, rare conditions are investigated, including Familial Melanoma (FM), Hereditary Diffuse Gastric Cancer (HDGC), Familial Adenomatous Polyposis (both FAP and MAP), and Li Fraumeni (LF) syndrome. Following genetic counselling, we search for germline mutations in cancer predisposition genes by means of DNA next generation sequencing (NGS)-based approaches. These approaches rely on cancer gene panels, allowing us not only to investigate the specific genes already known to be associated with each syndrome (*e.g. BRCA1/2* genes for HBOC; *TP53* gene for LF syndrome), but also to extend the analysis to other genes, simultaneously.

We retrospectively analyzed 2077 DNA samples sequenced at the Clinical Genomics Unit following the request for *BRCA1* and *BRCA2* genetic test. One-hundred-nineteen individuals (5.7%) proved to carry a pathogenetic variant (*i.e.* a disease causing mutation) in *BRCA1* and 120 (5.8%) had a pathogenic variant in *BRCA2*. Fifty-three (2.6%) patients carried a VUS (*i.e.* a Variant of Unknown Significance) in *BRCA1* and 97 (4.7%) had a VUS in *BRCA2.* Ninety-six percent of the individuals screened for *BRCA1* gene and carrying a pathogenic variant were affected with breast/ovarian cancer, according to the expected genotype-phenotype correlations. Two percent of the individuals showed a tumor type "non corresponding" to the mutated gene: one subject was affected with melanoma and the other with squamous cell carcinoma. The remaining two percent of the sequenced individuals were carrying a *BRCA1* pathogenic mutation but were healthy at the time of the genetic test. Among individuals with *BRCA2* variants, one pathogenic mutation carrier was affected with melanoma, while one individual with a VUS was affected with lymphoma. All the remaining patients were affected with breast/ovarian cancer, as expected according to genotype-phenotype correlations. In our cohort, 81.2% of the individuals did not show any mutations in *BRCA1*/*BRCA2* genes. Of relevance, these individuals had been selected for strong family history of cancer and 96% of them had already developed breast and/or ovarian cancers. This finding could be partially explained by the presence of deep-intronic pathogenic variants (lying in DNA regions not analyzed with the gene panel), or by epigenetic mechanisms inactivating the investigated genes. Alternatively, some cases could be explained by pathogenic variants in cancer predisposition genes not included in the panel.

We evaluated the status of *MMR* genes in 282 cases with suspected Lynch Syndrome, following the request for *MLH1, MSH2, EPCAM, MSH6, PMS2* gene

testing. All the individuals screened for *MMR* genes and carrying a pathogenic variant (10.6% - 30/282) proved to be affected with a tumor of the LS tumor spectrum (colorectal, endometrial, and gastric cancers). Among individuals carrying VUS in *MMR* genes (9.5% 27/282) 6 were affected with breast cancer and one with a CNS tumor; the remaining patients were affected with a tumor of the *MMR* tumor spectrum (colorectal, endometrial, and intestinal cancer).

We sequenced 779 patient samples following the request for multi gene panel testing. Sixty out of 779 individuals (7.7%) showed a pathogenic variant in a cancer related gene, while 138 (17.7%) showed a VUS. The use of multi gene panels can uncover "incidental findings", showing the presence of pathogenic variants or VUS not consistent with the patient's family history or the presence of variants in cancer predisposition genes linked to the patient's disease but not requested by the clinician. This testing approach allowed us to identify new associations between the disease affecting the patient and the mutated gene, including: *MUTYH* (monoallelic pathogenetic variants) and breast, prostate, and thyroid cancers; *PALB2* and ovarian cancer; *CDKN2A* and breast cancer. The use of multi gene panels typically produces a great number of VUS. VUS are a challenging aspect in molecular diagnostics of hereditary cancers: they are variants for which evidence of pathogenicity is limited or contradictory, thus precluding a benign/likely-benign or pathogenic/likely-pathogenic classification. Of relevance, classifications may change over time, with molecular, genetic, and epidemiological evidence accumulating. Accordingly, it is mandatory that laboratories periodically reassess variants to ensure patients receive up-to-date genetic results, thus guaranteeing the most appropriate clinical management. In our lab, 13.4% of *BRCA1* and *BRCA2* variants initially classified as VUS were reclassified over a 4-year period. Among these, 75% were reclassified as benign/likely-benign and 25% as pathogenic/likely-pathogenic.

In parallel with DNA sequencing, we applied the MLPA method (Multiplex ligation-dependent probe amplification) to search in cancer genes for CNVs (Copy Number Variants), *i.e.* for gross rearrangements, including deletions and/or duplications. Among individuals with genetic alterations, we identified 47 subjects carrying a CNV. These individuals represented the 7% of all sequenced cases. *BRCA1* was the gene showing the highest number of CNVs (26/47 individuals). All deletions we identified were classified as "pathogenic" since leading to a nonfunctional protein product. On the other side, most duplications we identified were classified as VUS. We deeply analyzed a duplication of *CHEK2* exon 2, not reported in literature and initially classified as VUS. We were able to demonstrate that the rearrangement consisted in a *tandem* duplication creating a premature stop codon in the DNA sequence. On this basis we could classify the duplication as "pathogenic".

NGS is highly effective for detecting single nucleotide variations but can fail in discriminating genes from their pseudogenes (as in the case of *PMS2* or *CHEK2* genes). Misalignment of sequencing reads to a pseudogene can lead to false positive variant calls if not confirmed with an orthologous methodology. By applying alternative molecular strategies (*e.g.* long-range PCR and Sanger sequencing), we

were able to validate pathogenic variants in *PMS2* and *CHEK2*, discriminating both genes from their pseudogene counterparts.

The selection of patients for *BRCA1/BRCA2* testing has long relied on the presence of a strong family history of breast and ovarian cancer. We developed a new approach ("*Persona project"*) to evaluate the presence of variants in "Triple Negative Breast Cancer" (estrogen and progesterone receptors, and HER2 protein are absent) patients, regardless of family history, to assess the prognostic and therapeutic predictive role of Homologous Recombination gene defects *vs.* other genetic alterations. By sequencing 318 TNBC patients with a multi gene panel, we found 90 patients (28.3%) with pathogenic variants in one or more "actionable" genes. The genetic information we got through the NGS analysis will now be evaluated and discussed in a multidisciplinary board to address individuals to a personalized cure strategy.

2. Abbreviations

BC	breast cancer
CGS	clinical genomics unit
CNV	copy number variant
CRC	colorectal cancer
DGC	diffuse-type gastric cancer
ER	estrogen receptor
FAP	familial adenomatous polyposis
FM	familial melanoma
GWAS	genome-wide association study
HBOC	hereditary breast and ovarian cancer
HDGC	hereditary diffuse gastric cancer
HER2	human epidermal growth factor receptor 2
HNPCC	hereditary nonpolyposis colorectal cancer
HSCT	hematopoietic stem cell transplantation
IF	incidental finding
IGCLC	International Gastric Cancer Linkage Consortium
Indel	short insertion or deletion
LFS	Li-Fraumeni syndrome
MAP	*MUTYH*-associated polyposis
MLPA	multiplex ligation-dependent probe amplification
MMR	mismatch repair
NCCN	clinical practice guidelines in oncology

Abbreviations

NGS	next generation sequencing
PR	progesterone receptor
SEER	surveillance, epidemiology, and end results
SNP	single nucleotide polymorphism
SNV	single nucleotide variant
WES	whole exome sequencing
WGS	whole genome sequencing
WT	wild type

3. Introduction

3.1 Genetic predisposition to cancer

Cancer is a common disease in the general population and, occasionally, it may even seem to run in families. This might be because family members share environmental exposures, behaviour patterns or conditions, such as smoking habit or obesity, that increase cancer risk. Alternatively, this might be because family members share, at the germline level, low/moderate-penetrance genetic variants that slightly increase cancer predisposition. However, in most familial cancer cases the disease is caused by a combination of both environmental and genetic factors. Less frequently, cancer is associated with the presence, at the germline level, of a high-penetrance defective gene that is being passed from generation to generation. Although this genetic condition is often referred to as "inherited cancer", what is inherited is the defective gene that greatly increases the risk of developing a particular type or different types of tumors.

Cancer remains a leading cause of morbidity and mortality across the globe, with increases in mortality of approximately 25.0% since the 1990s and projections of $\geq$ 23 million cases annually by 2030. An estimated 20% of cancer patients has a positive family history of cancer and about 10% of cases is ascribable to specific inherited conditions, *i.e.* to specific high-penetrance genes. Indeed, the hereditary cancer syndromes are defined as an elevated risk of cancer that runs in the family, with the risk being associated with germline lesions in specific genes. The type/types of tumors are dependent on the mutated gene.

For example, the Hereditary Breast and Ovarian Cancer syndrome (HBOC) originates from *BRCA1* or *BRCA2* gene mutations that significantly increase the likelihood of developing breast, ovarian, prostate, and other types of cancer. The most common inherited condition that increases the risk for colon cancer is the Lynch syndrome, associated with germline mutations in DNA MisMatch Repair (*MMR*) genes. Patients with the Li-Fraumeni syndrome, associated with *TP53* mutations, have a heightened risk of cancer before age 30, and are almost guaranteed to suffer from cancer by the age of 60. Of relevance, carriers of cancer syndrome related genes also have a higher risk of multiple malignancies and rare tumors and are more likely to develop cancer at a younger age.

Our understanding of cancer syndromes has greatly advanced with recent progresses in DNA sequencing technology. Indeed, the high throughput sequencing is allowing geneticists to find new cancer predisposition genes and to improve the efficiency of genetic screening. The most common germline alterations associated with cancer syndromes include: single nucleotide variants (SNVs); duplications, insertions, or deletions (Indels); exon and gene copy number variants (CNVs); structural variants (SVs). The molecular analysis of inherited cancer genes ranges from simple assessments of known hotspot mutations in single genes, to more complex tests

that simultaneously detect alterations in different gene by Next Generation Sequencing (NGS). Moreover, allele-specific PCR can be applied to test the expression of each allele, Sanger sequencing to confirm mutations emerged from NGS, and multiplex ligation-dependent probe amplification (MLPA) to assess the presence of gross gene rearrangements. More specifically, NGS technologies have revolutionized the molecular profiling allowing sequencing of multiple target genes (multi-gene panel), of whole exome (WES), thus profiling all protein-coding regions, and of whole genome (WGS), thus profiling both protein-coding and non-coding regions. These technologies can permit the identification and classification of families with cancer predisposition and help to detect individuals at-risk, likely enhancing surveillance and early disease detection to prevent mortality. Accordingly, increased public awareness that cancer can be heritable, and that the heritable risk can be evaluated, has increased the demand for genetic counselling and genetic testing (Alharthi *et al*., 2020).

Basically, in laboratory settings, genetic testing consists of DNA analysis by appropriate methods to search for the presence of germline alterations in cancer-related genes, including SNVs, Indels; CNVs and SVs.

3.2 Single Nucleotide Variants

A Single Nucleotide Variant (SNV) is the substitution of a single nucleotide with another along the DNA. A given SNV can be a rare mutation or a common variant, *i.e.* a Single Nucleotide Polymorphism (SNP). The precise definition of SNP requires the substitution to be present in at least 1% of the population (Fig. 1).

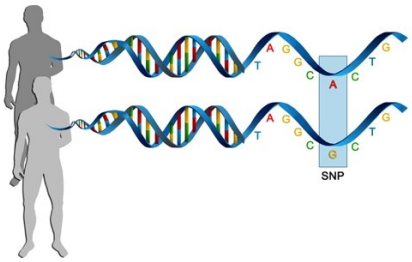


Fig. 1. Example of SNP: A has been replaced by G along the DNA sequence. In the human genome, the G nucleotide may appear in most individuals, but in a minority of individuals, the position is occupied by an A. This condition generates in the population 3 different genotypes, GG, GA, and AA.

SNVs occur naturally, are responsible for much of the genetic variability found between humans (SNPs), and most have no discernible affect on health. They can occur in coding and noncoding regions of DNA.

If a SNV occurs in a protein coding region, this can result in either:

- a nucleotide substitution that does not result in an amino acid change;

- a nucleotide substitution leading to an amino acid substitution. This may or may not result in a phenotypic effect (including pathogenicity) depending on the impact of the amino acid substitution on protein function and structure;

- a nucleotide substitution resulting in a stop codon and, consequently, in a truncated protein product. A truncated protein is generally non-functional.

3.3 Short insertions or deletions

Indel stands for short insertion or deletion and refers to a small sequence of DNA (usually less than 50 base pairs) that has either been inserted into or deleted from the genome. During translation, the mRNA sequence is read in groups of three bases (codons), with each codon corresponding to a given amino acid. If an Indel is not a multiple of three bases long, it will result in a frameshift that changes the reading of subsequent codons and, therefore, alters the entire amino acid sequence that follows the mutation. This has usually a strong impact on the protein product, often also leading to targeted decay of the mRNA. Indels are usually more harmful than a substitution in which only a single amino acid is changed.

3.4 Structural Variants

A Structural Variant is a large (more than 50 base pairs) rearrangement of part of the genome. It can be a deletion, a duplication, an insertion, an inversion, a translocation, or a combination of these changes (Fig. 2). Specifically, a Copy Number Variant (CNV) is a duplication or a deletion that changes the number of copies of a particular DNA sequence within the genome. As for SNVs and Indels, SVs have been implicated in several conditions, including cancer, cardiomyopathies, and intellectual disabilities.

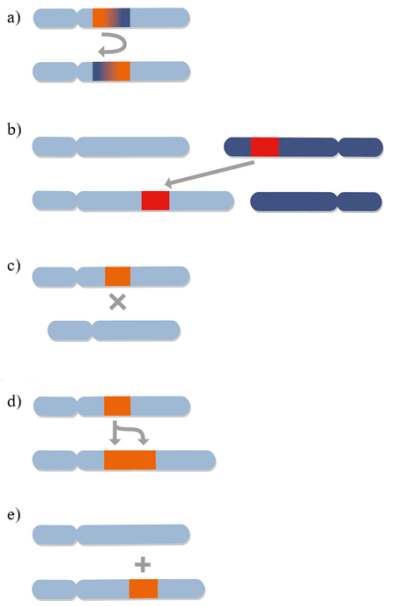


Fig. 2. Examples of tructural variants: a) inversion, b) translocation, c) deletion, d) duplication, e) insertion

3.5 Exome and gene copy number variation

Deletions and duplications of chromosomal segments (CNVs) are a major source of genetic variability among individual and are an underlying factor in human evolution and diseases, including mental illness, developmental disorders, and cancer. CNVs seem to be as important as SNPs in determining differences among individuals and represent a major driving force in evolution, especially in the rapid evolution that has occurred, and continues to occur, within the human and great apes' lineage. CNVs form at a faster rate than other types of mutations and are known to be generated by non-homologous end-joining recombination mechanisms. Recent models also focus on perturbation of DNA replication and replication of non-contiguous DNA segments. For example, cellular stress can induce repair of broken replication forks to switch from high-fidelity homologous recombination to non-homologous repair, thus promoting copy number change. An important fraction of the genome, currently estimated at up to 12%, is subject to CNVs. CNVs can arise both meiotically and somatically, as evident by the finding that identical twins can have different CNVs and that repeated sequences in different organs and tissues from the same individual can vary in copy number.

CNVs harbor coding regions and non-coding regulatory regions and may confer profound phenotypic effects relative to effects caused by SNPs. They have a multitude of effects based on their genomic location including gene dosage effects and cis-regulatory functions. CNVs that overlap protein coding genes offer insights into disease phenotypes and associated biology. Nearly 80% of cancer loci harbor CNVs and support the above premise (Kumaran *et al*., 2017). Changes in copy number

might change the expression levels of genes lying within the variable regions, allowing transcription levels to be higher or lower than those that can be achieved by control of transcription of a single copy gene per haploid genome. Additional copies of genes also provide redundancy that allows some copies to evolve new/modified functions or expression patterns while other copies maintain the original function. The non-homologous recombination events that underlie copy number changes also allow new combinations of exons between different genes by translocation, insertion, or deletion, so that proteins might acquire new domains, and hence new or modified activities.

However, much of the variation in copy number is disadvantageous. Copy number changes are involved in cancer genesis and progression and contribute to cancer proneness. In many cases, a change in copy number of any one of many genes is not well tolerated and leads to pathological conditions and genomic disorders. Most of our knowledge comes from studies of complete genomes (*e.g.* by array CGH-array comparative genomic hybridization), and from genome-wide surveys of CNVs by using techniques such as Multiplex Ligation-dependent Probe Amplification (MLPA) and quantitative real-time PCR. At present, thanks to advances made in NGS sequencing, which generate millions of sequences of the same target genomic region, it is possible to detect CNVs from NGS data through the appropriate bioinformatics tools. These latter usually apply a read depth approach based on counting the number of reads aligned to a particular region of the human genome.

Germline CNVs, *i.e.* constitutive loss or gain of genomic DNA ranging in size from 50 bp to 1 Mb, can represent genetic determinants of cancer susceptibility. This is the case of familial breast cancer and cancers of prostate, ovary, pancreas, colon, rectum, endometrium, as well as of familial melanoma. Breast cancer is the most common malignancy in women worldwide with about 2.09 million new cases diagnosed per year. It has been estimated that 5-10% of all breast cancers are hereditary cases. Family-based linkage analysis, gene re-sequencing and genome wide association studies allowed the identification of high, moderate, and low penetrant variants. However, these variants collectively explained only half of the breast cancer genetic predisposition. Thus, the genetic component of a substantial part of hereditary cases is yet to be discovered. CNVs can contribute to this remaining component of breast cancer risk, leading to disease development through their impact on gene expression and protein structure. As far as *BRCA1/2* predisposition genes is concerned, pathogenic CNVs are more frequent in *BRCA1* than in *BRCA2* and represent about 27% and 8% of *BRCA* genetic variants, respectively (Boujemaa *et al*., 2021). Such a high frequency can be explained by the higher number of Alu repeated sequences along *BRCA1* gene, and by the homologous recombination events occurring between *BRCA1* and its pseudogene (Boujemaa *et al*., 2021). Apart from *BRCA1* and *BRCA2*, deleterious CNVs have also been found in a series of highly penetrant cancer-predisposing genes, including mismatch repair genes (*MLH1, MSH2, MSH6, PMS2*) and *TP53* (Krepischi *et al*., 2012). Given that CNVs make a major contribution to the burden of all genetic diseases, we can expect that many new CNVs related to cancer susceptibility will be disclosed in the near future and that genome sequencing will become a tool of medical practice and differential diagnosis, leading to new approaches in management of cancer families (Krepischi *et al*., 2012).

3.6 Genetic risk assessment and counseling

Genetic testing is part of a more complex process, *i.e.* cancer risk assessment and genetic counseling. This is a multi-step process involving clinical assessment, genetic susceptibility testing (if appropriate) and risk management recommendations delivered in counseling sessions to individuals at risk for familial or hereditary cancer. The purpose of cancer genetic counseling is: to inform individuals about genetic, biological, and environmental factors related to a cancer diagnosis and/or to a risk for the disease; to provide meaning from genetic testing and genetic results; to identify appropriate strategies for cancer prevention and management.

Genetic testing should be considered in individuals for whom there is a personal/family history suggesting genetic susceptibility to cancer and for whom results will aid in risk management and treatment. In other words, the selection of appropriate candidates for genetic testing is based on the personal and familial characteristics that determine the individual's prior probability of being a carrier of a pathogenic or likely pathogenic variant, and on the psychosocial degree of readiness of the individual to receive the results of the genetic test.

The genetic risk assessment is a dynamic process: its accuracy is enhanced by relevant information retrieved from population and pedigrees studies and can change depending on how many and which relatives of the individual are diagnosed with cancer.

The genetic testing strategy is greatly facilitated when a pathogenic or likely pathogenic variant has already been identified in another family member of the individual under analysis. In this case, the testing laboratory can limit the search for the pathogenic or likely pathogenic variant to the same location in the same gene. However, if there is reason to suspect more than one disease-associated variant in the family, then broader testing may be considered. For most of the families in whom the presence of a pathogenic or likely pathogenic variant is unknown, it is best to start testing an affected family member with early-onset disease or with multiple primary tumors (with bilateral disease if testing for BC predisposition); indeed, this individual has the highest likelihood for a positive test result. The testing of unaffected individuals should only be considered when no affected family member is available for testing. Obviously, in such cases, a negative test result is considered indeterminate and does not provide the same level of information as when there is a known pathogenic or likely pathogenic variant in the family. Thus, one should be mindful that, when testing unaffected individuals (in the absence of having tested affected family members), significant limitations may exist; testing multiple family members may be indicated since the absence of a mutation in one unaffected relative does not rule out a mutation in others. The maternal and paternal sides of the family should be considered independently for familial patterns of cancer (Weitzel *et al*., 2017). Individuals who have received allogeneic hematopoietic stem cell transplantation (HSCT) should not have molecular genetic testing performed on blood samples, as blood cells would represent donor-derived DNA. In these cases, the DNA of the individual being tested should be extracted from a buccal swab (Tran *et al*., 2003;

Hong *et al*., 2007).

A counseling issue is posed by the finding of a so-called VUS (Variant of Unknown Significance). This genetic alteration may represent a benign variant unrelated to an increased cancer risk or may represent a variant associated with an increased cancer risk. The patients carrying a VUS should be considered for referral to research studies that aim to define the functional impact of genetic variants (variant re-classification studies performed through clinical labs and registries). It is important to note that there may be inconsistencies among research programs and registries about the significance and clinical actionability of some VUS and this, in turn, may lead to confusion regarding clinical management (Balmana *et al*., 2016; Vail *et al*., 2015; Lincoln *et al*., 2017). In the case where there are discrepancies in VUS classification, careful consideration must be taken to analyze family history and to test family members. Moreover, functional studies must be reviewed that possibly aid to interpret VUS pathogenic relevance (Eccles *et al*., 2015).

Carriers of a VUS should be managed based on family history of cancer. Carriers of a pathogenic or likely pathogenic variant should be encouraged to participate in clinical trials or genetic registries. They should be encouraged to recontact their medical genetics providers every few years for updates: laboratories may issue amended reports as the knowledge on hereditary cancer risk expands.

Tumor profiling, *i.e.* the pattern of genetic lesions present in tumor tissue, can be considered complementary to germline testing. In this context, it is mandatory to evaluate variants "pathogenicity" at somatic and germline levels. Indeed, tumor testing is designed to address treatment actionability and prognosis (Robson *et al*., 2015). Accordingly, a given variant considered as pathogenic or likely pathogenic at the germline level, *i.e.* associated with cancer risk/predisposition, may be considered as non-pathogenic or as a VUS in the tumor if the variant has no obvious clinical implications for therapeutic treatment or prognosis.

3.7 Multi-gene testing

Next-generation sequencing allows for the analysis of multiple genes simultaneously. When performed to search for germline variants associated with genetic diseases this is referred to as multi-gene testing. Multi-gene testing can detect pathogenic or likely pathogenic variants that are not found in single-gene testing, for example by simply analyzing the gene expected to be mutated and to be associated with a given clinical phenotype (Walsh *et al*., 2006; Kurian *et al*., 2014; Kurian *et al*., 2018).

Multi-gene testing may be most useful when more than one gene can explain an inherited cancer syndrome. In these cases, phenotype-directed testing based on personal and family history through amulti-gene panel test may be more efficient and/or cost-effective compared to the sequential analysis of each sigle gene (Hall *et al*., 2014; Hall *et al*., 2016; Manchanda *et al*., 2018). Multi-gene testing may also be considered for patients with personal and family history suggestive of an inherited

syndrome but tested negative for the gene/genes known to be associated with the syndrome (Walsh *et al*., 2010).

It is becoming more and more routine in many institutions to perform phenotypically directed multi-gene panel tests to assess for pathogenic changes in multiple relevant genes simultaneously (LaDuca *et al*., 2020). A problem regarding multi-gene testing is that there are limited dataand a lack of guidelines regarding the degree of cancer risk associated with some of the genes included in the panels. Therefore, the big challenge is how to communicate and manage the risk for carriers of these genes (Bombard *et al*., 2016; Rainville *et al*., 2014; Blazer *et al*., 2016; Tung *et al*., 2016; van Marcke *et al*., 2016). This issue is compounded by the low incidence of hereditary syndromes, leading to a difficulty in performing adequately powered studies (Bombard *et al*., 2013). Moreover, multi-gene tests, in addition to high-penetrance genes, often include low to moderate-penetrance genes; for these last, there are even less data regarding degree of cancer risk and guidelines for risk management (Pal *et al*., 2020; Hall *et al*., 2014; LaDuca *et al*., 2014). Also, certain variants in a gene may be associated with different degrees of risk compared to other variants within the same gene. For example, certain *ATM* genetic variants are associated with an increased risk for early-onset breast cancer and for bilateral cancer occurrence, while the association between other *ATM* variants and breast cancer susceptibility is unclear (Brunet *et al*., 2008; Heikkinen *et al*., 2005; Thompson *et al*., 2005; Tommiska *et al*., 2006).

3.8 Hereditary Breast Ovarian Cancer (HBOC) syndrome; high-penetrance breast/ovarian cancer susceptibility genes

Specific patterns of hereditary breast and ovarian cancers have been found to be linked to pathogenic/likely-pathogenic variants in *BRCA1/2* genes that give rise to the Hereditary Breast Ovarian Cancer (HBOC) syndrome (Blackwood *et* al., 1998; Venkitaraman *et* al., 2002). In addition to HBOC, there are two rare syndromes characterized by an increased risk for breast cancer, the Li-Fraumeni syndrome (LFS) and the Cowden syndrome, which are associated to germline mutations in the *TP*53 and *PTEN* genes, respectively (Pilarski *et* al., 2009; Schneider *et al*., 2013).

These hereditary syndromes share several features beyond an increased risk of breast cancer. They arise from germline pathogenic/likely-pathogenic variants that are not within sex-linked genes; hence, the variants can be inherited from either parent. The disease-related variants are highly penetrant. The syndromes are characterized by breast cancer onset at an early age (often bilateral) and development of other types of cancer and exhibit an autosomal dominant inheritance pattern. Offspring of an individual with one of these hereditary diseases have a 50% chance of inheriting the disease-related variant. The manifestations of the syndromes are often variable in individuals within a single family (e.g., age of onset, tumor site, number of primary tumors). The risk of developing cancer in individuals with one of the syndromes depends on numerous variables including the gender and age of the individual. Prior to 2020, the NCCN Guidelines for “Genetic/Familial High-Risk Assessment: Breast and Ovarian” focused largely on testing criteria for *BRCA1/2* and appropriate risk

management for carriers of a *BRCA1* or *BRCA2* pathogenic or likely pathogenic variants.

There is now strong evidence that genes beyond *BRCA1/2*, *TP*53 and *PTEN* confer markedly increased risk of breast and/or ovarian cancers.

In a recent paper (Peleg Hasson *et al*., 2020) Peleg Hasson and collaborators reviewed the relevant data on genetic testing for breast cancer. The results are summarized in Fig. 3. Most breast cancer cases are sporadic rather than inherited. Approximately 10-15% of patients are associated with hereditary syndromes (identified germline mutations), and most of them are carriers of a disease-causative variant in *BRCA1/2*.

When analyzing genetic tests of more than 60,000 patients with breast cancer and after excluding *BRCA*-positive patients and patients with syndromic *PTEN* and *TP53* gene mutations, more than 6% of individuals were found to have a pathogenic variant in other genes, including *CHEK2*, *PALB2* and *ATM* (Couch *et al*., 2017). Pathogenic variants in *PALB2*, *RAD51D*, *ATM*, *BARD1*, and *CHEK2* (Couch *et al*., 2017) were associated with high or moderately increased risk.

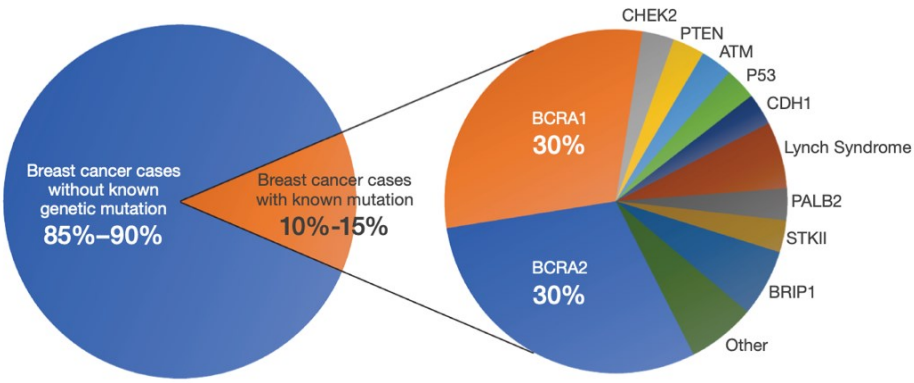


Fig. 3. Breast cancer patients with germline pathogenic variants in cancer-predisposition genes (from Peleg Hasson et al., 2020).

3.9 BRCA-Related Breast/Ovarian Cancer Syndrome

Both *BRCA1* and *BRCA2* genes encode for proteins involved in tumor suppression, with a specific role in the control of homologous recombination (HR) and double-strand break repair in response to DNA damage (Piri *et al*., 2001). *BRCA1/2* variants that are classified as pathogenic or likely pathogenic can be highly penetrant; however, the probability of cancer development for the carriers of these variants is quite variable, even within families with the same genetic lesion (Petrucelli *et al*., 2011). At present, it is unclear whether penetrance is only related to the specific variant identified in a family or to additional factors, either genetic or environmental (likely both), affecting disease expression. At any rate, it is generally accepted that carriers of

BRCA1/2 pathogenic or likely pathogenic variants are at high risk of both breast and ovarian cancer and warrant consideration of preventive strategies.

3.10 Breast Cancer Risk in BRCA1/2 mutation carriers

Estimates of penetrance range from 41% to 90% lifetime risk for breast cancer, with an increased risk for contralateral breast cancer (Mavaddat *et al*., 2013; van den Broek *et al*., 2016; Hu *et al*., 2020). A prospective cohort study including 9856 unaffected *BRCA1/2* mutation carriers showed that the cumulative risk of breast cancer by 80 years of age was 72% for *BRCA1* variant carriers and 69% for *BRCA2* variant carriers (Kuchenbaecker *et al*., 2017). Other estimates of cumulative risk for contralateral breast cancer 20 years after breast cancer diagnosis are 40% and 26% for *BRCA1* and *BRCA2* mutation carriers, respectively (Kuchenbaecker *et al*., 2017).

A systematic review including 68 papers reported that *BRCA1* and *BRCA2* germline mutations are associated with the development of contralateral breast cancer (Akdeniz *et al*., 2017).

Some histopathologic features have been reported to occur more frequently in breast cancers associated with a *BRCA1/2* pathogenic or likely pathogenic germline variant. For example, several studies have shown that a fraction of *BRCA1/2* associated breast cancers are “triple negative”. This term refers to cancers that don't have estrogen or progesterone receptors (ER-/PR-) and don't express the HER2 protein (Hu *et al*., 2020; Young *et al*., 2009). Specifically, from 7% to 28% of patients with triple-negative breast cancer have been reported to be *BRCA1* mutation carriers (Buys *et al*., 2017; Shimelis *et al*., 2018), while the incidence of *BRCA2* germline mutations ranges from 1% to 17% in triple-negative breast cancer cases unselected for age or family history (Meyer *et al*., 2012; Buys *et al*., 2017; Shimelis *et al*., 2018).

3.11 Ovarian Cancer Risk in BRCA1/2 mutation carriers

In carriers of *BRCA1/2* mutations an increased risk for cancers of the ovary, fallopian tube, and peritoneum has been observed (Arts-de Jong *et* al., 2016). Specifically, germline mutations in *BRCA1/2* genes are responsible for at least 10% of epithelial ovarian cancers (Norquist *et al*., 2016). Depending on the population studied, *BRCA1/2* carriers have an estimated 8% to 62% lifetime risk for ovarian cancer (Finch *et* al., 2006). An analysis of 2222 patients with ovarian cancer showed that among subjects with high-grade serous carcinoma 11% were carriers of a *BRCA1/2* mutation (Song *et al*., 2014). Among invasive ovarian cancer cases, as many as 13% to 20% of women have been reported to have a *BRCA1/2* germline mutation (Schrader *et* al., 2012). A prospective cohort study including 9856 unaffected *BRCA1/2* mutation carriers showed that the cumulative risk of ovarian cancer by 80 years of age was 44% and 17% for *BRCA1* and *BRCA2* mutation carriers, respectively (Kuchenbaecker *et al*., 2017).

3.12 Prostate Cancer Risk in BRCA1/2 mutation carriers

BRCA1/2 germline mutations are also associated with an increased risk for prostate cancer (Giri *et al*., 2019), with this association being strongest for advanced or metastatic prostate cancer (Lung *et* al., 2019). A study of a large cohort of Spanish patients with prostate cancer showed that carriers of *BRCA1/2* mutations had significantly higher frequencies of aggressive cancer, nodal involvement, and distant metastasis compared with non-carriers (Castro *et al.,* 2013). In a sample of 692 men with metastatic prostate cancer, unselected for family history or age at diagnosis, 5.3% had a *BRCA2* mutation, and 0.9% had a *BRCA1* mutation (Pritchard *et al*., 2016). In addition, analyses from a treatment center database showed that the combined *BRCA1/2* and *ATM* mutation carrier rate was significantly higher in lethal prostate cancer patients (6.07%) than localized prostate cancer patients (1.44%). This study also showed that mutation carriers with prostate cancer had significantly decreased survival, compared with non-carrier patients (5 years *vs.* 16 years, respectively) (Na *et al*., 2017). A fraction of prostate cancer patients with Ashkenazi Jewish ancestry has been found to be associated with *BRCA1/2* mutations, with gene mutation frequencies ranging from 0% to 2% and from 1% to 3%, for *BRCA1* and *BRCA2*, respectively (Gallagher *et al*., 2010).

3.13 Pancreatic Cancer Risk in BRCA1/2 mutation carriers

Prior to the increasingly common use of panel testing, studies showed that *BRCA1/2* germline mutation frequencies in pancreatic cancer cases ranged from 1% to 11% for *BRCA1* and from 0% to 17% for *BRCA2* (Holter *et al*., 2015). However, these studies included only patients with familial pancreatic cancer (Zhen *et al*., 2015) or those of Ashkenazi Jewish ancestry (Lucas *et al*., 2013) both of whom may have a greater likelihood of testing positive for a *BRCA1/2* germline mutation. More recently, gene panel testing confirmed that some pancreatic cancers harbor actionable *BRCA1/2* pathogenic or likely pathogenic variants (Hu *et al*., 2018). Moreover, patients with pancreatic cancer and Ashkenazi Jewish ancestry proved to have a greater likelihood of testing positive for a *BRCA1/2* mutation, with prevalence of detected mutations in this group ranging from 5.5% to 19% for various genes, with mutations being more common for *BRCA2* (Salo-Mullen *et al*., 2015).

3.14 Gene variants other than BRCA1/2 associated with Breast/Ovarian Cancer

Prior to 2020, the NCCN Guidelines for "Genetic/Familial High-Risk Assessment: Breast, Ovarian, and Pancreatic" focused largely on testing criteria for *BRCA1/2* and appropriate risk management for carriers of a *BRCA1* or a *BRCA2* pathogenic or likely pathogenic variant. There is now strong evidence that genes beyond *BRCA1/2* confer markedly increased risk of breast and/or ovarian cancers. These genes include *ATM, BARD1, BRIP1, CDH1, CHEK2, NBN, PALB2, PTEN, RAD51C, RAD51D, STK11*, and *TP53*. Pathogenic/likely-pathogenic variants associated with Lynch syndrome and neurofibromatosis type 1 (*NF1*) can also increase breast/ovarian cancer risk. The recommendations for cancer risk management intervention for carriers of pathogenic variants associated with breast and/or ovarian

cancer risk are based on absolute lifetime risk estimates. Cancer risk management intervention may be recommended when a carrier's absolute risk exceeds that of the average-risk population (i.e., 12% - 13% for breast cancer and 1% - 2% for ovarian cancer) (Domchek *et al*., 2019). Quality of the evidence supporting risk estimates should also be evaluated when determining appropriate risk management for carriers of a pathogenic or likely pathogenic variant. Lower penetrance genes that may be included as part of multi-gene testing but for which there is currently insufficient evidence of an association with breast and/or ovarian cancer include: *FANCC, MRE11A, MUTYH* heterozygosity, *RECQL4, RAD50, RINT1, SLX4, SMARCA4* and *XRCC2*. Risk management recommendations for these genes should consider family history and other clinical factors (Apostolou *et al*., 2013).

3.15 Colorectal Cancer predisposition

Colorectal cancer (CRC) is the fourth most frequently diagnosed cancer and the second leading cause of cancer death in the United States. The incidence of CRC per 100,000 individuals decreased from 60.5 in 1976 to 46.4 in 2005 (Cheng *et al*., 2011). In addition, mortality decreased by almost 35% from 1990 to 2007 (Siegel *et al*., 2011), and by 53% from 1970 to 2016 (Siegel *et al.*, 2019). These improvements in incidence and mortality are thought to be in part a result of cancer prevention and earlier diagnosis through screening and treatment modalities. Despite the observed improvements in the overall CRC incidence rate, a retrospective cohort study of "Surveillance, Epidemiology, and End Results (SEER)" colorectal cancer registry found that the incidence of CRC in patients younger than 50 years has been increasing (Bailey *et* al., 2014). The authors estimate that the incidence rates for colon and rectal cancers will increase by 90.0% and 124.2%, respectively, for patients 20 to 34 years of age by 2030. The cause of this trend is currently unknown.

CRC often occurs sporadically, but familial cases are also common. Genetic susceptibility to CRC associated with high-penetrance genes includes well-defined syndromes such as Lynch syndrome (also known as Hereditary Non Polyposis Colorectal Cancer, or HNPCC), familial adenomatous polyposis (FAP), and *MUTYH*-associated polyposis (MAP) (see Fig. 4).

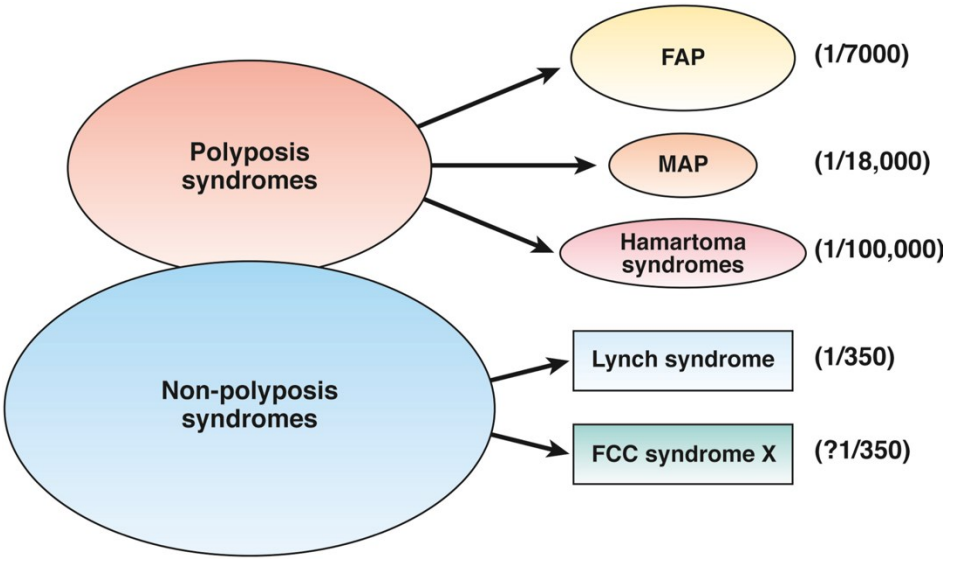


Fig. 4. Main hereditary CRC syndromes with population frequency estimates (from Stoffel and Boland, 2015)

3.16 Lynch Syndrome

The Lynch Syndrome (LS) is the most common form of genetically determined CRC predisposition, accounting for 2% to 4% of all CRC cases. Estimates suggest as many as 1 in every 250-1000 individuals in the general population may be carriers of a lesion in a gene associated with LS. In LS cancer predisposition is dominantly transmitted and results from a germline pathogenic variant in a DNA *MMR* gene. The causative genes are *MLH1*, *MSH2*, *MSH6*, and *PMS2* (Tanakaya, 2019). In addition, LS can be caused by deletions in the *EPCAM* gene, which lead to hypermethylation of the *MSH2* promoter and subsequent *MSH2* gene silencing (Rumilla *et al*., 2011).

MMR genes correct DNA mismatches generated during DNA replication, thereby preventing mutations from becoming permanent in dividing cells. Since *MMR* reduces the number of replication-associated errors, a defective system increases the spontaneous mutation rate, thus leading to microsatellite instability (MSI-H: insertions/deletions within simple repeated sequences) and to frameshift mutations in genes containing microsatellite sequences (Fig. 5). Indeed, *MMR* plays a role in the DNA damage response pathway that eliminates severely damaged cells and prevents both mutagenesis in the short term and tumorigenesis in the long term (Li, 2007).

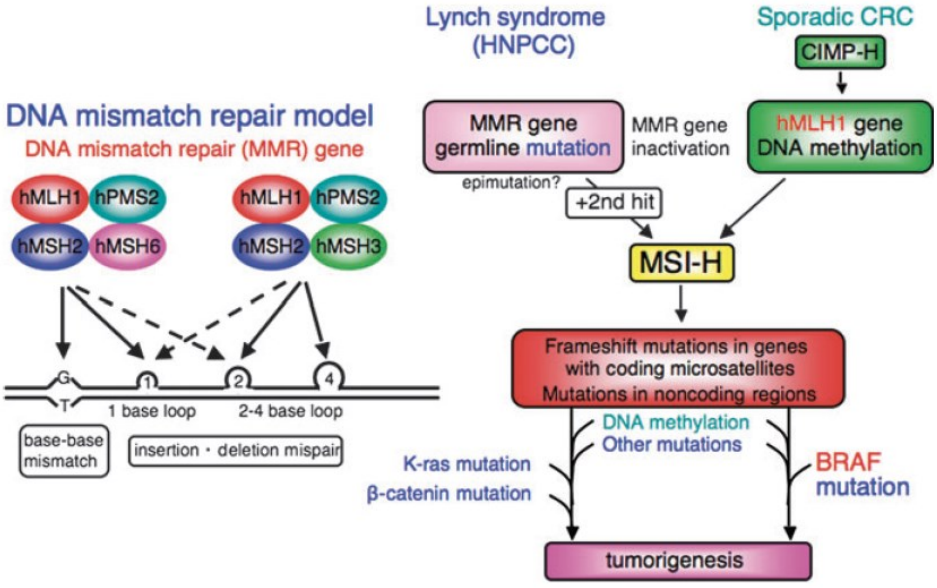


Fig. 5. Molecular model of CRC carcinogenesis associated to DNA MMR deficiency and MSI-H tumor phenotype (from Imai and Yamamoto, 2008)

The identification of the syndrome is important both for patients with cancer (because of high risk for metachronous cancers), and for their family members. After LS identification, surveillance offers an opportunity for early detection and even prevention of cancer among mutation carriers. Further, cancer site-specific evaluation and heightened attention to symptoms is also advised for all cancers that occur with increased frequency in mutation carriers, including colorectal, endometrial, gastric, ovarian, pancreatic, ureter and renal pelvis, biliary tract, brain (glioblastoma), and small intestinal cancers, as well as sebaceous gland adenomatous polyps and keratoacanthomas.

3.17 Familial adenomatous polyposis (FAP)

Familial adenomatous polyposis (FAP) is a rare, dominantly inherited condition characterized by the development of numerous precancerous polyps (adenomas) in the large intestine (colon and rectum). In the severe form, polyps develop in teen years or early 20s. The number of polyps varies from less than 100 (attenuated FAP) to thousands, and with increasing age the polyps get larger and more problematic. Eventually, one or more of these adenomas will become cancerous. Without treatment, patients with FAP have a nearly 100% lifetime risk of colorectal cancer. The chance of developing colorectal cancer increases with age; the average age at which people are diagnosed with cancer is 39. Also, other organs are susceptible to growths, either benign or malignant: bones, retinas of the eyes, stomach, duodenum, small intestine, thyroid gland, and brain.

Introduction

FAP is associated with germline mutations of *APC* tumor suppressor gene that plays a central role in WNT signaling pathway, especially regarding the degradation of β-catenin within the cell cytoplasm. When *APC* is mutated, the β-catenin-Tcf complex is not suppressed, leading to constitutive activation of several genes and oncogenes controlling cell growth and division. Ultimately, *APC* mutations affect the ability of the cell to maintain normal growth and function, which results in cell overgrowth and adenoma formation (Talseth-Palmer, 2017)

3.18 MUTYH-associated polyposis (MAP)

MUTYH-associated polyposis (MAP) is a rare, recessively inherited condition, characterized by the development of polyps in the colon and rectum and by an increased risk of colorectal cancer. People with harmful mutations in *MUTYH* gene can develop different types of polyps in the large intestine, including adenomas, sessile serrated polyps, and hyperplasic polyps. Most people with MAP develop 10 to 100 polyps; in rare cases, *MUTYH* mutation carriers can develop CRC without any polyps, while others have more than 100 polyps. The colorectal polyps are often found in people in their 40s and most CRC occur between the ages of 40 and 60. Approximately half of people with MAP have CRC at the time of MAP diagnosis. Patients are also at an increased risk of duodenal and thyroid cancer.

MAP, unlike the other cancer syndromes, is inherited in an autosomal recessive manner; of relevance, people who have a single *MUTYH* gene mutation have a slightly increased risk of CRC.

MUTYH is a gene involved in the repair of oxidative damage to the DNA. Accordingly, *MUTYH* defective cells accumulate mutations along DNA sequences including genes that are relevant for cell growth regulation in colonic epithelium, such as *APC* and *KRAS.* Thus*,* somatic mutations in these target genes are the crucial events associated with the carcinogenic process (Church *et al*., 2012).

3.19 Hereditary diffuse gastric cancer (HDGC)

Hereditary diffuse gastric cancer (HDGC) is a dominantly inherited form of diffuse-type gastric cancer (DGC), a highly invasive tumor characterized by late presentation and poor prognosis. HDGC syndrome is defined by the presence of a pathogenic germline variant in the E-cadherin (*CDH1*) gene in either an isolated individual with DGC, or in a family with one or more DGC cases in first- or second-degree relatives (Blair *et al*., 2020). E-cadherin is a key protein in adherens junctions and is involved in major cellular processes such as embryogenesis and maintenance of tissue architecture. E-cadherin is recognized as a tumor suppressor gene, and it is well established that *CDH1* germline inactivation causes diffuse gastric cancer as well as lobular breast cancer. In addition, in the last decade, evidence has emerged demonstrating that *CDH1* mutations can be associated with lobular breast cancer and/or congenital abnormalities, without any personal or family history of diffuse gastric cancer (Figueiredo *et al.,* 2019). The lifetime risk of gastric cancer in carriers from families with pathogenic or likely pathogenic variants in *CDH1* is very high, and

diagnosis has been made in individuals in their teens and early 20s. As a result, prophylactic total gastrectomy (PTG) is usually advised, generally between ages 20 and 30.

Germline pathogenic variants of *CDH1* were originally described in three Maori pedigrees from New Zealand with family members affected by DGC at a young age. Subsequently, germline *CDH1* pathogenic variants were identified in approximately 15 to 50 percent of kindreds selected according to the clinical criteria for HDGC as initially defined by the International Gastric Cancer Linkage Consortium (IGCLC) (Fitzgerald *et al*., 2010). These selection criteria have been progressively updated and relaxed, mainly through less restrictive age limits of patients (Blair *et al*., 2020).

3.20 Li-Fraumeni syndrome

Li-Fraumeni syndrome (LFS) is a cancer predisposition syndrome associated with high risk of a spectrum of childhood- and adult-onset malignancies. The lifetime risk of cancer in individuals with LFS is ≥ 70% for men and ≥ 90% for women. Five types of cancers account for the majority of LFS tumors: adrenocortical carcinomas, breast cancer, central nervous system tumors, osteosarcomas, and soft-tissue sarcomas. LFS is associated with an increased risk of several additional cancers including leukemia, lymphoma, gastrointestinal cancers, cancers of head and neck, kidney, larynx, lung, skin (e.g., melanoma), ovary, pancreas, prostate, testis, and thyroid. Criteria for clinical diagnosis of LFS have been established and the syndrome diagnosis is established in a proband who meets specific criteria and/or has a heterozygous germline pathogenic variant in *TP53* tumor suppressor gene (Schneider et al., 2013).

TP53 encodes p53 protein, which has been termed the “guardian of the genome” and has many important functions including DNA replication and repair processes control, epigenetic patterning of the genome, cell cycle arrest, apoptosis, autophagy, senescence, differentiation, antioxidant stress responses, and cellular energy metabolism (Schluer et al., 2017; Zerdoumi et al., 2017). In normal (unstressed) cells, p53 protein levels are kept low by a negative-regulatory feedback mechanism that is mediated by the MDM2 protein. MDM2 binds to p53, marking it for degradation. However, following the exposure to genotoxic stressors, such as ionizing radiation or other carcinogens, p53 and MDM2 become phosphorylated, which weakens the MDM2-p53 bond. The weakened MDM2-p53 interaction lessens the degradation of p53, which allows p53 to accumulate in the cell. The absence of normal p53 and/or the accumulation of abnormal p53 adversely impacts the expression of many downstream genes that regulate critical cellular processes including cell cycle arrest, DNA repair, apoptosis, and senescence and ultimately leads to genomic instability and malignant transformation (Valdez *et al*., 2017; (Zerdoumi *et al*., 2017).

3.21 Familial Melanoma

Melanoma is the most aggressive of the common skin cancers, being responsible for 75% of deaths from skin cancer. Melanoma incidence is rapidly increasing especially in Caucasian populations. Although development of melanoma during childhood is rare, it can appear at any age and is the second most diagnosed cancer among patients under 30 years old. For this reason, melanoma is one of the cancers with more years of productive life lost. The identification of individuals at high risk of developing melanoma is essential to reduce melanoma mortality, and prevention and early detection programs can be implemented. Melanoma etiology is complex and heterogeneous as it involves environmental and genetic risk factors.

Personal history of melanoma increases 5-8% the risk of developing a second melanoma, while family history of melanoma has been widely associated with an increased melanoma risk. Approximately 5-10% of cases occur in a familial context. In cancer families, melanoma susceptibility is inherited following an autosomal dominant pattern, with incomplete penetrance. Melanoma high risk genes are defined as genes that when mutated in an individual confer a high risk of melanoma and are usually associated with multiple melanoma cases within the family (Ribeiro Moura Brasil Arnaut *et al*., 2021).

Cyclin-dependent kinase inhibitor 2A (*CDKN2A*) was the first gene associated with melanoma susceptibility. The *CDKN2A* gene maps in 9p21 region and encodes two tumor suppressor proteins, p16INK4A and p14ARF, *via* differential splicing and alternative reading frames. The p16INK4A protein, encoded by the α-transcript (composed by exon 1α, 2 and 3), promotes the arrest of the cell cycle in the G1 phase by inhibiting RB (retinoblastoma protein) phosphorylation through cyclin-dependent kinase 4 (CDK4). The β-transcript (composed by exon 1β, 2 and 3) encodes p14ARF and acts through the p53 pathway inducing cell cycle arrest or promoting apoptosis. Furthermore, both p53 and p16INK4A play an important role on cell damage response and senescence (Fig. 6). *CDKN2A* mutations are found in around 20% of melanoma-prone families, but the mutation frequency can range from 5% to 72% depending on the criteria used to select probands and families and on the populations and geographic areas (Ribeiro Moura Brasil Arnaut *et al*., 2021).

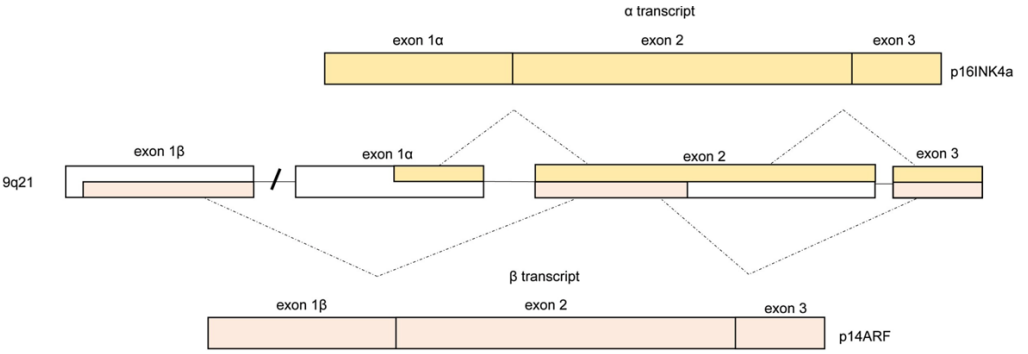


Fig. 6. CDKN2A gene encodes two alternatively spliced transcripts, p16INK4a and p14ARF. Exons 1α, 2 and 3 encode p16INK4a, while exon 1β, spliced to exons 2 and 3 in a different reading frame and transcribed using a different promoter, encodes p14ARF (from Ribeiro Moura Brasil Arnaut et al., 2021).

CDK4 was the second high risk melanoma susceptibility gene identified. *CDK4* is an oncogene located within the 12q14 chromosomal region and encodes a protein that controls cell cycle progression through the G1 phase. Mutations in this gene have been described in 17 melanoma-prone families and in all of them the mutation affects the same amino acid (Arginine 24) (Ribeiro Moura Brasil Arnaut *et al*., 2021). This amino acid is in the p16INK4A binding domain of the CDK4 protein. Thus, when *CDK4* is mutated, p16INK4A cannot inhibit the CDK4 kinase activity resulting in the progression of the cell cycle. CDK4 mutation carriers phenotypically behave similarly to p16INK4A mutation carriers: this is consistent with the functional impact that mutations in both proteins have at the cellular level, which results in the activation of the same pathway (Potrony *et al*., 2015).

4. Aims

My PhD thesis experimental work has been performed at the Clinical Genomics Unit - CGS - European Institute of Oncology, Milan. In this lab I perform genetic tests in patients with suspected genetic predisposition to cancer with the aim to identify germline pathogenic variants associated with the disease.

Literature reports that hereditary cancer predisposition syndromes are responsible for approximately 10% of all diagnosed cancer cases (Tsaousis *et al*., 2019). Patients investigated at the Clinical Genomics Unit are selected according to specific consensus criteria, including clinical features, family history of cancer and age of the disease onset. Genetic testing is specifically focused on the two most common cancer syndromes, *i.e.* Hereditary Breast and Ovarian Cancer syndrome (HBOC) and Lynch Syndrome (LS). In addition, more rare conditions are investigated, including Familial Melanoma (FM), Hereditary Diffuse Gastric Cancer syndrome (HDGC), Familial Adenomatous Polyposis (both FAP and MAP), and Li Fraumeni (LF) syndrome. Following genetic counselling, we search for germline mutations in susceptibility genes with DNA next generation sequencing (NGS)-based approaches. These approaches rely on cancer gene panels, allowing us not only to investigate the specific genes already known to be associated with each syndrome (*e.g. BRCA1/2* genes for HBOC, or *TP53* gene for LF syndrome), but also to extend the analysis to other genes, simultaneously.

The identification of the disease-causative genetic variants is important both for patients and at-risk relatives, with clinical management implications for both affected and unaffected individuals. In affected subjects, the identification of the genetic cause of the inherited condition can guide surgical management and, in some cases, systemic treatment. Furthermore, the definition of the specific syndrome can guide a personalized follow-up program to incorporate surveillance and prevention strategies of secondary malignancies associated with the syndrome itself. In unaffected family members, the presence of the disease-causative variant means an increased risk of developing the disease, allowing mutation carriers to benefit from enhanced prevention strategies.

We retrospectively analyzed 2641 patients from 2019 to 2022 to assess the presence of pathogenic genetic alterations responsible of the diagnosed cancers. By the NGS analysis of our patient cohort, we aimed at verifying the correspondence between the tumor type affecting the patient and the mutated gene, according to the known and expected genotype-phenotype correlations. Of relevance, data derived from recent NGS studies show that a notable subset of hereditary cancer cases is missed if screening is solely performed by syndrome-based specific gene testing. Multi-gene panels usually include high and moderate penetrance genes and, in many cases, some low or yet unknown risk genes. Data concerning their contribution to

cancer risk is constantly increasing, allowing a more accurate gene penetrance stratification.

In our study, multigene NGS analysis was carried out to identify cancer-susceptibility variants and to assess the applicability and utility of this type of analysis. In addition to mutations in expected genes associated with the different syndromes, the use of multi-gene panels allowed us to detect unexpected germline mutations in genes not related to the tumor type, thus offering new preventive screenings for patients and new guidelines for optimization of medical surveillance and therapy.

Moreover, the application of multi-gene panels led to the identification of many variants of unknown significance (VUSs). It is very important to deeply investigate novel germline variants and to assess their pathogenicity in those patients who have a strong family history of cancer. Of relevance, many VUSs undergo category reclassifications over time (*e.g.* from likely pathogenic to benign) as more information is available from data bases reporting genetic variants in the general population and in underrepresented minorities. By our screening approach, we succeeded to reclassify some *BRCA1*/*BRCA2* VUSs identified in patients who underwent a genetic test in the past. These patients had an amended genetic report for physicians who, in turn, used this new information in clinical decision making.

Besides single nucleotide variants (SNVs), we also searched for gross genomic rearrangements (copy number variants - CNVs), such as deletions and duplications. Gross deletions in cancer predisposition genes are known to be pathogenic, while the significance of most of gross duplications in cancer predisposition genes is unknown. We were able to map and characterize an unknown duplication in *CHEK2* gene to give it a precise clinical significance and, consequently, to offer a new medical surveillance program to the carrier patient and to his relatives.

The selection of patients for germline genetic testing has long relied on the presence of a strong family history of breast/ovarian or colon cancer, for HBOC and LS, respectively. However, it is now clear that this criterion results in substantial numbers of subjects with a germline mutation in cancer predisposition genes being missed (George *et al*., 2015). In this context, we started a new study, called "Persona project", that aims at identifying germline pathogenic variants in patients affected with triple negative breast cancer (TNBC), regardless of their family history. Various studies have demonstrated that 15%-20% of women with TNBC carry *BRCA1* or *BRCA2* germline mutations, and that 70% of breast cancers in *BRCA1* mutation carriers are "triple-negative" (Stevens *et al*., 2013). In general, *BRCA1* and *BRCA2* appear to account for approximately one third of germline mutations in breast cancer, with smaller contributions from other cancer predisposition genes (Dite *et al*., 2003). We were interested in assessing the presence of *BRCA1* or *BRCA2* mutations in individuals affected by TNBC, without applying any other selection criterion, and in evaluating the prognostic and therapeutic predictive role of these gene defects

(affecting homologous recombination) *vs.* other genetic alterations. This study is expected to allow the identification of subjects at risk for the development of secondary malignances and to offer genetic testing to family members of mutation carriers. Moreover, the result of genetic testing can influence the decision for medical therapy (*e.g.* the use of platinum derivatives *vs*. PARP inhibitors) and dictate risk-reducing strategies that may imply bilateral salpingo-oophorectomy and mastectomy or long term medical approaches. Of relevance, the genetic heterogeneity characterizing breast cancer predisposition creates the need for comprehensive testing with panels of multiple genes, all at once, to discover mutations in genes other than those known to be associated with the disease. We were interested in characterizing every single patient carrying mutations in new genes that can influence cancer predisposition, disease progression and therapy. We were also interested in evaluating the occurrence of the most frequent *BRCA1* and *BRCA2* pathogenic variants and of the low penetrance disease-associated variants (daSNPs). Such a deep genetic analysis, together with follow-up data, is expected to improve therapeutic approaches and to refine personalized treatment and care strategies.

5. Materials and Methods

5.1. Data Set

We collected and analyzed clinical and genetic data of 2641 patients who received Oncogenetic Counseling at The European Institute of Oncology in Milan from 2019 to 2022. The patient samples were tested for germline variants in cancer predisposition genes following the Clinical Practice Guidelines in Oncology (NCCN Guidelines®) testing criteria. The following data were collected for each patient: gender, age (calculated from birthdate to the end of data collection), date and place of birth, health status (cancer diagnosis), cancer site.

5.2. Molecular Analysis

A peripheral blood sample was taken from each patient for germline genetic testing. Mutational analysis of coding sequences and intron-exon boundaries of cancer predisposition genes was carried out with next generation sequencing (NGS) and Sanger Sequencing. The presence of large genomic rearrangements - copy number variants (CNVs) - was detected with Multiplex Ligation-dependent Probe Amplification (MLPA).

All the detected *BRCA1* and *BRCA2* gene variants were analyzed following the Evidence-based Network for the Interpretation of Germline Mutant Alleles (ENIGMA) Consortium guidelines. The same variants in *BRCA1* and *BRCA2* genes and the ones identified in all the other cancer predisposition genes were checked following *Varsome* database and *Clinvar* public report archive, and classified according to the International Agency for Research on Cancer (IARC) recommendation, using a system dividing the variants into 5 classes:

benign (class I),

likely benign (class II),

variant of uncertain significance (VUS, class III),

likely pathogenic (class IV),

pathogenic (class V).

All the mutations identified were named in accordance with the Human Genome Variation Society (HGVS) nomenclature.

Materials and Methods

5.3 Sample processing

A DNA sample was obtained from peripheral blood of each patient by using the MagCore Super Automated Acid Nucleic Extractor (RBSBioscience, Taiwan). The High Sensitivity Qubit 3.0 Fluorometer (Life Technologies, Carlsbad, CA) was used to quantify the DNA by measuring the fluorescence intensity of a dye binding to the double-stranded DNA (dsDNA).

Following the request of either the patient's physician or a genetics professional to identify variants in specific genes, three different NGS sequencing libraries were generated. The NGS systems were the following:

a) Trusight Sequencing Cancer Panel (Illumina, San Diego, CA), which includes 94 genes associated with both common (*e.g.* breast, colorectal) and rare cancers. The Panel also includes 284 SNPs found to be associated with cancer through genome-wide association studies (GWAS);

b) Sophia Genetics Custom Panel (Sophia Genetics, Geneva, CH), which includes the 29 most clinically relevant genes associated with breast and ovarian cancer, Lynch and intestinal polyposis syndromes, Li-Fraumeni syndrome, hereditary Melanoma;

c) Devyser BRCA NGS (Stockholm, Sweden), to directly sequence *BRCA1* and *BRCA2* genes for very urgent cases.

Materials and Methods

Here below are reported:

1) the gene sets of a), b), and c) Panels

a)

AIP	CEBPA	FANCA	KIT	PRF1	SLX4
ALK	CEP57	FANCB	MAX	PRKAR1A	SMAD4
APC	CHEK2	FANCC	MEN1	PTCH1	SMARCB1
ATM	CYLD	FANCD2	MET	PTEN	STK11
BAP1	DDB2	FANCE	MLH1	RAD51C	SUFU
BLM	DICER1	FANCF	MSH2	RAD51D	TMEM127
BMPR1A	DIS3L2	FANCG	MSH6	RB1	TP53
BRCA1	EGF3	FANCI	MUYH	RECQL4	TSC1
BRCA2	EPCAM	FANCL	NBN	RET	TSC2
BRIP1	ERCC2	FANCM	NF1	RHBDF2	VHL
BUB1B	ERCC3	FH	NF2	RUNX1	WRN
CDC73	ERCC4	FLCN	NSD1	SBDS	WT1
CDH1	ERCC5	GATA2	PALB2	SDHAF2	XPA
CDK4	EXT1	GPC3	PHOX2B	SDHB	XPC
CDKN1C	EXT2	HNF1A	PMS1	SDHC	
CDKN2A	EZH2	HRAS	PMS2	SDHD	

b)

APC	CDKN2A	MUTYH	RAD51C
ATM	CHEK2	NBN	RAD51D
BARD1	EPCAM	PALB2	STK11
BRCA1	FAM175A	PIK3CA	TP53
BRCA2	MLH1	PMS2	XRCC2
BRIP1	MRE11A	PMS2CL	
CDH1	MSH2	PTEN	
CDK4	MSH6	RAD50	

c)

BRCA1
BRCA2

2) the list of the **predisposition genes** associated with the most common cancers; the **syndromes/tumor-types** associated with the different genes:

BRCA1* and *BRCA2: (*HBOC, Hereditary Breast Ovarian Cancer syndrome*) - breast and ovarian cancer; *BRCA2*: prostate cancer, pancreatic cancer

TP53: (*LF, Li Fraumeni syndrome*) breast cancer, osteosarcomas, soft tissues sarcomas, leukemia, brain cancer, and adrenocortical carcinoma

***MMR* genes**: (*MLH1, MSH2, EPCAM, MSH6, PMS2*): (*LS, Lynch syndrome*) - colorectal, uterine (endometrial), stomach, kidney, bladder, and ovarian cancers

APC: (*FAP, Familial Adenomatous Polyposis*) - colorectal, small intestine and pancreatic cancer

MUTYH: (*MAP, MUTYH associated polyposis*) - colorectal, gastric, and liver cancer

CDH1: (*HDGC, Hereditary Diffuse Gastric Cancer syndrome*) - gastric and breast cancer (rarely: colorectal, thyroid, and ovarian cancer)

ATM: breast cancer

CHEK2: breast and colon cancer

PALB2: breast cancer

CDKN2A*/*CDK4: melanoma

5.4 Sequencing library preparation

5.4.1 Hybridization capture-based sequencing library preparation - Nextera flex for enrichment - Illumina

100 ng of genomic DNA was fragmented and tagged with adapter sequences enzymatically; adapter-tagged DNA was then washed on before a limited-cycle PCR amplification. The PCR step added pre-paired 10 base pair Index 1 (i7) adapters, Index 2 (i5) adapters, and sequences required for sequencing cluster generation. Double-sided bead purification was performed to purify the amplified libraries. The libraries were pooled by mass before an overnight hybridization step at 62°C in which targeted regions of the DNA were bound with capture probes of the panel genes. After this step Streptavidin Magnetic Beads were used to capture probes hybridized to the targeted regions of interest and a final PCR cycle amplified the enriched library. The final library was then purified, and unwanted products were removed using AMPure XP Beads. The enriched library was qualified by Agilent Tapestation system (Agilent Technologies, Santa Clara, Ca) that measured the average size of the DNA fragments. The library was diluted to a final loading concentration of 16 pM (starting from 24 samples) or 10 pM (starting from 12 samples) and prepared for sequencing using Illumina Miseq Reagent Kit (600 cycles for 24 samples with V3 flow cell / 300 cycles for 12 samples with Standard V2 flow cell) (Illumina) according to the manufacturer's instructions to generate paired-end reads. Following denaturation with NaOH 2N and at 95°C for 3 minutes, DNA was loaded with 3% PhiX control (Illumina) on Miseq instrument (Illumina).

5.4.2 Hybridization capture-based sequencing library preparation - Hereditary cancer solution - Sophia Genetics

Two hundreds ng of genomic DNA underwent enzymatic fragmentation, end-repairing, and A-tailing in a unique process. Tagged DNA was ligated to dual index adapters (i5-i7) and then cleaned up with freshly prepared ethanol 80%. Dual size selection was performed with AMPure XP beads to remove shorter than 300 bp and longer than 700 bp DNA fragments. Cleaned DNA was amplified by PCR and a library pooling was done. The library was hybridized with probes specific for 29 genes for 4 hrs at 65°C. Streptavidin Magnetic Beads were used to capture probes hybridized, which underwent three wash cycles to remove unbound DNA. DNA with probes was amplified by PCR and then cleaned up with ethanol 80% and AMPure XP beads. The

final library was qualified by Agilent Tapestation system and diluted to a final loading concentration of 10 pM (starting from 24 samples) or 12 pM (starting from 8 samples). The library was prepared for sequencing using Illumina Miseq Reagent Kit (500 cycles for 24 samples with Standard V2 flow cell/ 500 cycles for 8 samples with Micro flow cell) according to the manufacturer's instructions to generate paired-end reads. Following denaturation with NaOH 2N at 95°C for 3 minutes, DNA was loaded with 3% PhiX control on Miseq instrument.

5.4.3 Amplicon based sequencing library preparation – Devyser BRCA

Ten ng of genomic DNA was used to amplify *BRCA1* and *BRCA2* genes in a single-tube multiplex reaction; a total of 188 amplicons with a mean amplicon target length of 196 bp (range, 120 to 278 bp) were amplified to create sequencing libraries of the complete *BRCA1/2* genes. This PCR-based library was diluted and used to incorporate the molecular barcodes and adapter sequences into each amplicon by a second PCR reaction. Amplicon libraries were pooled to generate a sequencing library that was purified using the Devyser Library Clean kit and quantified using the High Sensitivity Qubit 3.0 fluorometer. The final library was normalized to a concentration of 20 pM and prepared for sequencing using Illumina MiSeq with MiSeq Reagent Kit (300 cycles) according to the manufacturer's instructions to generate paired-end reads. Following denaturation with NaOH 0.2N (illumina), DNA was loaded with 1% PhiX control (Illumina) on Miseq instrument (Illumina).

5.5 Automatic sample preparation system

Microlab Star liquid handling robot (Hamilton, Reno, NV) was used for Illumina and Sophia Genetics library preparation. The system was fine-tuned to the lab's specific NGS workflow. Air displacement pipetting using Hamilton tips with CO-RE (Compressed O-Ring Expansion) technology provided superior measurement accuracy, precision, and reproducibility without tip distortion or aerosol generation during tip pick-up and ejection. Each aspirate and dispense step were dynamically tracked by an internal software and the indexed and multiplexed samples were processed according to the worklist and highly reproducible.

5.6 Miseq NGS system

The MiSeq can perform genomic DNA sequencing and data analysis with base calling, alignment, variant calling, and reporting in a single run. This instrument utilizes a double-sided, single-lane flow cell and reagent cartridge supplied in kit form. Sequencing was performed by recording the synthesis of DNA strands in clusters of sample templates attached to a flow cell. NGS library was loaded into the flow cell and the fragments hybridized to the flow cell surface. Each bound fragment was amplified into a clonal cluster through bridge amplification (Fig 1.1). Sequencing reagents, including fluorescently labeled nucleotides, were added and the first base was incorporated. The flow cell was imaged and the emission from each cluster was recorded. The emission wavelength and intensity were used to identify the base. This cycle was repeated "n" times to create a read length of "n" bases (Fig. 1.2). Reads were

aligned to a reference sequence with bioinformatics software. After alignment, differences between the reference genome and the newly sequenced reads could be identified (Fig. 1.3).

Fig. 1.1) cluster amplification

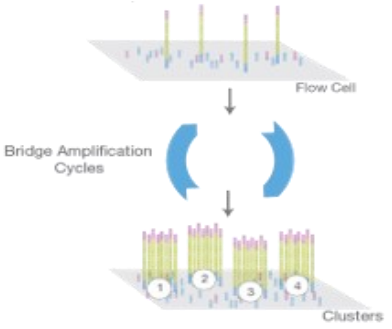


Fig. 1.2) sequencing

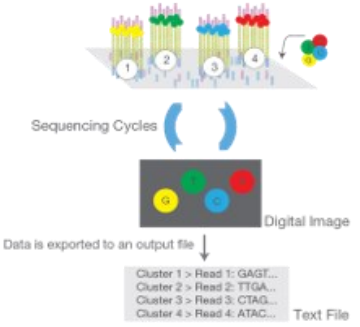
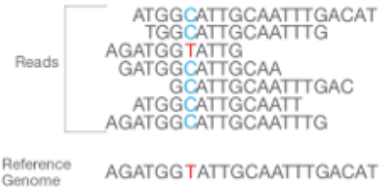


Fig. 1.3) alignment and data analysis



5.7 Bioinformatic analysis

Primary data processing (base calling, demultiplexing, and FastQ file generation) was conducted directly in the MiSeq analysis pipeline. Secondary data analysis was performed using three different bioinformatic platforms: an internal pipeline developed by our laboratory, the Sophia DDM, and Devyser Amplicon Suite software-as-a-service tool. This analysis included read alignment to the human reference genome (Genome Reference Consortium GRCh37/hg19), variant calling, visualization of the sequence reads, and report generation. The reference transcript sequences used were NM_007294.3 for *BRCA1* and NM_000059.3 for *BRCA2*; the reference transcripts used for all the other genes were the canonical ones. A high mean read quality score (Q score >35) for each amplicon and a minimun of 50× read coverage per amplicon were considered to ensure high confidence variant calling.

However, variants with lower read quality (Q score of 20 to 35) and variant fraction >20% were listed apart; any low coverage regions, if present, were also reported.

5.8 Sanger sequencing

All candidate pathogenic or likely pathogenic changes and all variants of uncertain significance (VUS) identified by NGS were confirmed by direct Sanger sequencing to allow false-positive detection. Moreover, once the disease-causing mutation was identified in a patient, the analysis was offered to its family members searching for the same mutation through Sanger sequencing. Mutation detection was performed by DNA amplification using AmpliTaq Gold® DNA Polymerase (Applied Biosystems, Waltham, MA) with 15 ng of initial DNA or Takara Taq DNA Polymerase (Kusatsu, JP) for long range amplification with 100 ng of initial DNA. PCR products were purified with EXOSAP-IT™ (Thermo Fisher Scientific, Waltham, MA) and amplified using Big Dye 3.1 Terminator chemistry (Thermo Fisher Scientific). Following denaturation with Hi-Di Formamide (Thermo Fisher Scientific), samples were loaded into the capillary array of a 3500xL Dx Genetic Analyzer (Applied Biosystems). The average PCR product contained 500 nucleotides and the maximum length that could be sequenced by the Sanger method was about 800 nucleotides. The amount of data generated by Sanger sequencing for each patient varied depending on the size of the exon to be sequenced (Fig. 2).

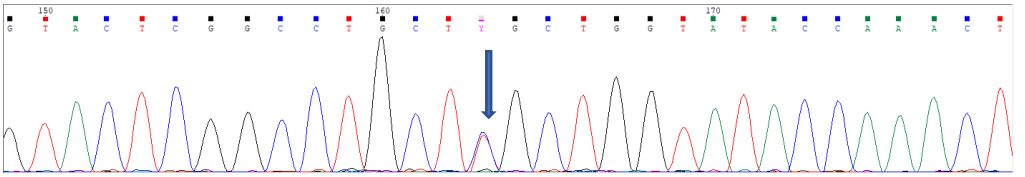


Fig. 2) Example of Sanger sequencing electropherogram: the arrow indicates the presence of a heterozygous mutation (in the same position one allele has a cytosine while the other allele has a thymine).

5.9 Multiplex ligation-dependent probe amplification (MLPA)

Deletions and/or duplications of exons or entire genes were detected by MLPA method. This was applied to confirm copy number variations identified by NGS, to detect the same gross variants in relatives of the affected carrier, and to find *de novo* large genomic rearrangements. Hundred ng of purified sample DNA was denatured, and this was followed by overnight incubation (16-20 hrs.) with MLPA probe oligos for all exons of specific genes to be investigated (*BRCA1*, *BRCA2*, *MMR* genes, *TP53*, *CDH1*, *ATM*, *RAD50*, *RAD51C*, *RAD51D*, *STK11*, *PALB2*) (MLPA kits: MRC-Holland, Amsterdam, The Netherlands).

The principle of MLPA is reported in Figure 3. Two adjacent DNA oligonucleotides are directly hybridized to their complementary target sequences on the template DNA. This reaction is followed by ligation of the two oligonucleotides to

form one probe. One oligo contains a target specific sequence with an M13 forward tail, the other contains a target specific sequence linked to a variable length 'stuffer' sequence and an M13 reverse tail. The two oligonucleotides can only be ligated if both specific parts can hybridize to adjacent template sequences. A PCR reaction is then performed making use of the M13 primer sequences on the ligation product. The amplification can only start after successful adjacent hybridization and ligation of both oligonucleotides, which ensures specificity of probe amplification. Thus, MLPA is characterized by PCR performed on the ligated oligonucleotides and not on the template DNA. The amount of ligated probes is related to the number of specific primer binding sites, making this method suitable for the detection of chromosomal deletions or amplifications.

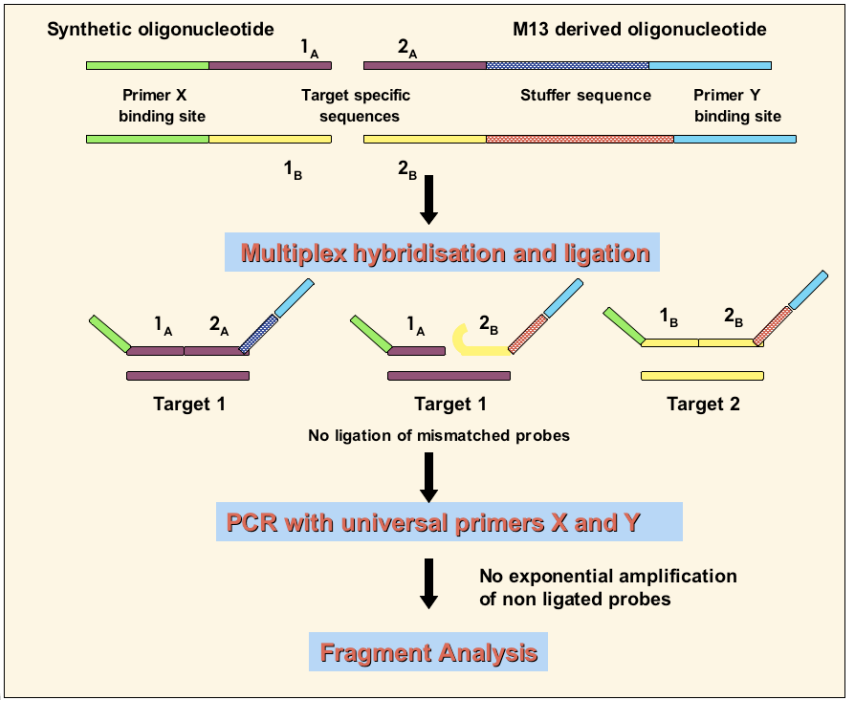


Fig. 3) Principle of MLPA. For each specific target, a set of two probes (1A&2A or 2B&2B) was designed that hybridize immediately adjacent to each other on the same target strand. Both probes consist of a short target-specific sequence and a universal forward or reverse PCR primer-binding site. In addition, one of the probes contains a so-called stuffer sequence. For each probe, the 'stuffer' part has a specific length and sequence. The long probes are M13-derived. The short probes are synthetic. After an overnight hybridization to the target DNA, the two parts of each hybridized probe are joined by a ligation reaction. Next, a PCR is carried out with a single fluorescent-labeled primer pair, which ensures that the relative yield of the PCR products is proportional to the amount of target.
The fragment analysis is carried out on an automated capillary sequencer. The multiple fragments can be distinguished based on different length. The peak area value of each product is used to calculate the relative quantity.

Following a denaturation cycle at 95°C for 3 minutes PCR products were loaded onto a capillary electrophoresis device (3500xL Dx - Applied Biosystems) and separated by length. Each fragment corresponded to a specific MLPA probe. The final step was data analysis by a specific software (Coffalyser.Net): relative copy numbers were determined by comparing the relative peak heights of reference probes and target probes in the test samples with those in reference samples with a known normal copy number (Fig. 3.1, Fig. 3.2).

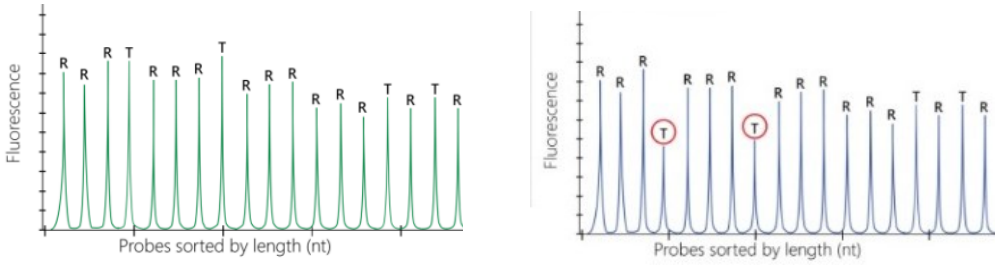


Fig.5.1) MLPA profile. Electropherogram of a test sample A (right) compared to that of reference sample (left). A relative decrease of two picks (corresponding to two specific probes) is seen in test sample A (circled in red). The peaks indicated with R correspond to the reference probes.

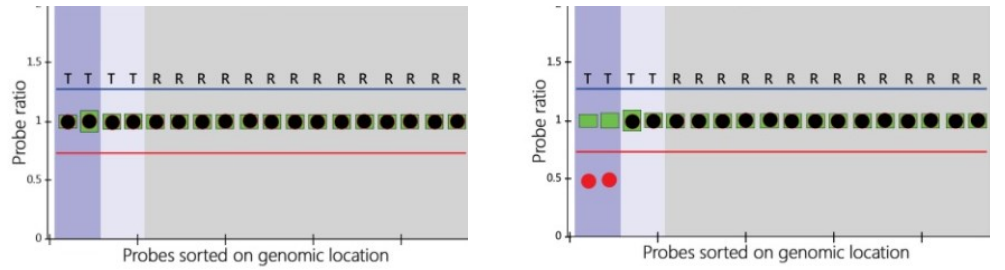


Fig. 5.2) Same results of Fig. 3.1 after normalization: probes' ratio of test sample A (right) normalized to reference sample (left). A heterozygous deletion, probe ratio 0.5, is seen in test sample (red dots). T: target probes; R: reference probes

5.10 Taqman Real-Time PCR assay

A Copy Number Assay, a copy number Reference Assay, a Master Mix (TaqPath ProAmp Master Mix, Applied Biosystems) and gDNA samples were combined in a single well or tube. The Copy Number Assay detects the target gene or genomic sequence of interest, and the Reference Assay detects a sequence that is known to be present in two copies in a diploid genome. Four replicates of each sample were performed. Ten ng of DNA template was denatured, and each set of assay primers were annealed to its specific target sequences in *BRCA1, BRCA2 and MSH2* genes.

Each TaqMan probe was annealed specifically to its complementary sequence between forward and reverse primer binding sites. During each round of PCR, the target and reference sequences were simultaneously amplified by the DNA polymerase, cleaving probes that were hybridized to each amplicon sequence. When an oligonucleotide probe was cleaved by the DNA polymerase 5' nuclease activity, the quencher was separated from the reporter dye, increasing the fluorescence of the reporter. Accumulation of PCR products could be detected in real time by monitoring the increase in fluorescence of each reporter dye at each PCR cycle. Raw data were collected, and relative quantitation analysis was performed with Applied Biosystems CopyCaller Software to determine the copy number of each genomic DNA target: lower amounts of input material (the presence of a gross deletion) resulted in higher quantification cycle (Cq) values compared to a reference sample with no copy number variation.

6. Results

I retrospectively analyzed 2641 DNA samples sequenced at the Clinical Genomics Unit (CGS) to search for DNA variants by using the Trusight Sequencing Cancer Panel, the Diagnostic Sophia Panel, and/or the Devyser *BRCA* Panel; samples were also screened by MLPA method to search for extended deletions/duplications along the sequenced genes.

Six hundred and eighty-five individuals (685/2641: 26%) proved to carry genetic alteration: 50% showed a pathogenic point mutation, 40% showed a variant of uncertain significance (VUS), while 3% showed both a pathogenic point mutation and a VUS. Among individuals with a genetic alteration, 47 (7%) proved to carry copy number variants (CNVs) and no pathogenic point mutations in the same genes (Fig. 1).

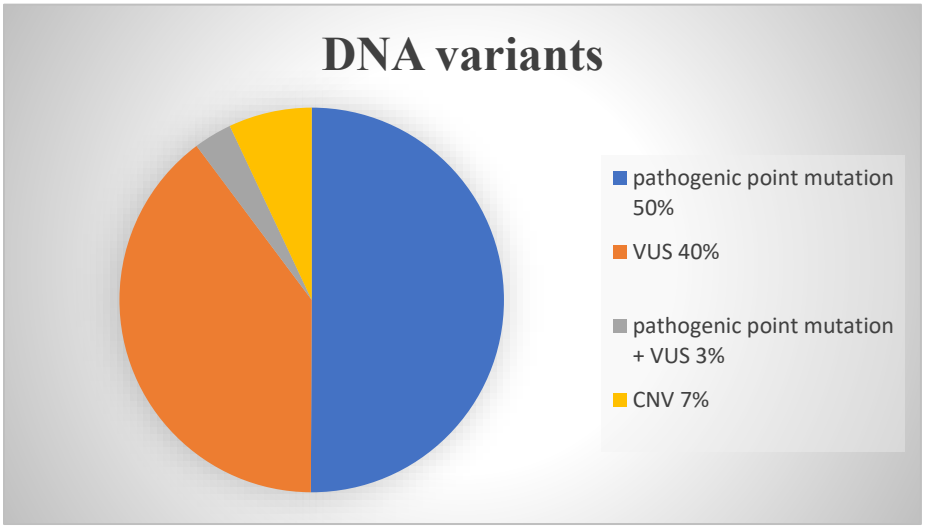


Fig.1 Type of variants identified in 685/2641 individuals screened at the CGS by NGS-based approaches and MLPA method. Subjects were investigated following the request of patient's physicians or of genetics professionals.

6.1 Copy Number Variants

Among the 47 CNV carriers, 36 (77%) proved to carry a gross deletion while 11 (23%) showed an extended duplication (Fig. 2).

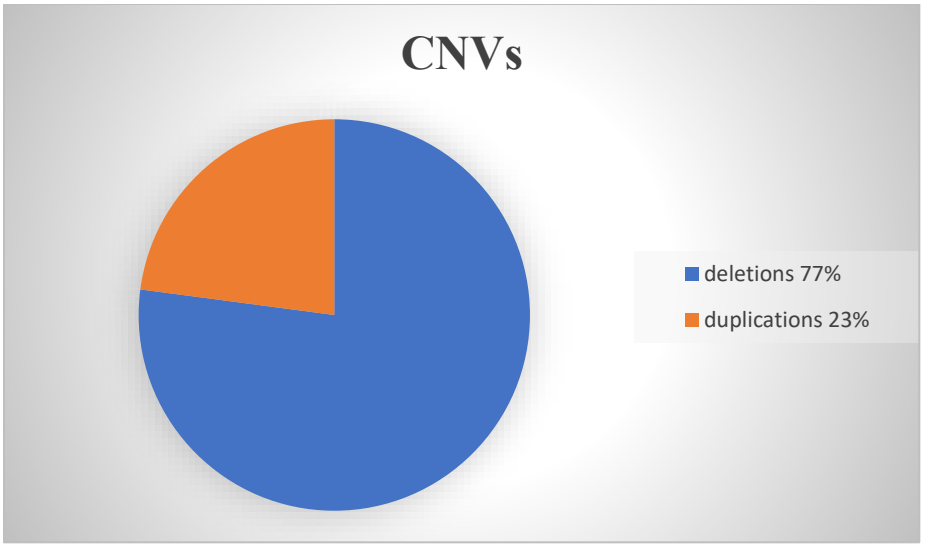


Fig.2 Rearrangements identified by MLPA along the sequenced genes in 48 carriers of CNVs.

Among the 47 individuals with CNVs, 26 (55.3%) carried deletions/duplications along the *BRCA1* gene and two showed deletions along *BRCA2*. Three individuals proved to carry deletions in *MLH1*, two in *MSH2*, two in *EPCAM* and *MSH2* and one in *MSH6* genes. Three subjects showed a duplication in *PALB2,* two in *ATM*, one showed a previously undescribed duplication in *CHEK2,* one in *BARD1*, one in *CDH1* and another in *RAD51D* gene. One individual showed a deletion in *ATM* and another in *BRIP1* gene. We identified two patients with a rearrangement in *BRCA2* gene consisting in a deletion of exons 12 and 13. Of relevance, *BRCA2* is not currently analyzed by MLPA in diagnostic settings. All the above lesions are summarized in Fig. 3. Fig. 4 shows, as for example, the MLPA of one the two patients carrying the *BRCA2* deletion of exons 12 and 13.

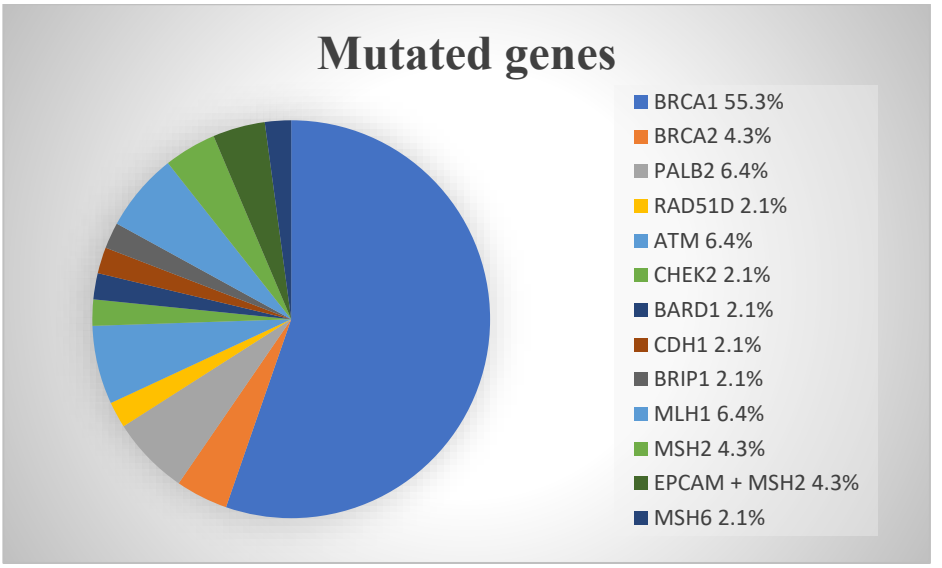


Fig.3 Genes showing gross rearrangements in the 47 CNV carriers.

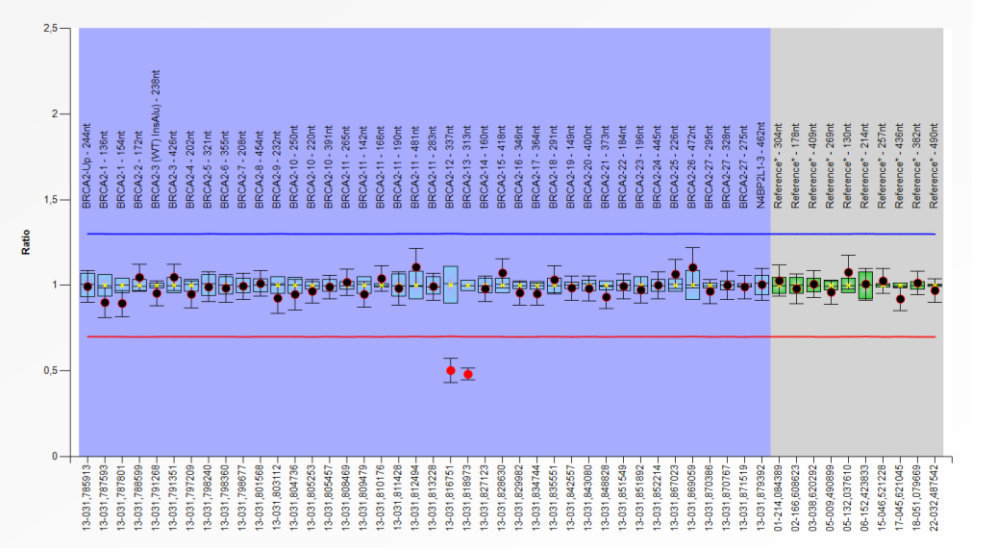


Fig.4 MLPA of a patient carrying a deletion of exons 12 and 13 of BRCA2 gene. The MLPA assay was performed with a kit containing probes for BRCA2 exons and for control sequences (in green). The deletion is denoted by the red spot below the deletion cut-off line (red) in the ratio chart. Longitudinal axis represents the final ratio after comparing the patient and control sample peak intensities by a dosage ratio and transverse axis represents the sequences covered by the MLPA kit.

Among the 26 individuals with CNVs in *BRCA1* gene, 5 (19.2%) proved to carry deletions of exons 21-22, 7 (27%) carried deletions of 1-2 exons and upstream region and one individual showed the same deletion also including exon 3. Five individuals had deletions of exon 20, 3 subjects carried deletions of exons 16-17, and 1 showed deletions of exons 19. One subject had a deletion of exons 5-6-7-8-9-10

while another carried a deletion of exon 13. One individual showed a duplication of exon 13 and another four copies of the entire *BRCA1* gene. *BRCA1* rearrangements are summarized in Fig. 5.

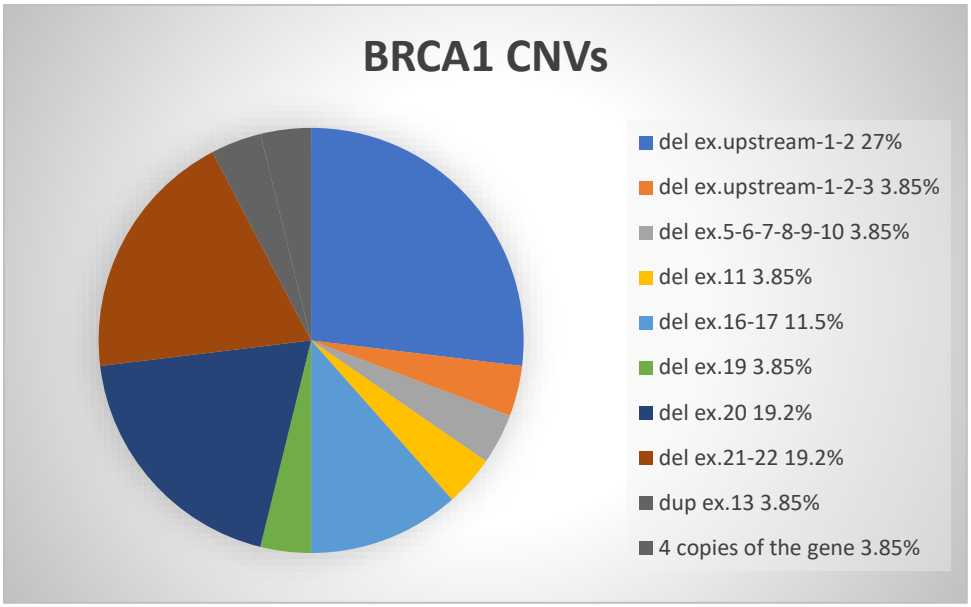


Fig.5 BRCA1 rearrangements identified.

All deletions, as well as exon-13 duplication, affecting *BRCA1* gene are disease causative: they include critical domains for the protein function and are classified as "Class 5 - pathogenic variants" according to the ClinGen Sequence Variant Interpretation (SVI). On the contrary, the four copies of the *BRCA1* locus represent a still uncharacterized CNV which is, therefore, classified as variant of unknown significance (VUS) (Fig. 6).

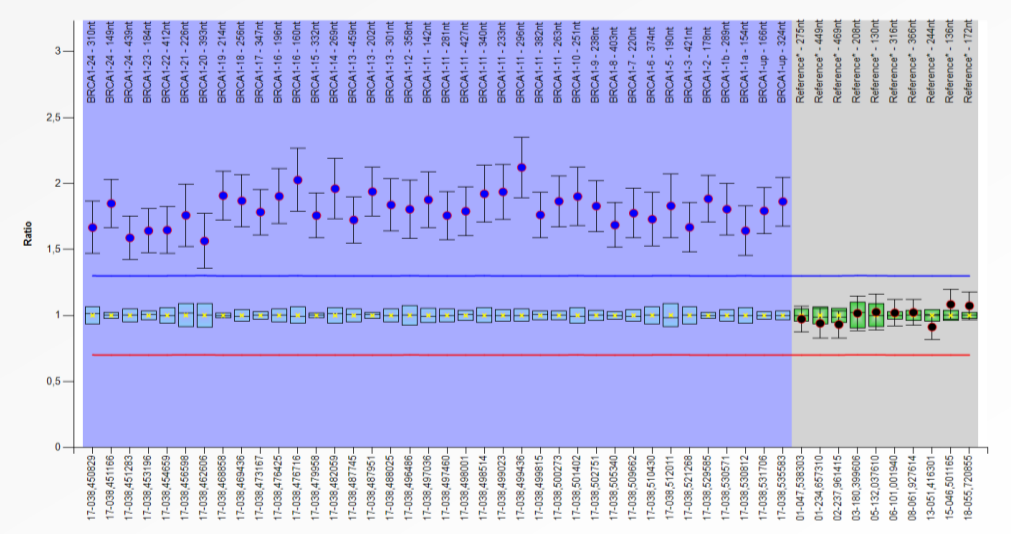


Fig.6 MLPA of the patient carrying 4 copies of the BRCA1 gene. The MLPA assay was performed with a kit containing probes for BRCA1 exons and for control sequences (in green). The 4 copies of the entire gene in this individual are highlighted by probes (in blue) corresponding each of them to single exons.

Among individuals with CNVs in *MLH1* gene, one proved to carry a deletion of the entire gene, one had deletions of exons 10 to 19, and another carried deletions of exons 14 to 19. Regarding *MSH2* gene, one individual had a deletion of exon 7 and another individual of exon 10. One subject had a deletion of the entire *EPCAM* gene together with the upstream region of *MSH2*; another subject had the deletion of the entire *EPCAM* gene together with the exons 1-2-3-4-5-6-7 of *MSH2* gene. One individual was carrier of an exons 3-4-5-6 deletion in *MSH6* gene. All deletions in *MMR* genes are classified as “Class 5 - pathogenic variants”.

One patient had a duplication of exons 1 to 12 in *PALB2* gene (Fig. 7): this rearrangement is currently classified as VUS. Another patient had a duplication of exon 11 in *PALB2* gene and another in exon 13. These two duplications are classified as “pathogenic”.

One subject showed a duplication of the entire *BARD1* gene: this gross CNV is currently classified as VUS.

One patient had a duplication from exon 3 to 3’UTR region in *CDH1* gene: this gross CNV is classified as VUS.

One individual had a deletion of exon 8 in *BRIP1* gene: this rearrangement is classified as “pathogenic”.

One patient had a deletion of exon 63 in *ATM* gene: this rearrangement is classified as “pathogenetic”. Two patients had a duplication of *ATM* exons 62-63 which is considered as a VUS.

One individual showed the duplication of the entire *RAD51D* gene: this rearrangement has never been described and is currently classified as VUS.

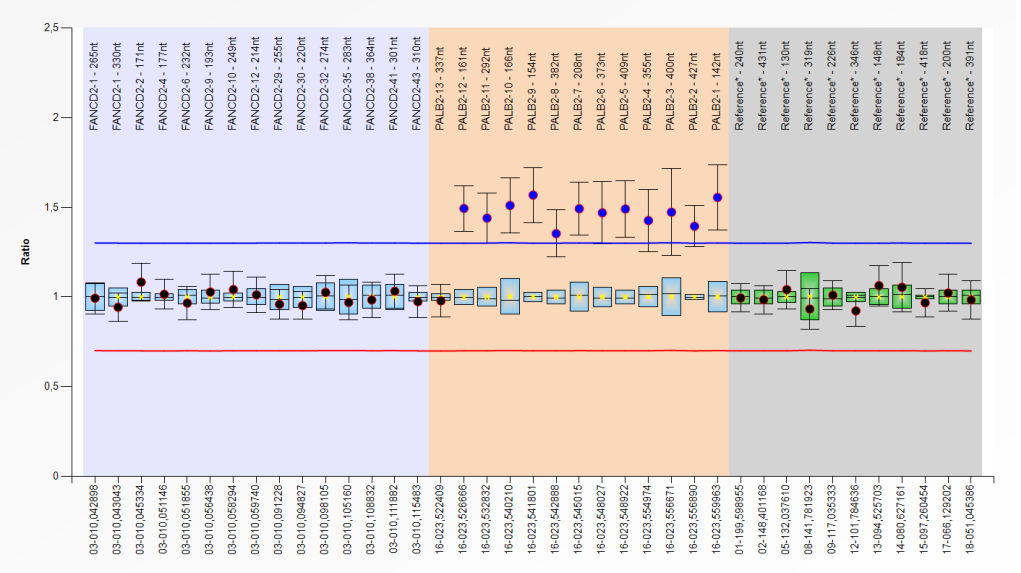


Fig.7 MLPA of a patient carrying a duplication within PALB2 gene. The MLPA assay was performed with a kit containing probes for exons of PALB2 and FANCD2 genes and for control sequences (in green). The duplication of exons 1-12 of PALB2 gene in this individual is highlighted by probes (in blue) corresponding each of them to single exons.

One individual had a duplication of exon 2 in *CHEK2* gene. This rearrangement has never been reported in other patients; however, it can be classified as "pathogenetic" since it causes a premature termination codon along *CHEK2* mRNA.

Among individuals with CNVs in *BRCA1* gene, 17 were affected with breast cancer, one with ovarian cancer, one with prostate cancer, while the remaining 6 were healthy. The individual carrying four copies of *BRCA1* gene was affected with breast cancer. The two individuals with a deletion in *BRCA2* gene were affected with breast cancer. Among patients with CNVs in *MMR* genes, 3 were affected with colon cancer, one with endometrial cancer, one with breast cancer while three were healthy. Two patients carrying CNVs in *PALB2* gene were affected with breast cancer while another was affected with colon cancer. Two individuals with CNVs in *ATM* gene were affected with breast cancer while another was healthy. The individuals with rearrangements in *BARD1*, *CDH1*, *RAD51D* and *CHEK2* genes were affected with breast cancer while the patient with a deletion in *BRIP1* gene was affected with ovarian cancer.

Overall, 75% of CNV pathogenic variant carriers had developed a cancer, whereas 25% were asymptomatic when tested (Fig. 8). Overall, in 84% of the patients with CNVs the tumor type corresponded to the mutated gene accordingly to the

expected genotype-phenotype correlations. Of relevance, in six cases the obtained results were unexpected: *BRCA1* gene was associated with prostate cancer, *RAD51D* gene with breast cancer, *MSH2* gene with breast cancer, *PALB2* gene with colon cancer, *BARD1* gene with ovarian cancer and *CHEK2* gene with colon cancer. All VUS carriers had developed a cancer. The average ages at diagnosis were 47 and 45 for subjects with pathogenic CNVs and for subjects with VUS, respectively.

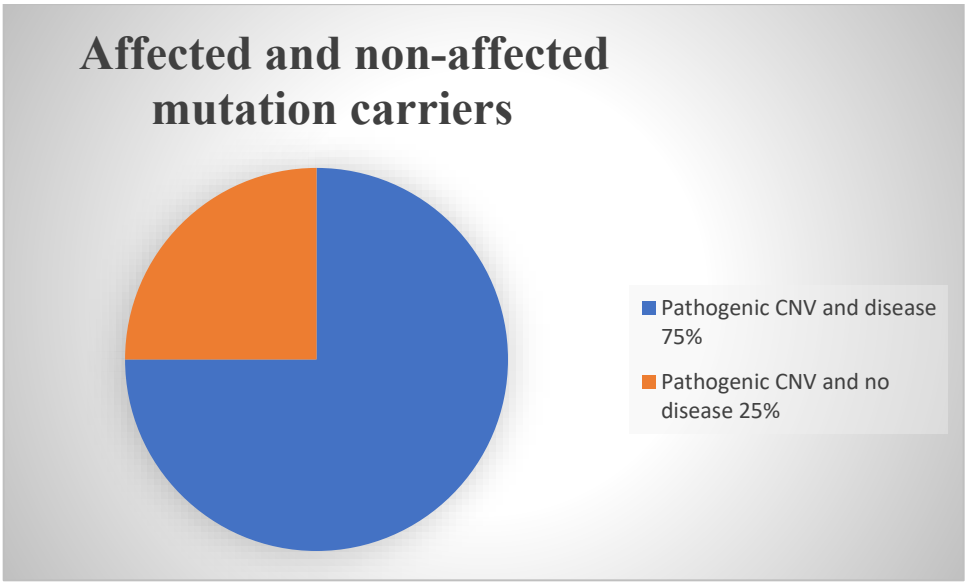


Fig.8 Affected and non-affected subjects among carriers of pathogenic CNVs.

6.2 Point mutations

6.2.1 BRCA1/BRCA2 genes sequencing

I retrospectively analyzed 2077 DNA samples sequenced at the Clinical Genomics Unit to search for point mutations following the request for *BRCA1* and *BRCA2* genetic test.

One-hundred-nineteen individuals (5.7%) proved to carry a pathogenetic variant in *BRCA1* and 120 (5.8%) had a pathogenic variant in *BRCA2*. Fifty-three patients (2.6%) carried a VUS in *BRCA1* and 97 (4.7%) had a VUS in *BRCA2* (Fig. 9).

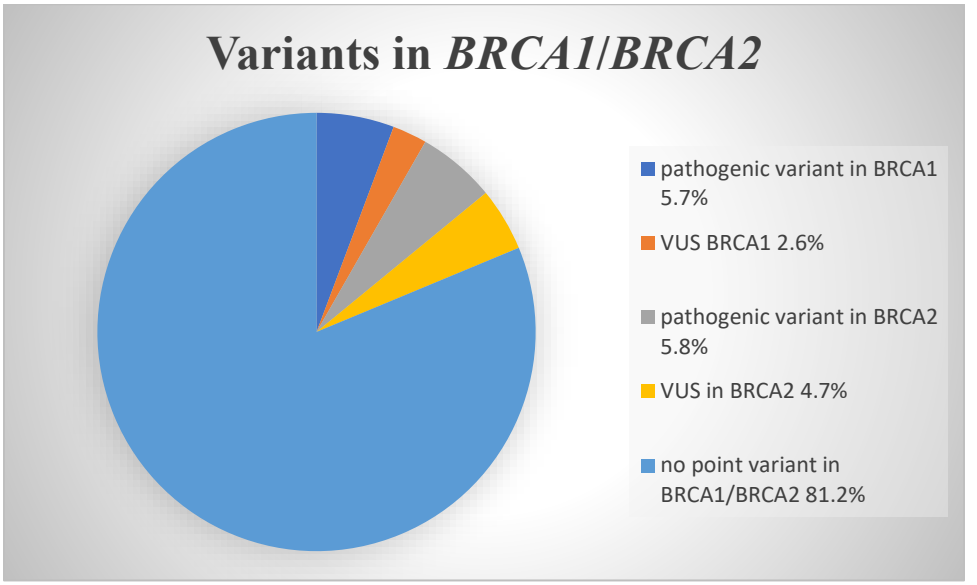


Fig.9 BRCA1/BRCA2 identified variants.

Ninety-six percent of the individuals screened for *BRCA1* gene and carrying a pathogenic mutation were affected with breast/ovarian cancer, accordingly to the expected genotype-phenotype correlations. Two percent of the individuals showed a tumor type "non corresponding" to the mutated gene: one individual was affected with melanoma and the other with squamous cell carcinoma. The remaining two percent of the sequenced individuals were carrying a *BRCA1* pathogenic mutation but were healthy at the time of the genetic test. Among patients carrying a *BRCA1* VUS one showed an unexpected association between apudoma (endocrine tumor) and gene mutation.

Among individuals with a *BRCA2* variant, one pathogenic mutation carrier was affected with melanoma while one individual with a VUS was affected with lymphoma. As expected, the remaining patients were affected with breast/ovarian cancer. Two percent of patients were healthy individuals at the time of the genetic test.

6.2.2 MMR genes sequencing

I analyzed 282 DNA samples sequenced at the Clinical Genomics Unit to search for point mutations following the request for *MMR* genetic test (*MLH1, MSH2, EPCAM, MSH6, PMS2* genes). The 30 identified pathogenic variants were 6 in *MLH1,* 12 in *MSH2,* 7 in *MSH6,* and 5 in *PMS2,* while the 27 identified VUS were 6 in *MLH1*, 7 in *MSH2,* 8 in *MSH6,* 6 in *PMS2* (Fig. 10).

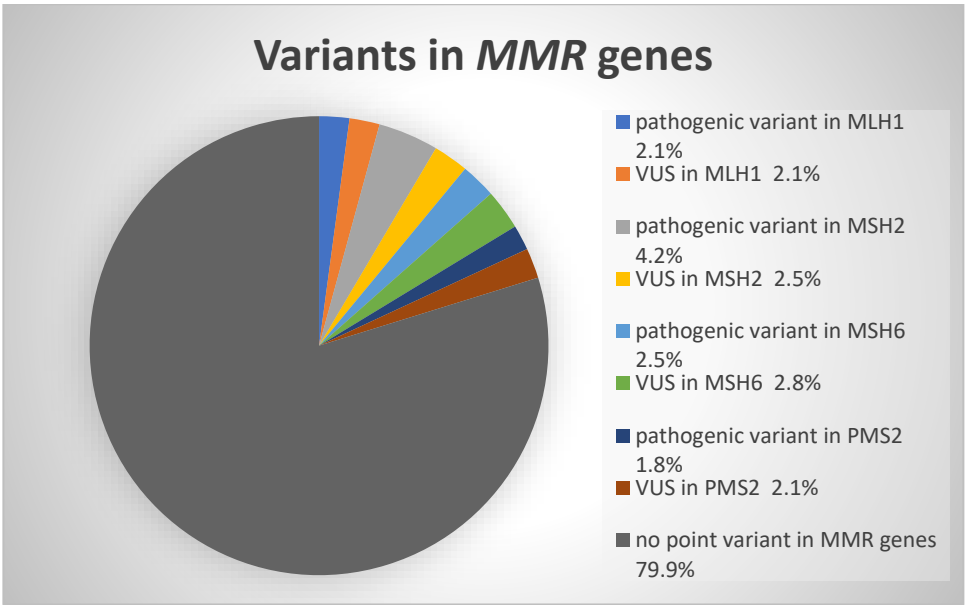


Fig.10 MMR genes variants identified by DNA sequencing.

All individuals screened for *MMR* genes and carrying a pathogenic variant were affected with a tumor of the Lynch syndrome tumor spectrum, including colorectal, intestinal, endometrial, and pancreatic cancer. Among individuals carrying *MMR* gene VUS, 6 were affected with breast cancer (2 with a mutation in *MLH1,* two in *MSH2,* two in *MSH6*). One individual with a VUS in *MLH1* was affected with a CNS tumor. The remaining patients were affected with a tumor of the *MMR* tumor spectrum (colorectal, endometrial, and intestinal cancer).

6.2.3 Multi gene panel testing

I analyzed a series of genes that are known to be associated with rare hereditary cancer syndromes (LF, FAP, MAP, HDGC and familial melanoma) following the physician' request for multi gene panel testing. Out of 779 individuals, 60 (7.7%) showed a pathogenic mutation in a cancer related gene (Fig. 11), 138 (17.7%) showed a VUS (Fig. 12) and 405 (74.6%) were wild-type and healthy when tested.

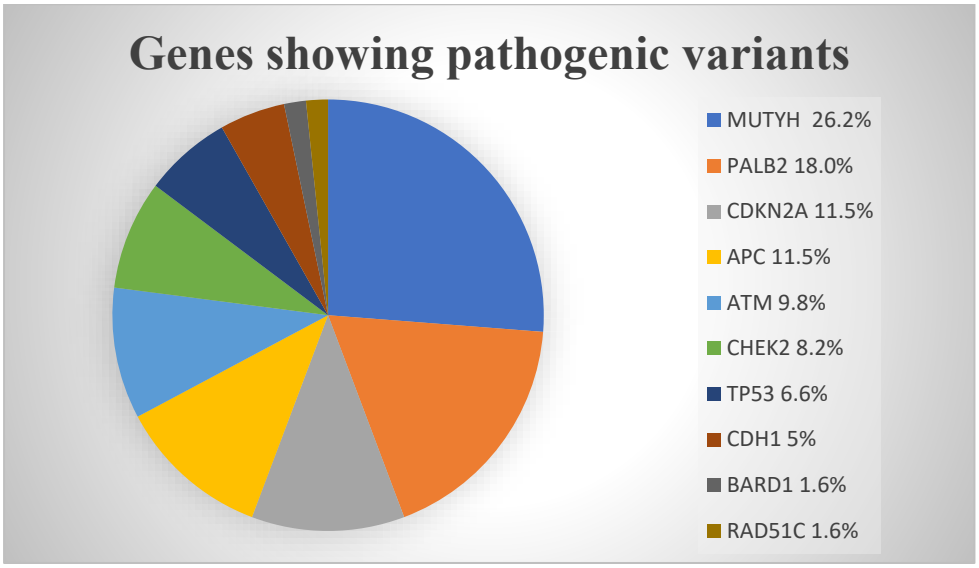


Fig.11 Pathogenic variants identified in different genes following multi gene panel sequencing.

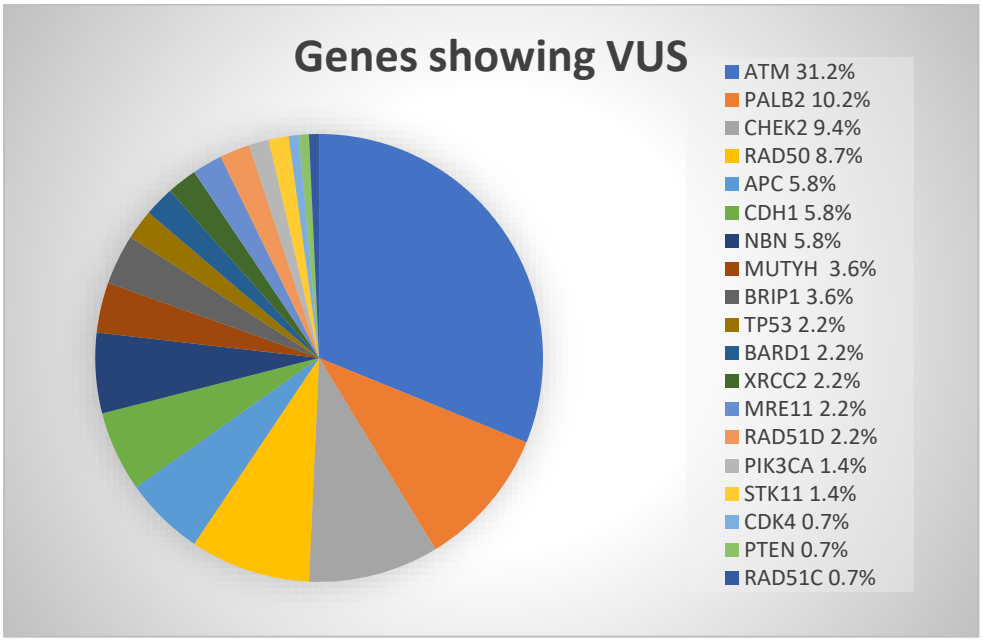


Fig.12 Variants of uncertain significance identified in different genes following multi gene panel sequencing.

Results

Among patients with pathogenic variants, 10 carried a *MUTYH* monoallelic variant: three were affected with breast cancer, one with breast and colon cancer, one with prostate cancer, one with FAP, one with endometrial cancer and three were healthy. In addition, three patients affected with colon cancer carried two different *MUTYH* pathogenic variants, consistently with compound heterozygosity.

Eleven patients carried variants in *PALB2* gene: 7 had breast cancer, one had ovarian cancer, one had breast cancer and endometrial cancer, one had breast and thyroid cancer and one was healthy.

Seven patients showed *CDKN2A* gene variants: one was affected with breast cancer, 5 had melanoma, and one was healthy.

Seven patients were *APC* variants carriers: two had *FAP*, two had colon cancer, one had rectum cancer and one had liver cancer.

Four individuals were *TP53* mutation carriers: one had a sarcoma, one had breast cancer, one had breast cancer and sarcoma, and one had breast and ovarian cancer.

Out of 3 *CDH1* mutation carriers, 2 were affected with breast and one with stomach cancer.

Six patients showed *ATM* gene variants and were affected with breast cancer. Six patients had a *CHEK2* pathogenic mutation: 5 were affected with breast cancer and one with endometrial cancer.

One subject had a *BARD1* genetic lesion and breast cancer; another patient had a variant in *RAD51C* gene and was affected with breast cancer.

Overall, in 90% of the analyzed patients the tumor type was corresponding to the mutated gene, accordingly to the expected genotype-phenotype correlations. Of relevance, in the remaining 10% of cases the gene-disease associations were unexpected: *MUTYH* with breast, prostate, and thyroid cancers; *PALB2* with ovarian cancer; *CDKN2A* with breast cancer.

6.2.4 Incidental findings

In clinical exome and genome germline sequencing, incidental findings (*ifs*) are variants of potential medical relevance that are unrelated to the medical indication for which the genetic test was originally required. The American College of Medical Genetics (ACMG) adopted the terminology of *ifs* in its initial recommendations of 2013 and defined them as "[…] results that are not related to the indication for ordering the sequencing but that may nonetheless be of medical value or utility […]" (Green *et al*., 2013). However, labeling findings as incidental in a context of genomic sequencing has been terminologically criticized as paradoxical, because discovering numerous variants is intrinsic to these techniques.

Results

In CGS laboratory, in 1431 patients sequenced by Sophia Genetics Custom Panel, we discovered 42 subjects (3%) carrying pathogenic mutations and 188 subjects (13.1%) carrying VUS in cancer predisposition genes the analysis of which had not specifically requested by the clinician. Apart from one patient screened for *BRCA1* and *BRCA2* genes, affected with breast cancer and carrying a VUS in *BRCA2* gene and a pathogenic *if* in *CDH1* gene, all the other individuals showing a pathogenic *if* were WT for the requested genes. In 17 subjects (40%) affected with breast cancer, we found pathogenic *ifs* in genes not related to the tumor type: 11 individuals carried a monoallelic pathogenic mutation in *MUTYH* gene, 3 in one of the *MMR* genes, 2 in *RAD50* gene and one in *MRE11* gene. In 18 breast cancer patients for whom the sequencing of *BRCA1/BRCA2* only had been ordered, pathogenic *ifs* were identified in other cancer predisposition genes related to breast cancer: 6 in *ATM*, 4 in *PALB2,* 3 in *CHEK2*, one in *CDH1*, one in *PTEN*, one in *TP53*, one in *ABRAXAS*, and one in *RAD51C* gene (Fig. 13)

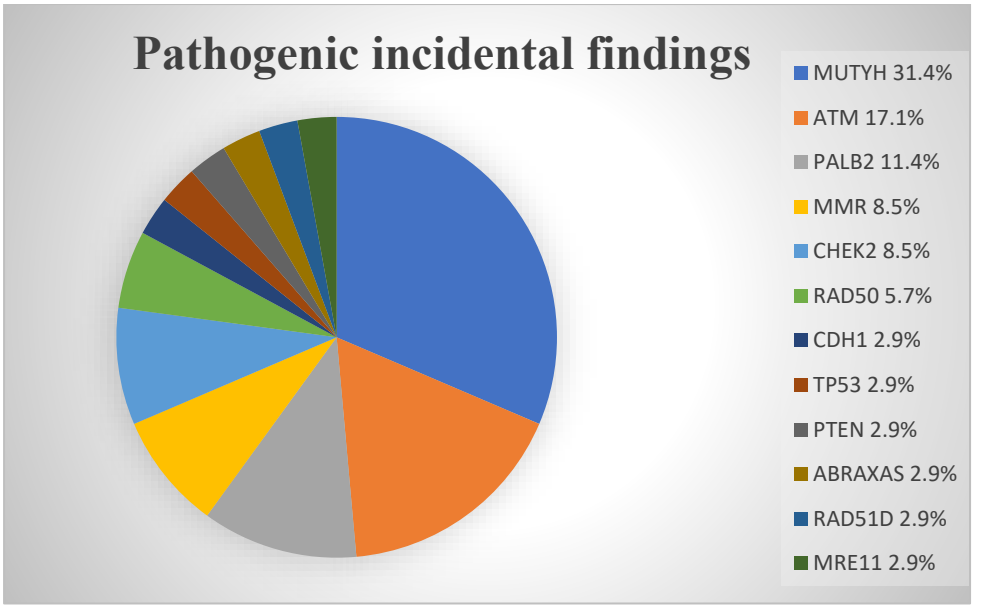


Fig.13 Pathogenic ifs identified in breast cancer patients tested for BRCA1 and BRCA2 genes.

One individual affected with colon cancer and analyzed for *MMR* genes was found to carry a pathogenic *if* in *CDKN2A* gene, while another patient affected with colon cancer and analyzed for *APC* and *MUTYH* genes had a pathogenic *if* in *PALB2* gene. One subject affected with endometrial cancer, analyzed for *BRCA1/BRCA2* and *MMR* genes, proved to carry a pathogenic variant in *APC* gene. One patient affected with melanoma and tested for *CDKN2A* gene was found to have a monoallelic pathogenic *if* in *MUTYH*. One healthy individual with strong family history of breast cancer and screened for *BRCA1* and *BRCA2* genes was found to have a pathogenic mutation in *PMS2*. One individual with colon cancer and loss of MSH6 expression in

tumor tissue had a pathogenic mutation in *MSH2* gene.
Overall, in 50% of cases in which pathogenic *ifs* were found we did not observe the expected association between tumor type and mutated gene.

6.3 VUS reclassification

Our laboratory reclassified some *BRCA1/BRCA2* VUS identified in patients who underwent genetic testing in the past and issued amended reports to physicians who, in turn, disclosed the results to their patients. VUS were re-evaluated using: the American College of Medical Genetics (ACMG) updated guidelines for interpretation of genetic results; clinical information; publically available databases.

Among 1841 samples sequenced by NGS for *BRCA1/BRCA2* genes, we identified 115 cases with VUS, including 40 with a *BRCA1* variant, 73 with a *BRCA2* variant, and 2 with variants in both genes (Fig. 12).

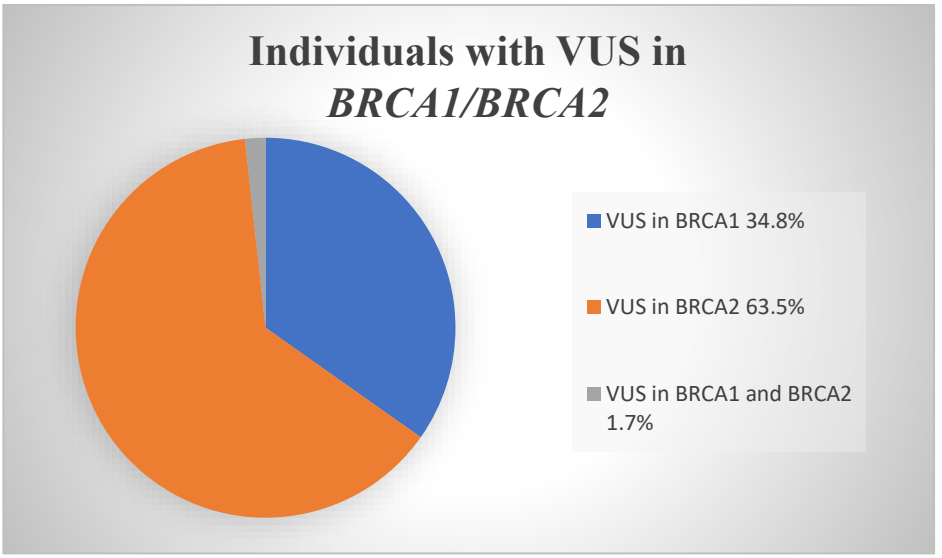


Fig.12 VUS carriers in BRCA1/BRCA2 genes.

Six *BRCA1* variants, previously classified as "Class 3" (*i.e.* of uncertain significance), have been reviewed as follows: c.3082C>T (p.Arg1028Cys), c.3708T>G (p.Asn1236Lys), c.3344_3346delAAG (p.Glu1115del), and c.5074+6C>G mutations have been reclassified as benign (Class 1); c.43A>C mutation as likely benign (Class 2); c.5017_5019delCAC (p.His1673del) mutation as likely pathogenic (Class 4).

Ten *BRCA2* variants, previously classified as "Class 3" (*i.e.* of uncertain significance), have been reviewed as follows: c.7994A>G (p.Asp2665Gly), and c.6322C>T (p.Arg2108Cys) mutations have been reclassified as benign (Class 1);

c.9649-6dupT, c.9104A>C (p.Tyr3035Ser), c.1022G>T (p.Cys341Phe), c.1247T>G (p.Ile416Ser), and c.5508T>G (p.Asn1836Lys) mutations as likely benign (Class 2); c.8375T>C (p.Leu2792Pro) mutation as likely pathogenic (Class 4); c.632-3C>G, and c.7975A>G mutations as pathogenic (Class 5).

Fig. 13 shows the reclassification of specific VUS identified in *BRCA1* and *BRCA2* genes: 6 variants (37.5%) have been reviewed as benign, 6 (37.5%) as likely benign, 2 (12.5%) as pathogenic and 2 (12.5%) as likely pathogenic. Overall, 13.4% (16/119) of the *BRCA1/2* analyzed variants initially classified as VUS over a 4-year period were reclassified. Among these, 75% were reclassified as benign/likely benign and 25% as pathogenic/likely pathogenic.

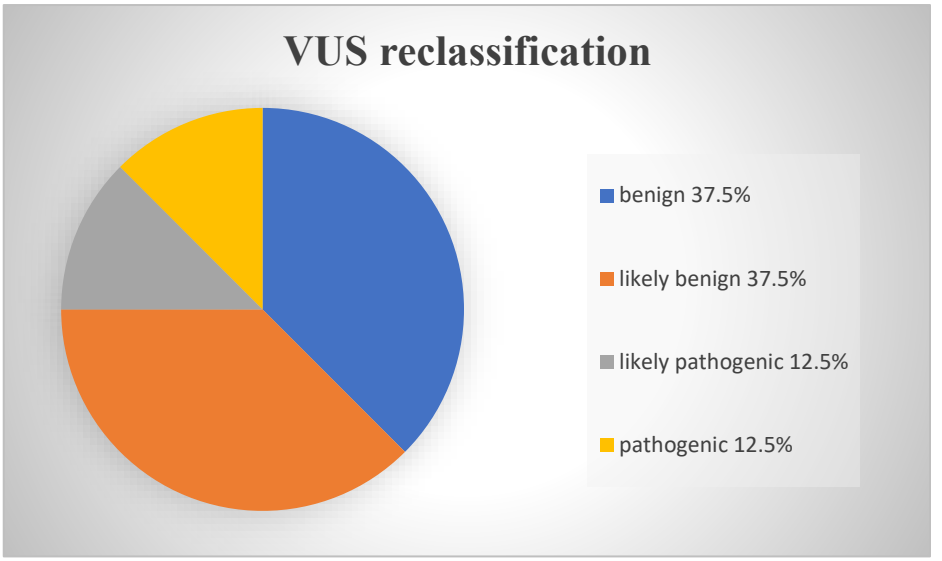


Fig.13 Reclassification of variants of uncertain significance.

6.4 PMS2 gene and PMS2CL pseudogene

Pseudogenes are complete or partial copies of genes, unable to code for functional polypeptides. These copies over time randomly accumulate mutations (insertions, deletions, and substitutions) that often cause disruptions of the original reading frame. Since pseudogenes are characterized by high sequence similarity with their corresponding genes, an ambivalent mapping in NGS data analysis cannot be avoided. Exons 11 through 15 of *PMS2* gene have highly homologous counterparts in the pseudogene *PMS2CL*. The differences in the reference genome between the corresponding regions of the gene and the pseudogene are dubbed paralogous sequence variants (PSVs). Sometimes PSVs can be exchanged between the gene and the pseudogene locations (gene conversion). Because of this, special attention is needed when assigning reads, variants, and CNVs to the gene or to the pseudogene.

In our survey, we identified 6 patients carrying a pathogenic mutation and 7 patients showing a VUS in *PMS2*. Nine of these variants (2 pathogenic and 7 VUS)

were in exons from 11 to 14, requiring deeper analysis to verify that *PMS2* rather than *PMS2CL* was affected. Of relevance, the presence of mutations in *PMS2* can influence the type of cure strategy, surgery, or surveillance. The 2 patients with a pathogenic variant were affected with colon cancer. Among individuals carrying a VUS 2 were affected with colon cancer, 2 with breast cancer, one with endometrial cancer, one with rectum cancer and one was healthy.

To investigate the critical variants, we developed a long-range PCR to amplify a sequence of about 10 kb by using primers targeting a specific region unique to *PMS2* (Fig. 14).

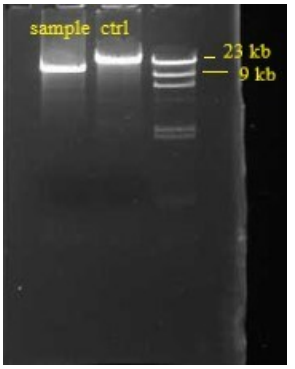


Fig.14 Long-range PCR: amplified region from exon 10 to exon 15 of PMS2

We sequenced the amplified fragment by Sanger method, obtaining the electropherogram corresponding to the region containing the variant we wanted to validate (Fig. 15). By this type of approach, we were able to validate one pathogenic variant (as explained in legends of Fig. 16 and Fig. 17) and 3 VUSs. The patient carrying the pathogenic variant in *PMS2* was affected with colon cancer, showing the expected genotype-phenotype correlation. One individual with a VUS was affected by breast cancer, one with rectum cancer and another with colon cancer.

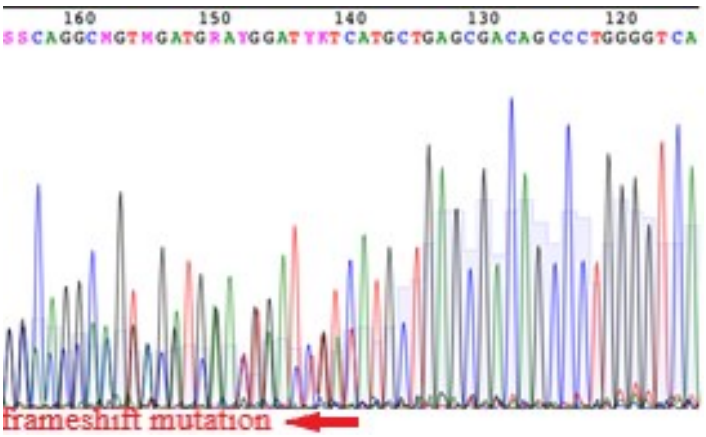


Fig.15 Sanger sequence of DNA from the patient carrying a pathogenic heterozygous frameshift mutation in exon 14 of PMS2 gene. Exon 14 has highly homologous counterparts in PMS2CL pseudogene.

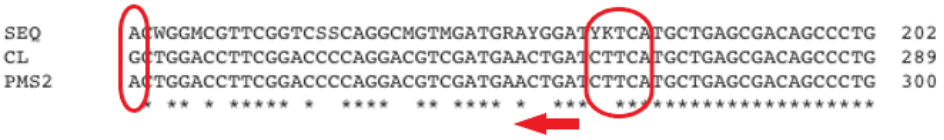


Fig.16 Exon 14 sequences: the first one belongs to the patient carrying a frameshift mutation in PMS2 gene; the second one is the sequence of PMS2CL pseudogene; the third one is the sequence of PMS2 gene. On the left, the red circle indicates that, in that genomic position, the patient shows an adenine as in the sequence of PMS2 gene; instead, PMS2CL shows a guanine in that same genomic position. The red circle on the right indicates the start point of the frameshift.

Fig.17 Intron 12 sequences: the first one belongs to the patient carrying the frameshift mutation in PMS2 gene; the second one is the sequence of PMS2CL pseudogene; the third one is the sequence of PMS2 gene. The red circles indicate four genomic positions in which PMS2CL pseudogene differs in one base from PMS2 gene and the sequence of the individual under examination, indicating that the pathogenic mutation affects PMS2 gene.

6.5 CHEK2 gene and pseudogenes

CHEK2 pseudogenes overlap with exons 10 through 14 and share 95-98% sequence homology with *CHEK2* (Chromosome 22). These pseudogenes are located on chromosome 15 (*CHEK2 p2*), chromosome 2 (*CHEK2 p3*), chromosome 22 (*CHEK2 p4*), chromosome 10 (*CHEK2 p5*). In our survey, we identified 3 patients with a pathogenic mutation and 3 patients with a VUS located in exons 11 and 12 of *CHEK2* gene. These variants required a deeper analysis to verify if they affected *CHEK2* gene or its pseudogenes. Of relevance, the presence of mutations in *CHEK2* can influence the type of cure strategy or surveillance. The 3 patients with a pathogenic variant were affected with breast cancer. Also, the individuals carrying a VUS were affected with breast cancer.

We developed a long-range PCR to amplify a sequence of about 1600 kb by using primers specifically targeting a region unique to *CHEK2* gene. We sequenced the amplified fragment by Sanger method, and we obtained an electropherogram corresponding to the region containing the variant we wanted to validate. By this type of approach, we were able to validate three pathogenic variants (see one example in Fig. 18) and three VUSs.

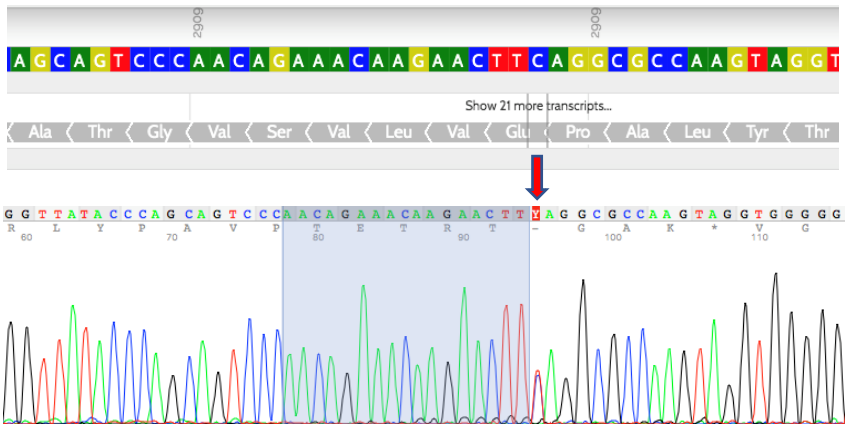


Fig. 18 Sanger sequence of DNA from the patient carrying a pathogenic heterozygous missense mutation in exon 11 of CHEK2 gene. Exon 11 has highly homologous counterparts in CHEK2 pseudogenes

Fig.19 CHEK2 exon 11 sequences: the first one is the sequence of CHEK2p2 pseudogene; the second is the sequence of CHEK2 gene; the third is the sequence belonging to the patient carrying a heterozygous missense mutation in CHEK2 gene. The red circles indicate that, in those genomic positions, the patient shows the same bases of CHEK2 gene; instead, in the same genomic positions, CHEK2p2 pseudogene has different bases. The green circle indicates the heterozygous variant located on the sequence of the patient.

6.6 Gross duplication in CHEK2 gene

A patient affected by breast cancer (expected to carry a *BRCA1* or a *BRCA2* mutation) proved to carry a gross duplication of exon 2 in *CHEK2* gene (Fig. 20). This rearrangement was not reported in the literature and was of unknown significance.

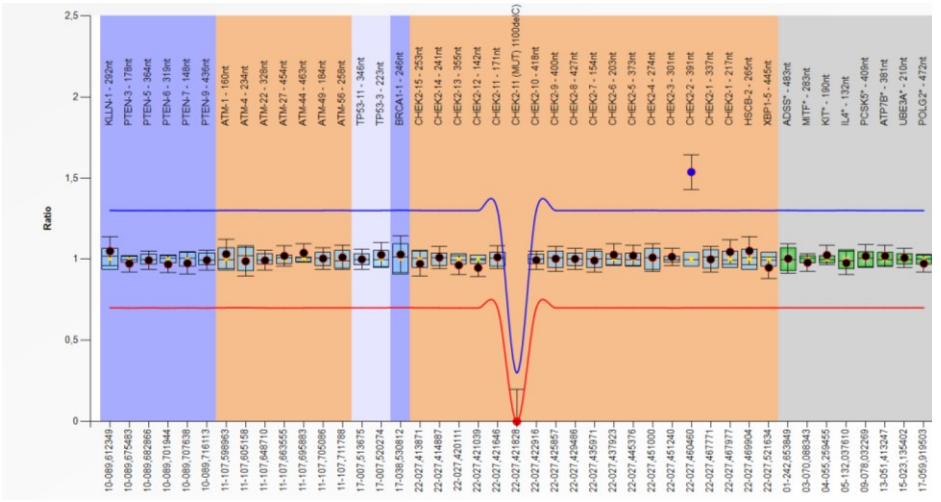


Fig.20 MLPA assay performed by a kit containing probes for CHEK2 exons. The probe in blue corresponds to exon 2 that is duplicated in this subject. The peak corresponds to the c.1100delC mutation detected by a CHEK2 specific probe; this mutation is absent in the subject here analyzed.

We perform a PCR on patient's DNA and on a DNA control sample by using a couple of primers specific to *CHEK2* exon 2, spaced each other by 15pb only. In the control subject we did not observe any amplification product, while in the patient DNA we observed a 1.5 kb DNA fragment (Fig. 21). Indeed, the mutated allele (containing the duplication) gave rise to a fragment, while the WT allele could not be amplified with primers so close each other.

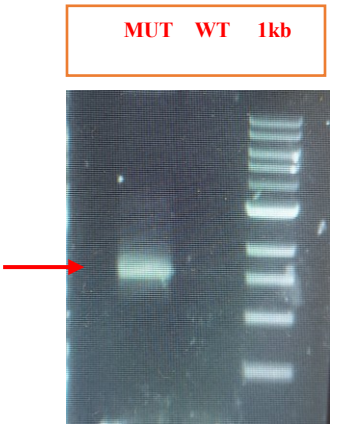


Fig.21 CHEK2 PCR performed on mutated and WT DNA samples. A 1.5 kb fragment (red arrow) was only generated in the DNA sample of patient carrying a exon 2 duplication.

The presence of the 1.5 kb fragment indicated that the duplication of *CHEK2* exon 2 was *in tandem*, whereby another copy of exon 2 was inserted serially to the original segment (Fig. 22). We sequenced the fragment by Sanger method and compared the obtained sequence with that of *CHEK2* intron 1 and 2 by *Clustal Omega* bioinformatic tool. We observed the presence of *CHEK2* intron 2 (after exon 2), then a 24pb omology region between intron 2 and intron 1, and again the sequence of intron (Fig. 23, Fig. 24).

Fig.22 The mutant allele carrying two copies of CHEK2 exon 2 in tandem

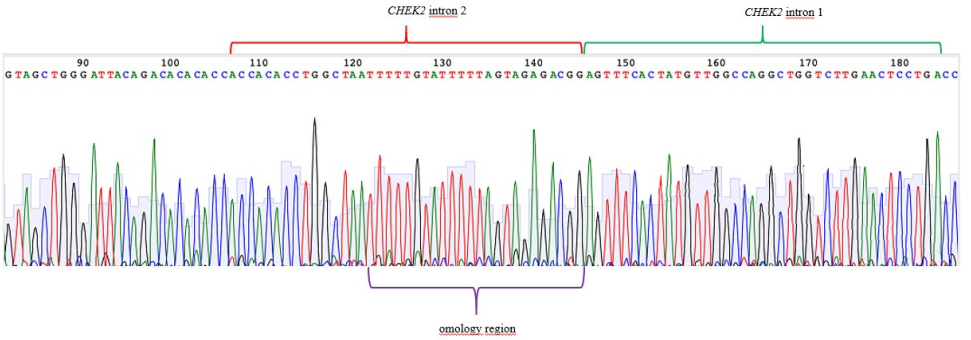


Fig.23 Sanger sequencing of 1.5kb DNA fragment of the patient carrying a CNV in CHEK2 gene.

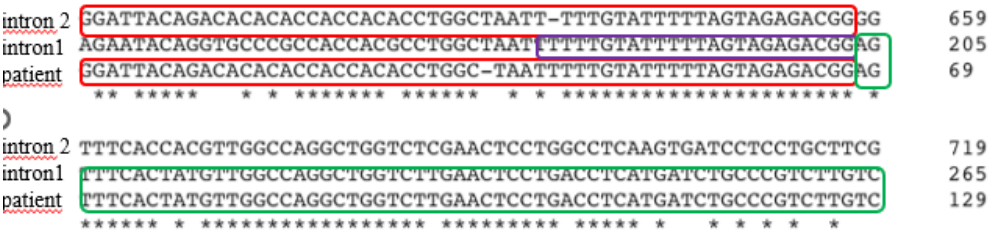


Fig.24 Clustal Omega alignment among CHEK2 intron 2, intron 1 and the patient 1.5 kb fragment sequences. The red outline indicates the presence of intron 2 after exon 2 in the patient DNA, then a 24pb omology region (violet outline) between intron 2 and intron 1, and again the sequence of intron 1 (green outline).

Using *Expasy Translate* bioinformatic tool, we translated the DNA sequence of the patient into the protein sequence. We found a premature stop codon (TGA), absent in WT *CHEK2* exon 2, generating a non-functional truncated protein (Fig. 25). Accordingly, we could classify the duplication of exon 2 as pathogenic.

Fig.25 Protein sequence of the duplicated CHEK2 exon 2 (on the left) compared to the WT sequence of CHEK2 exon 2 (on the right)

*6.7 The "**PERSONA PROJECT**": analysis of triple negative breast cancers*

The prevalence of *BRCA1* and *BRCA2* mutations in breast cancer population is 15% and 8%, respectively. Various studies demonstrated that 15%-20% of women with "triple-negative" breast cancer (TNBC) carry *BRCA1* or *BRCA2* germline mutations, and that 70% of breast cancers that develop in *BRCA1* mutation carriers are "triple-negative" (TNBCs) (Stevens *et al*., 2013). TNBCs are negative for the expression of ER (*estrogen receptors*) / PgR (*progesterone receptor*) / HER2 (*human epidermal growth factor receptor 2*) proteins and differ from other types of invasive breast cancer in that they grow and spread faster, have limited treatment options (they do not benefit from endocrine or anti-HER2 therapy) and a worse prognosis (Lips *et al*., 2015). *BRCA1/2* mutation prevalence is significantly higher in TNBC patients of younger age. Accordingly, patients of young age with TNBCs are predicted to carry pathogenic variants; as such, several guidelines point to these patients as targets for screening tests (the National Comprehensive Cancer Network guidelines recommend genetic testing of all TNBC patients aged ≤60 years, regardless of family history).

The vast majority of breast cancer genetic screening focuses on *BRCA1-2* alterations, due to their high prevalence. However, *BRCA1* and *BRCA2* account for approximately one third of germline mutations in breast cancer, with contributions of other genes (Easton *et al*., 2015). Among these, there are genes of the Fanconi anemia pathway (*e.g. RAD51D, NBN, ATM*) that, as *BRCA1/2*, are involved in homologous recombination (HR) mediated DNA repair (Domagala *et al.,* 2015). Inherited and acquired defects in HR DNA repair (the so-called "BRCAness" phenotype) represent a very important biomarker for therapy response since HR-deficiency confers sensitivity to DNA damaging therapy.

At the Clinical Genomics Unit, we are interested in evaluating the prevalence of mutations in breast cancer risk-associated genes and the feasibility of genetic testing for clinical decision making in TNBC cases; more specifically, we are interested in

assessing the prognostic and therapeutic predictive role of HR defects *vs.* the other genetic alterations. Accordingly, I searched for pathogenic germline variants and VUSs in patients selected on the basis of the following criteria: age between 18 and 60 years, histologically confirmed TNBC (ER < 5%, PgR < 5%, HER2 negative), I-III tumor stage, suitable to undergo surgery. To address the primary aim of the study (*i.e.* to evaluate the frequency of pathogenic variants in breast cancer risk-associated genes in TNBC cases) with sufficiently precise estimate, and considering as the measure of precision the 95% confidence interval of the estimated proportion, a sample size of 318 patients with informative tests is needed. The sample size calculation is based on the hypothesis that the true prevalence of mutations is approximately 20%. Therefore, I sequenced and analyzed 318 samples by using the Trusight Sequencing Cancer Panel (94 cancer-related genes and 286 SNPs). Specifically, I searched for variants in 14 actionable genes related to health conditions with known medical cure and/or risk cancer prevention recommendations. These genes were: *BRCA1, BRCA2, APC, MUTYH, MLH1, MSH2, MSH6, EPCAM, PMS2, CDH1, CDKN2A, CDK4, TP53, PTEN*. I also annotated pathogenic variants and VUSs identified in the other genes of the panel. Two-hundred-twenty individuals (69.2%) proved to be WT in the 14 actionable genes, did not carry any pathogenic variant but had VUS in other genes of the panel. Only 8 individuals (2.5%) were WT in all genes of the panel. Ninety-hundred patients (28.3%) had pathogenic variants in one or more genes of the panel. Sixty-four patients (20.1%) showed pathogenic variants affecting one of the 14 actionable genes, 11 (3.5%) had a pathogenic variant and a VUS affecting one or more of the 14 genes, and 45 (14.2%) had at least one VUS affecting one or more of the 14 genes (Fig. 26, Fig. 27).

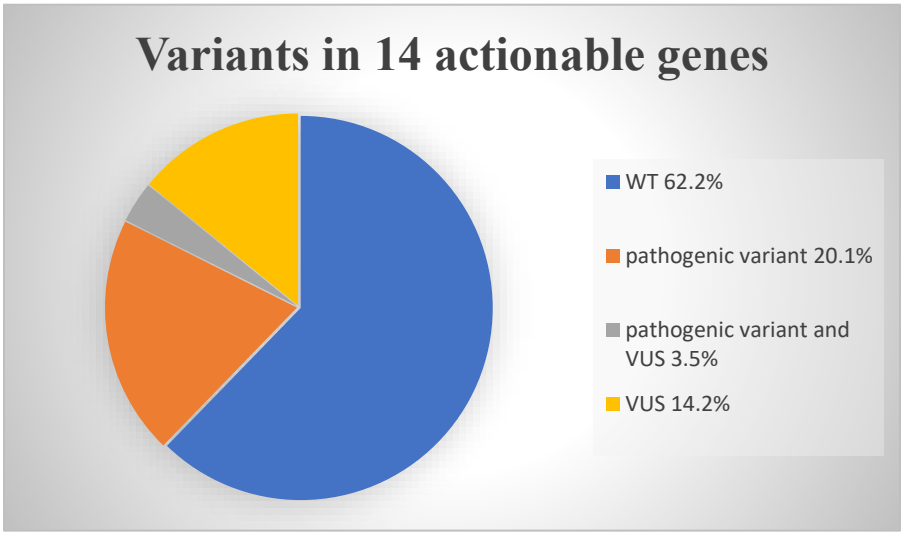


Fig. 26. Germline variants in the 14 actionable genes tested on TNBC patients.

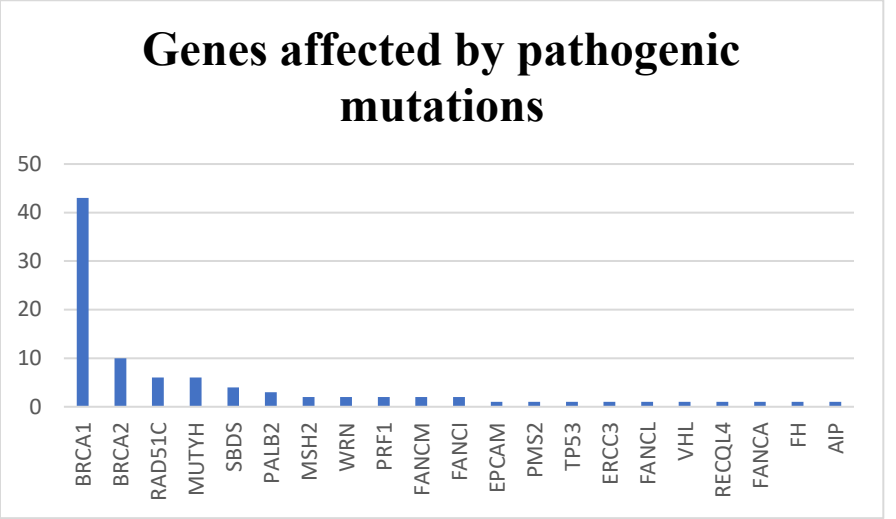


Fig. 27 Genes affected by pathogenic variants: BRCA1 was the gene affected by the highest number of pathogenic variants, followed by BRCA2, RAD51C, MUTYH, SBDS, PALB2, MSH2, and WRN. Forty-three individuals (42.6%) had a pathogenic mutation in BRCA1 and 10 individuals (9.9%) in BRCA2.

As shown in Fig. 27, pathogenic variants were present not only in breast cancer related genes (*BRCA1, BRCA2, PALB2, ERCC3, FANCM, FANCA, FH, AIP)*. The panel allowed detection of pathogenic variants also in genes linked to other cancer types: six individuals (6%) showed a pathogenic variant in *RAD51C*, associated with ovarian cancer; 6 patients (6%) had a monoallelic pathogenic mutation in *MUTYH,* associated with *MAP*; 4 individuals (4%) had pathogenic mutations in *MMR* genes linked to Lynch Syndrome; 4 (4%) individuals showed a pathogenic variant in *SBDS* gene linked to Shwachman-Bodian-Diamond syndrome, 2 individuals (3%) carried a mutation in *WRN*, linked to colorectal cancer. The remaining patients showed pathogenic variants in hemangioblastomas, osteosarcomas, hematological and renal cancer predisposition genes. Among pathogenic mutations identified, 38% affected genes not linked to breast cancer.

Among variants in *HBOC* predisposition genes, c.5266dupC represented the 11.6% of all *BRCA1* pathogenic lesions; c.1088delA and the c.798_799delTT represented the 9.3% and the 7% of all *BRCA1* pathogenic variants, respectively. The c.5266dupC variant entered the Ashkenazi Jewish population through affecting the Polish population about 400 years ago and it is now the most prevalent *BRCA1* breast cancer related mutation in European countries (Karami *et al.*, 2013). The c.1088delA is a common Italian variant linked to breast cancer predisposition (Figlioli *et al.* 2021), while the c.798_799delTT mutation has been reported in Moroccan, Algerian and Tunisian breast cancer families as one of the founder mutations in Northern Africa (Zoure *et al*. 2018).

I also analyzed the TNBC survey for two disease-associated SNPs (daSNPs): rs721048 and rs2046210. The first daSNP is an intronic variant (c.1290+30064G>A)

in *EHBP1* gene: the risk allele (A) is associated with prostate cancer (Ao *et al.,* 2015). Up to 35% of patients of "*Persona project*" affected with TNBC proved to carry this risk variant. The second daSNP (G>A) is in proximity to the C6orf97 and estrogen receptor 1-ESR1 genes. The risk allele (A) is recurring in breast cancer patients with anti-estrogen therapy resistance and poor prognosis. As reported by literature data (Jin et al. 2019), this polymorphism may serve as a potential genetic biomarker of BC in both Asians and Caucasians. Of relevance, 52% percent of the individuals of "*Persona project*" proved to carry this disease-associated variant.

7. Discussion

Hereditary cancer syndromes are characterized by Mendelian patterns of inheritance and caused by pathogenic germline variants that are transmitted in families. These syndromes are clinically suspected when there are multiple individuals within a family with the same cancer type or related cancers, and/or when cancer is developed at an early age, and/or when there are multiple primary cancers in one individual (Nagy *et al*., 2004). The identification of a germline pathogenic variant can confirm the genetic predisposition, leading to personalized risk assessment, clinical management and decision-making (*e.g.* cancer treatment, high-risk surveillance, surgery), and can enable cascade testing to identify at-risk family members. For some cancer syndromes (*e.g.* HBOC and Lynch syndromes), heightened surveillance or preventive strategies (prophylactic surgery) can enable early detection, and reduce morbidity and mortality (Li *et al*., 2016).

In this frame, I retrospectively analyzed 2641 human DNA samples to identify germline pathogenic variants in cancer predisposition genes including small sequence changes and gross genomic rearrangements.

Until recently, the role of structural mutations in cancer syndromes has not adequately considered, probably because genomic deletions or duplications are not readily detected by PCR-based sequencing. New multiplex PCR methods, especially multiplex ligation-dependent probe amplification (MLPA), have allowed targeted copy number assessment of single genes or exons (Schouten *et al*., 2002). This has recently led to an upsurge in discoveries of patients and families with rare pathogenic CNVs that strongly predispose to cancer. Indeed, 4-28% of inherited *BRCA* mutations have been estimated to be due to large genomic rearrangements, which are missed by using sequencing alone (Kwong *et al*., 2015).

Aim of this work was to evaluate the pick-up rate of extended deletions and duplications in our cohort. Among individuals with genetic alterations, we identified 47 subjects carrying a CNV. These individuals represented the 7% of all sequenced cases, a value which is in line with literature data. *BRCA1* was the gene showing the highest number of CNVs (26/47 cases), while a series of other genes, including *BRCA2*, were more rarely involved. The difference in CNVs between *BRCA1* (55%) and *BRCA2* (4.3%) can be explained by the higher number of short interspersed nuclear elements (SINE), specifically *Alu* repeated sequences, along *BRCA1*; indeed, these elements are known to promote genomic rearrangements (Wang *et al*., 2019; Smith *et al*., 1996). Similarly, the other genes (*ATM*, *BRIP1*, *MMR* genes) where we found rare CNVs are characterized by a limited number of *Alu* elements. All gross deletions identified in our patients were classified as “Class 5 - pathogenic variants” since leading to nonfunctional or truncated/unstable protein products. On the contrary, most gross duplications we identified were classified as VUS. In principle, additional copies of genes could provide redundancy, allowing some copies to evolve new/modified functions or expression patterns, while other copies maintain the

original function. Regarding duplications, microarray technology could be useful to investigate and evaluate the mRNA expression profile in samples carrying more copies of genes, including *BRCA1* and *RAD51D*, the effect of which is still unknown. In some of our cases (duplications of: *BRCA1* exon 13*; PALB2* exons 11 and 13; *CHEK2* exon 2), the variants we found were classified as pathogenic as they were *tandem* duplications in which two consecutive copies of one or more exons were present in the genome. In these cases, the additional copy created a premature stop codon along the mRNA sequence thus generating a nonfunctional protein. Seventy-five percent of pathogenic CNVs carriers with a strong family history of cancer had developed a cancer, whereas 25% were asymptomatic when tested. This fraction of healthy individuals was represented by subjects with pathogenic mutations in high penetrance (*BRCA1* and *MMR* genes) or in moderate penetrance (*ATM)* genes. For germline pathogenic mutations in *BRCA1* there is an estimated likelihood of 55% to 72% for developing breast cancer in the lifetime (Petrucelli *et al*., 2022), while for *ATM* the breast cancer risk is likely greater than 25% (Jerzak *et al*.,2018). For *MMR* genes the estimated cumulative risk of colorectal cancer is up to 41% (Bonadona *et al*., 2011). Therefore, the mutated healthy individuals of our cohort have a higher cancer risk compared to the general population and they probably will develop a tumor in their lifetime. Regarding CNV carriers, the average ages at diagnosis were 47 and 45 for subjects with pathogenic CNVs and for subjects with VUS, respectively. In 84% of the patients with pathogenic CNVs the tumor type corresponded to the mutated gene accordingly to the expected genotype-phenotype correlations. In the remaining subjects with pathogenic variants, we observed the following associations: *BRCA1* and prostate cancer, *MSH2* and breast cancer, *PALB2* and colon cancer, *CHEK2* and colon cancer. Regarding *BRCA* and prostate cancer, a recent study reported that the frequency of *BRCA2* mutations is higher than that of *BRCA1* among patients with prostate cancer and that *BRCA2* (but not *BRCA1*) mutations are associated with higher prostate cancer mortality (Oh *et al*., 2019). Some specific mutations in *CHEK2* gene seem to be related to colorectal cancer but more studies are needed to confirm this hypothesis (Cybulski *et al*., 2007; Liu *et al*., 2012; Xiang *et* al 2011). The prognostic significance of pathogenic mutations of *PALB2* in colorectal cancer remains unclear (Pan *et al*., 2020). An increased risk of breast cancer in *MSH2* mutation carriers emerged from the analysis of a Canadian familial cancer registry (Goldberg *et al*., 2017). At any rate, even if these new associations between cancer type and mutated gene are not reported in the NCCN Clinical Practice Guidelines in Oncology, our patients should follow intensive cancer surveillance programs that are specific for the mutated gene.

A far as small sequence changes are concerned, I retrospectively analyzed 2077 DNA samples to search for point mutations, following the request for *BRCA1* and *BRCA2* genetic test. One-hundred-nineteen individuals (5.7%) proved to carry a pathogenetic variant in *BRCA1* and 120 (5.8%) had a pathogenic variant in *BRCA2*. Fifty-three (2.6%) patients carried a VUS in *BRCA1* and 97 (4.7%) had a VUS in *BRCA2*. These results are consistent with previous reports (Ayed-Guerfali *et al.,* 2021*;* Caglayan *et al.,* 2019). More than one thousand six hundred (81.2%) cases of our cohort were represented by individuals in which no point mutations in *BRCA1* or *BRCA2* genes were identified. These individuals were strictly selected for their strong

family history of cancer and most of them (96%) were affected with breast/ovarian cancer. We can explain this finding by assuming that these patients have pathogenic mutations in cancer predisposition genes other than *BRCA1/BRCA2*; accordingly, these subjects should be tested by adequate exome sequencing. Alternatively, these patients may be carriers of pathogenic mutations in still unknown predisposition genes or in deep intronic regions of known predisposition genes. Deleterious DNA variants located more than 100 base pairs away from exon-intron junctions could lead to pseudo-exon inclusion due to activation of non-canonical splice sites or changes in splicing regulatory elements; moreover, deep intronic mutations could disrupt transcription regulatory motifs and non-coding RNA genes (Vaz-Drago *et al*., 2017). Finally, the above patients may be carriers of genes silenced by epigenetic mechanisms, such as promoter methylation. Epimutations *(i.e.* the abnormal transcription repression of active genes or the abnormal activation of usually repressed genes) are common events arising in somatic cells and have been directly associated with carcinogenic process. Less frequently, epimutations can arise in the germline and rare constitutive epimutations (*MLH1* silencing by promoter methylation) have been associated with Lynch syndrome (Gazzoli *et al*., 2002; Hitchins *et al*., 2007; Pineda *et al*., 2012; Zyla *et al*., 2021). Of relevance, NGS-based approaches are inadequate to detect the methylation status of promoter regions in DNA samples, while sodium-bisulfite-dependent, quantitative, fluorescence-based real-time PCR methods can be used to assess DNA methylation status.

As far as genotype-phenotype correlations are concerned, 2% of individuals carrying a pathogenic mutation in *BRCA1* showed a tumor type "non corresponding" to the mutated gene: one individual was affected with melanoma and the other with squamous cell carcinoma. Familial melanoma is commonly associated with variants in *CDKN2A/CDK4* genes and the association between melanoma and *BRCA1* gene is still controversial (Adams *et* al., 2020). On the other hand, NCCN guidelines do not mention any association between *BRCA1* and squamous cell carcinoma. Therefore, the causative role of *BRCA1* remains unclear and these two cases may be "phenocopies" due to "sporadic" tumors. Among subjects carrying pathogenic mutations in *BRCA2* gene we identified one patient affected with melanoma. Again, the involvement of *BRCA*2 in melanoma predisposition is matter of debate (Adams *et* al., 2020). Overall, only a constantly reexamination of gene-disease associations will contribute to understand the causative role of new germline mutations and to better counsel patients about their care.

I also evaluated the status of *MMR* genes in 282 patients: 80% of these subjects (affected with a tumor of the Lynch syndrome spectrum) did not carry any point mutation even if they were selected according to strict criteria of family history of cancer. These subjects are now eligible for clinical exome sequencing to search for new pathogenic mutations and for methylation profiling to search for *MMR* genes constitutive epimutations. All the individuals screened for *MMR* genes and carrying a pathogenic mutation were affected with a tumor of the known Lynch syndrome tumor spectrum, including colorectal, intestinal, endometrial, and pancreatic cancer. Among individuals carrying VUS in *MMR* genes, 6 were affected with breast cancer (2 with a variant in *MLH1,* 2 in *MSH2,* and 2 in *MSH6*). Interestingly, a recent study on the

association between *MMR* mutation type and breast cancer evaluated women with *MMR* mutations for their lifetime risk of breast cancer and determined their eligibility for breast-specific genetic testing. The authors' conclusion was that "many Lynch syndrome patients qualify for breast-specific genetic testing, and 7.5% of patients without breast cancer in this study qualified for enhanced surveillance for breast cancer" (Sheehan *et al*., 2020).

NGS technology is now commonly used for multigene panel germline testing. This methodology can sequence multiple genes simultaneously, resulting in dramatic reductions in the time and cost of genetic testing compared to traditional di-deoxy-capillary based sequencing or Sanger sequencing. Multigene panel testing is more efficient than single-gene testing when more than one gene may explain a patient's clinical manifestations or family history. In our lab, I sequenced 779 individuals following the request for multi gene panel testing. Out of 779 individuals, 60 (7.7%) showed a pathogenic mutation in a cancer related gene, and 138 (17.7%) showed a VUS. The analysis allowed us to identify new associations between the disease affecting the patient and the mutated gene. In 90% of the analyzed patients the tumor type was corresponding to the mutated gene, accordingly to the expected genotype-phenotype correlations. Of relevance, in the remaining 10% of cases the gene-disease associations were unexpected: *MUTYH* (monoallelic) with breast and prostate, and thyroid cancers; *PALB2* with ovarian cancer; *CDKN2A* with breast cancer. Monoallelic *MUTYH* mutations, inherited from only one parent, occur in 1-2% of the Caucasian population and are associated with a moderately increased risk of CRC (Win *et al*., 2016). Carriers of monoallelic mutation have on average an approximately 2.5-fold increased risk of CRC compared with the general population (Win *et al*., 2016). It is estimated that there are an elevated risk of liver and gastric cancers and a slightly increased risk of breast cancer for carriers with monoallelic *MUTYH* mutation (Zhu *et al*., 2011; Rennert *et al*., 2012). The associations between *MUTYH* and breast cancer, *MUTYH* and prostate cancer, *MUTYH* and thyroid cancer identified in our samples are not reported in NCCN Clinical Practice Guidelines in Oncology. NCCN guidelines report that *PALB2* pathogenic mutation carriers have a very low risk of developing ovarian cancer (3-5%) and there is no evidence for *CDKN2A* pathogenic mutation carriers to develop a breast cancer.

As already mentioned, the use of a multi gene panel allowed us to also identify many variants in cancer predisposing genes classified as VUS. As such, they cannot be used for clinical purposes either to identify individuals at risk or to drive treatment. VUS represent a particularly challenging topic in molecular diagnostics for hereditary cancers: they are variants for which evidence of pathogenicity is limited or contradictory, precluding a benign/likely-benign or pathogenic/likely-pathogenic classification. As gene panel size increases and exome/genome sequencing are increasingly used in molecular diagnostics, more VUS carriers will be identified. The yield of VUS partially depends on the panel size, and up to 40% of patients undergoing multigene panel testing for hereditary cancers have been found to have VUS (LaDuca *et al*., 2020). Variant classifications may change over time as evidence accumulates and laboratories reported that 3%-18% of variants have been reclassified over time (Turner *et* al., 2018). It has been observed that most reclassified variants are

downgraded to a more benign classification (Macklin *et al*., 2018); in particular, most reclassified VUS are downgraded to likely-benign or benign (Mersch *et al*., 2018; So *et al*., 2019). Strategies to generate evidence to reclassify VUS include family segregation studies (Tsai *et* al., 2019), functional studies to assess the biological impact of variants (Glazer *et al*., 2019) and data sharing (Lebo *et* al., 2018). Of relevance, variant reclassifications can affect patients' medical care (Chiang *et al*., 2021). Therefore, it is important that diagnostic laboratories periodically reassess variants to ensure patients receive up-to-date variant classifications to guide medical management. In our lab, 13.4% of *BRCA1* and *BRCA2* variants initially classified as VUS, were reclassified over a 4-year period. Among these, 75% were reclassified as benign/likely-benign and 25% as pathogenic/likely-pathogenic.

The use of multi gene panels sequencing can unexpectedly produce the so-called incidental findings (*ifs*), detecting the presence of pathogenic variants or VUS not consistent with the family history of patients or the presence of variants in cancer predisposition genes linked to the patient's disease but not requested by the clinician. In other words, the identified variants are outside of the original purpose for which the genetic test or procedure was conducted. The current debate among geneticists and clinicians focuses on the issue to disclose or not *ifs* to patients and if so, which *if* should be disclosed ("actionable mutations", variants of unknown significance or both). Regarding this debate, our lab decided to communicate to the clinicians the presence of pathogenic *ifs* and to reanalyze *ifs* of unknown significance over time: indeed, we considered *ifs* as a positive opportunity for adding information to improve the patient's care and health.

Out of 1431 patients sequenced by Sophia Genetics Custom Panel following the request of the clinician, I found 42 subjects (3%) carrying pathogenic mutations and 188 (13.1%) carrying VUS in cancer predisposition genes the analysis of which had not been requested. In 50% of cases in which pathogenic *ifs* were discovered there was no association between tumor type and mutated gene. In this scenario, it is very important for the patients to be informed about *ifs* identified since additional specific prevention screening procedures can be addressed to the patients to properly manage their condition. For example, the patient we found to carry a gross duplication of exon 2 in *CHEK2* was affected with breast cancer and was tested following the request for *BRCA1/BRCA2* gene sequencing. The patient proved to carry wt sequences of both *BRCA1* and *BRCA2* genes, while the duplication of *CHEK2* exon 2 was an *if*. This rearrangement was not reported in the literature and was initially classified as VUS from our lab. Following deep molecular analysis, we could demonstrate that this CNV was a *tandem* duplication creating a premature stop codon along the DNA sequence. Since the truncating mutation produced a nonfunctional protein, we could classify the duplication as pathogenic. Accordingly, we informed the clinician about the reclassification of the variant and the patient was addressed to a preventive screening program for colon cancer. The relatives of the patient as well will benefit from the information: they will undergo genetic testing aimed at searching for the CNV variant and, if carriers, they will follow a specific preventive program. Another patient affected with colon cancer and screened for *MMR* genes was found to carry a pathogenic *if* in *CDKN2A* gene: this patient, once informed, will be offered to follow

a preventive program for melanoma. Similarly, another subject affected with colon cancer and screened for *APC* and *MUTYH* genes, was found to carry a pathogenic *if* in *PALB*2 gene: this patient, once informed, will be offered to follow a preventive program for breast cancer.

An important challenge to consider is the discordant interpretation of variants' significance across laboratories. Laboratories may differ in how they apply interpretation criteria, may have analyzed a variant at different points in time (*e.g.* before and after release of population data), or may have access to nonpublic evidence about a specific variant. Specifically, differences in the application of the International Agency for Research on Cancer (IARC) recommendations, the American College of Medical Genetics and Genomics (ACMG) and the Association for Molecular Pathology (AMP) criteria, such as different definitions of mutation hotspots, or differences in the application of classification criteria related to variant frequency in control populations, are common sources of discordance (Amendola *et al*., 2020). Laboratories may also differ in their application of gene-specific variant interpretation criteria, which can lead to discordant interpretations (Amendola *et al*., 2020). Discordant variant interpretations can be particularly problematic if members of the same family are tested in different laboratories and receive different classifications for the same mutation. To improve consistency in variant interpretation, data sharing among laboratories and through public databases such as ClinVar (Landrum *et al*., 2018) is essential. In our lab, all identified variants are classified according to ENIGMA consortium (for *BRCA1* and *BRCA2* genes), ClinVar public archive, Varsome, InSIGHT variants and LOVD databases, and discussed in lab meetings.

Another important issue is the choice of the most appropriate molecular test and the understanding of the limits of technology and methods applied. While NGS is highly effective for detecting single nucleotide variations (SNVs) such as base substitutions and small insertions and deletions, it is inappropriate to detect some types of genomic variations, such as large chromosomal alterations (Piccinin *et al*., 2019). As already mentioned, alternative technologies are needed in particular cases. MLPA is used to characterize copy number variations; translocations and inversions are almost undetectable by using NGS and may require karyotyping or microarray to be identified (Piccinin *et al*., 2019). It can be particularly difficult to characterize some large duplications. In a large-scale study of chromosome duplications, Newman and collaborators (2015) found that most duplications (83%) are in tandem and in direct orientation relative to the original locus. Unfortunately, both NGS and MLPA appear inadequate to determine the precise structure and orientation of duplicated regions.

By NGS technology, bioinformatic read-based alignment issues can hamper the ability to align reads to the correct DNA reference sequence. This is the case of pseudogenes with regions of gene homology, repeat regions, and, in some cases, regions with GC-rich sequences. Misalignment of sequencing reads to a pseudogene can lead to false negative or positive variant calls if not confirmed with an orthologous methodology (Hegde *et al*., 2014) and validation of these variants by Sanger sequencing is also challenging. Alternative molecular testing strategies (*e.g.* analysis of long-range PCR products) are needed to resolve false positive calls. For example,

Discussion

PMS2 gene (implicated in Lynch syndrome) has regions (exons 11–15) (Li *et al*., 2015) with over 98% homology with *PMS2CL* pseudogene, and variants in the pseudogene are frequently identified. In our lab, confirmation testing using a 17 Kb long range PCR combined with Sanger sequencing allowed us to validate one pathogenic variant and three VUSs in *PMS2* gene. At the same time, it allowed us to confirm that one pathogenic variant and four VUSs were located in *PMS2CL* pseudogene. Another cancer-associated gene for which there are highly homologous regions is *CHEK2* gene. *CHEK2* pseudogenes overlap with exons 10 through 14 and share 95-98% sequence homology with *CHEK2* gene (Chromosome 22). These pseudogenes are located on chromosome 15 (*CHEK2 p2*), chromosome 2 (*CHEK2 p3*), chromosome 22 (*CHEK2 p4*), and chromosome 10 (*CHEK2 p5*). We developed a long-range PCR to amplify a sequence of about 1600 kb by using primers targeting a specific region unique of *CHEK2* and we were able to confirm that three pathogenic variants and three VUSs were in the gene.

For some genes, there are clinical practice guidelines to guide management of individuals with pathogenic variants, and evidence to support the clinical effectiveness of treatment or preventive interventions. New guidance is published or updated over time as evidence accrues. A diagnosis of a hereditary cancer syndrome can influence how cancer is treated (*e.g.* through targeted therapies such as PARP inhibitors) or prevented (*e.g.* through earlier or more intensive surveillance, risk-reducing surgeries, and chemoprevention). Numerous clinical management resources exist for *BRCA1/BRCA2*-associated breast and ovarian cancers (NCCN, 2021). NCCN guidelines also cover management of moderate penetrance breast cancer genes (*e.g CHEK2*, *ATM*), as well as other syndromes associated with breast cancer such as Cowden syndrome/*PTEN* Hamartoma syndrome (NCCN, 2021). A recent statement from the ACMG addresses the management of *PALB2* heterozygous subjects (Tischkowitz *et al*., 2021). Clinical practice guidelines and recommendations exist for the management of Lynch syndrome (Mishima *et al*., 2020; Holter *et al.,* 2022), as well as polyposis syndromes (Colas *et al*., 2020). However, for many rare hereditary cancer syndromes or genes, evidence-based guidelines for patient management and surveillance are not available (Hall *et al*., 2014). Germline testing may identify a genetic lesion for which clinical management is uncertain or for which interventions for prevention or early detection are not available. Indeed, in our screenings I identified pathogenic mutations in *ATM* (6/61 - 9.8%), *BARD1* (1/61 - 1.6%) and *RAD51C* (1/61 - 1.6%) genes for which clinical treatments are not defined.

The identification of pathogenic variants in moderate penetrance genes may raise doubts about patient management, due to lack of evidence on the effectiveness of surveillance or preventive interventions. In some cases, the lifetime cancer risk may not meet local thresholds for eligibility for high-risk surveillance. For example, heterozygous truncating variants in *ATM* are associated with a 20% lifetime risk of breast cancer, a value which may not meet cutoffs for high-risk breast cancer screening in some countries (Zhang *et al*., 2011). Over time, as new evidence emerges on penetrance, disease natural history, and effectiveness of preventive and therapeutic interventions, recommendations for the management of hereditary cancers will continue to evolve.

Discussion

Regarding management and treatment, a particular attention deserves the project we called "*Persona Project".* For patients with triple negative breast cancer (TNBC), the implementation of genetic testing in decision making might impact both risk management for the patient and her family, and therapeutic management. The identification of genetically predisposed subjects dictates risk-reducing strategies that may imply bilateral salpingo-oophorectomy and mastectomy or long-term medical approaches. Of relevance, the selection of patients for testing has long relied on the presence of a strong family history of breast and ovarian cancer. On the contrary, aims of our "*Persona Project"* were to evaluate the presence of germline variants in TNBC patients regardless of family history and to assess the prognostic and therapeutic predictive role of homologous recombination (HR) defects *vs.* other genetic alterations. In this frame, I sequenced and analyzed 318 samples by using the Trusight Sequencing Cancer Panel (94 cancer-related genes and 286 SNPs). Specifically, I searched for variants in 14 actionable genes related to health conditions with known medical cure and/or risk cancer prevention recommendations (*BRCA1, BRCA2, APC, MUTYH, MLH1, MSH2, MSH6, EPCAM, PMS2, CDH1, CDKN2A, CDK4, TP53, PTEN*). At the same time, I also annotated the variants identified in all the other genes of the panel. As already discussed, the use of a multigene panel can detect many VUS and variants in unexpected cancer predisposition genes (Eliade *et* al., 2017).

Indeed, only 8 individuals (2.5%) of our cohort were wild type for all genes of the panel while 90 patients (28.3%) were carriers of pathogenic variants in one or more genes of the panel. Out of 318 individuals, 43 (13.5%) had a pathogenic mutation in *BRCA1* and 10 (3%) in *BRCA2*. These results are consistent with various studies reporting that 15%-20% of women with TNBC carry *BRCA1* or *BRCA2* germline mutations, and that 70% of breast cancers that develop in *BRCA1* mutation carriers are "triple-negative" (TNBCs) (Stevens *et al*., 2013). Among individuals carrying pathogenic mutations in hereditary breast/ovarian cancer (HBOC) predisposition genes, the c.5266dupC variant represented the 11.6% of all *BRCA1* pathogenic lesions; the c.1088delA and the c.798_799delTT variants represented the 9.3% and the 7% of all *BRCA1* pathogenic mutations, respectively. The c.5266dupC is the most prevalent *BRCA1* breast cancer related mutation in European countries (Karami *et al.*, 2013) and is a very common variant in Apulian Southern Italy population. This prevalence is likely related to the historical/political changes occurred in Apulia over time, with the colonization of the region by the Normans around 1000 D.C. and the presence of numerous Jewish communities (founder effect) contributing to the spread of the variant in the region (Patruno *et* al., 2021). The c.1088delA is another common Italian variant (Figlioli *et al.* 2021), while the c.798_799delTT has been reported in Moroccan, Algerian and Tunisian breast cancer families as a Northern Africa founder mutation in (Zoure *et al*. 2018) then spread in European populations.

I found 4 TNBC patients with pathogenic variants in *MMR* genes: 2 in *MLH1*, 1 in *PMS2*, and 1 in *EPCAM* gene. Previous studies on large cohorts of Lynch syndrome patients already reported that *MMR* mutation carriers may present with a hereditary breast and ovarian cancer phenotype (Espenschied *et al*., 2017).

Discussion

I also analyzed our samples to search for two disease-associated SNPs (daSNPs): rs721048 and rs2046210. The first daSNP is an intronic variant (c.1290+30064G>A) in *EHBP1* gene and it is associated with prostate cancer (Ao et al., 2015). Up to 35% of patients affected with TNBC proved to carry the risk allele (A): we are now interested in comparing the allele frequencies in TNBC, in subjects diagnosed with different types of cancer, and in healthy individuals. The second daSNP is an intergenic variant (in proximity to the C6orf97 and estrogen receptor 1-ESR1 genes) and the risk allele is frequently recurring in breast cancer patients with anti-estrogen therapy resistance and poor prognosis. In agreement with literature data (Jin *et al.* 2019), this polymorphism may serve as a potential genetic biomarker of BC in both Asians and Caucasians. Fifty-two percent of the individuals of "*Persona project*" carried this same variant. In this case, we are interested in establishing a possible co-occurrence of the risk allele with other frequent pathogenic variants in *BRCA1* (c.5266dupC, c.188delA, c.798_799delTT), as well with other lesions in actionable genes.

Hereditary cancer genetic testing is an expanding field that promises to identify people at risk of cancer and has proven diagnostic and prognostic utility. However, using molecular testing as a tool in the armamentarium for the treatment and management of patients with hereditary cancer involves knowing various basic principles. The technology continues to evolve, and while often more sensitive, may also have limitations. Keeping pace with new technologies in the laboratory also means upgrading and validating informatics pipelines and algorithms routinely to be able to identify different types of sequence variation. With the advent of publicly accessible online databases and an increase in data sharing practices globally, knowledge is expanding at rate that requires routine examination. Guidelines for treatment and management of hereditary cancer will routinely be updated; clearly communicating genetic results and providing appropriate follow up recommendations will see continual refinement and further reduce the morbidity and mortality associated with hereditary cancers.

8. References

Adams DJ, Bishop DT, Robles-Espinoza CD. Melanoma predisposition-A limited role for germline BRCA1 and BRCA2 variants. *Pigment Cell Melanoma Res*. 33(1):6-7 (2020)

Akdeniz D, Schmidt MK, Seynaeve CM, et al. Risk factors for metachronous contralateral breast cancer: A systematic review and meta-analysis. *Breast.* 44:1-14 (2019)

AlHarthi FS, Qari A, Edress A, Abedalthagafi M. Familial/inherited cancer syndrome: a focus on the highly consanguineous Arab population. *npj Genomic Medicine* volume 5, Article number: 3 (2020)

Amendola LM, Muenzen K, Biesecker LG, et al. Variant classification concordance using the ACMG-AMP variant interpretation guidelines across nine genomic implementation research studies. *Am J Hum Genet*. 107(5):932-941 (2020)

Ao X, Liu Y, Bai XY, Qu X, Xu Z, Hu G, Chen M, Wu H. Association between EHBP1 rs721048(A>G) polymorphism and prostate cancer susceptibility: a meta-analysis of 17 studies involving 150,678 subjects. *Onco Targets Ther*. 8:1671-80 (2015)

Apostolou P, Fostira F. Hereditary breast cancer: the era of newsusceptibility genes. *Biomed Res Int*. 2013:747318 (2013)

Arts-de Jong M, de Bock GH, van Asperen CJ, et al. Germline BRCA1/2 mutation testing is indicated in every patient with epithelialovarian cancer: a systematic review. *Eur J Cancer*. 61:137-145 (2016)

Atchley DP, Albarracin CT, Lopez A, Valero V, Amos CI, Gonzalez-Angulo AM, Hortobagyi GN, Arun BK. Clinical and pathologic characteristics of patients with BRCA-positive and BRCA-negative breast cancer. *J Clin Oncol.* 26: 4282-8 (2008)

Ayed-Guerfali DB, Kridis-Rejab WB, Ammous-Boukhris N, Ayadi W, Charfi S, Khanfir A, Sellami-Boudawara T, Frikha M, Daoud J & Mokdad-Gargouri R. Novel and recurrent *BRCA1/BRCA2* germline mutations in patients with breast/ovarian cancer: a series from the south of Tunisia. *Journal of Translational Medicine* 19, Article number: 108 (2021)

Bailey CE, Hu CY, You YN, et al. Increasing disparities in the age- related incidences of colon and rectal cancers in the United States, 1975-2010. *JAMA Surg*. 1-6 (2014)

References

Balmana J, Digiovanni L, Gaddam P, et al. Conflicting interpretation of genetic variants and cancer risk by commercial laboratories as assessed by the prospective registry of multiplex testing. *J Clin Oncol*. 34:4071-4078 (2016)

Blackwood MA, Weber BL. BRCA1 and BRCA2: from molecular genetics to clinical medicine. *J Clin Oncol*.16:1969-1977 (1998)

Blair VR, McLeod M, Carneiro F, Coit DG, D'Addario JL, van Dieren JM, Harris KL, Hoogerbrugge N, Oliveira C, van der Post RS, Arnold J, Benusiglio PR, Bisseling TM, Boussioutas A, Cats A, Charlton A, Schreiber KEC, Davis JL, Pietro MD, Fitzgerald RC, Ford JM, Gamet K, Gullo I, Hardwick RH, Huntsman DG, Kaurah P, Kupfer SS, Latchford A, Mansfield PF, Nakajima T, Parry S, Rossaak J, Sugimura H, Svrcek M, Tischkowitz M, Ushijima T, Yamada H, Yang HK, Claydon A, Figueiredo J, Paringatai K, Seruca R, Bougen-Zhukov N, Brew T, Busija S, Carneiro P, DeGregorio L, Fisher H, Gardner E, Godwin TD, Holm KN, Humar B, Lintott CJ, Monroe EC, Muller MD, Norero E, Nouri Y, Paredes J, Sanches JM, Schulpen E, Ribeiro AS, Sporle A, Whitworth J, Zhang L, Reeve AE, Guilford P. Hereditary diffuse gastric cancer: updated clinical practice guidelines. *Lancet Oncol*. 21(8):e386 (2020)

Blazer KR, Slavin T, Weitzel JN. Increased reach of genetic cancer risk assessment as a tool for precision management of hereditary breast cancer. *JAMA Oncol.* 2:723-724 (2016)

Bombard Y, Bach PB, Offit K. Translating genomics in cancer care. *JNatl Compr Canc Netw*. 11:1343-1353 (2013)

Bonadona V, Bonaïti B, Olschwang S, et al. Cancer Risks Associated With Germline Mutations in *MLH1*, *MSH2*, and *MSH6* Genes in Lynch Syndrome. *JAMA.* 305(22):2304–2310 (2011)

Boujemaa M, Hamdi Y, Mejri N, Romdhane L, Ghedira K, Bouaziz H, Benna HE, Labidi S, Dallali H, Jaidane O, Nasr SB, Haddaoui A, Rahal K, Abdelhak S, Boussen H, Boubaker MS. Germline copy number variations in *BRCA1/2* negative families: Role in the molecular etiology of hereditary breast cancer in Tunisia. *PLoS One* (2021)

Brunet J, Gutierrez-Enriquez S, Torres A, et al. ATM germline mutations in Spanish early-onset breast cancer patients negative for BRCA1/BRCA2 mutations. *Clin Genet.* 73:465-473 (2008)

Buys SS, Sandbach JF, Gammon A, et al. A study of over 35,000women with breast cancer tested with a 25-gene panel of hereditary cancer genes. *Cancer*. 123:1721-1730 (2017)

Caglayan G., Nurgul Y., Abdullah S. Germline Mutations in BRCA1 and BRCA2 in Breast Cancer Patients with High Genetic Risk in Turkish Population. *Interational Journal of Breast cancer* (2019)

References

Castro E, Goh C, Olmos D, et al. Germline BRCA mutations are associated with higher risk of nodal involvement, distant metastasis, andpoor survival outcomes in prostate cancer. *J Clin Oncol*. 31:1748- 1757 (2013)

Cavenee WK, Dryja TP, Phillips RA, et al. Expression of recessive alleles by chromosomal mechanisms in retinoblastoma. Nature. 305(5937):779-784 (1983)

Cheng L, Eng C, Nieman LZ, et al. Trends in colorectal cancer incidence by anatomic site and disease stage in the United States from1976 to 2005. *Am J Clin Oncol*. 34:573-580 (2011)

Chiang J, Chia TH, Yuen J, et al. Impact of variant reclassification in cancer predisposition genes on clinical care. JCO Precis Oncol. 5:577-584 (2021)

Church J, Heald B, Burke C, Kalady M. Understanding MYH-associated neoplasia. *Dis Colon Rectum.* 55(3):359-62 (2012)

Colas C, Bonadona V, Baert-Desurmont S, et al. MUTYH-associated polyposis: review and update of the French recommendations established in 2012 under the auspices of the National Cancer institute (INCa). *Eur J Med Genet*. 63(12):104078 (2020)

Couch FJ, Shimelis H, Hu C, Hart SN, Polley EC, Na J, Hallberg E, Moore R, Thomas A, Lilyquist J, Feng B, McFarland R, Pesaran T, Huether R, LaDuca H, Chao EC, Goldgar DE, Dolinsky JS. Associations Between Cancer Predisposition Testing Panel Genes and Breast Cancer. *JAMA Oncol*. 3(9):1190-1196 (2017)

Cybulski, C., Wokołorczyk, D., Kładny, J. *et al.* Germline CHEK2 mutations and colorectal cancer risk: different effects of a missense and truncating mutations?. *Eur J Hum Genet.* 15, 237–241 (2007)

Deng G, Chen A, Hong J, Chae HS, Kim YS. Methylation of CpG in a small region of the hMLH1 promoter invariably correlates with the absence of gene expression. *Cancer Res*. 59(9):2029-2033 (1999)

Dite GS, Jenkins MA, Southey MC, Hocking JS, Giles GG, McCredie MR, Venter DJ, Hopper JL. Familial risks, early-onset breast cancer, and BRCA1 and BRCA2 germline mutations. *J Natl Cancer Inst*. Mar 19;95(6):448-57 (2003)

Domagala P, Jakubowska A, Jaworska-Bieniek K, Kaczmarek K, Durda K, Kurlapska A, Cybulski C, Lubinski J. Prevalence of Germline Mutations in Genes Engaged in DNA Damage Repair by Homologous Recombination in Patients with Triple-Negative and Hereditary Non-Triple-Negative Breast Cancers. *PLoS One.* 17; 10 (2015)

Domchek SM, Robson ME. Update on genetic testing in gynecologiccancer. *J Clin Oncol*. 37:2501-2509 (2019)

References

Easton DF, Pharoah PD, Antoniou AC, Tischkowitz M, Tavtigian SV, Nathanson KL, Devilee P, Meindl A, Couch FJ, Southey M, Goldgar DE, Evans DG, Chenevix-Trench G, Rahman N, Robson M, Domchek SM, Foulkes WD. Gene-panel sequencing and the prediction of breast-cancer risk. *N Engl J Med.* 372: 2243-57 (2015)

Eccles DM, Mitchell G, Monteiro AN, et al. BRCA1 and BRCA2 genetictesting-pitfalls and recommendations for managing variants of uncertain clinical significance. *Ann Oncol*. 26:2057-2065 (2015)

Eliade M, Skrzypski J, Baurand A, Jacquot C, Bertolone G, Loustalot C, Coutant C, Guy F, Fumoleau P, Duffourd Y, Arnould L, Delignette A, Padéano MM, Lepage C, Raichon-Patru G, Boudrant A, Bône-Lépinoy MC, Villing AL, Charpin A, Peignaux K, Chevrier S, Vegran F, Ghiringhelli F, Boidot R, Sevenet N, Lizard S, Faivre L. The transfer of multigene panel testing for hereditary breast and ovarian cancer to healthcare: What are the implications for the management of patients and families? *Oncotarget*. 10;8(2):1957-1971 (2017)

Espenschied CR, LaDuca H, Li S, McFarland R, Gau CL, Hampel H. Multigene Panel Testing Provides a New Perspective on Lynch Syndrome. *J Clin Oncol*. 35(22):2568-2575 (2017)

Farooqui SM, Ward R, Aziz M. Shwachman-Diamond Syndrome. 2022 Apr 30. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2022 Jan

Figlioli G, De Nicolo A , Catucci I , Manoukian S , Peissel B , Azzollini J , Beltrami B , Bonanni B, Calvello M ,Bondavalli D , Pasini B,6, Vignolo Lutati F , Ogliara P , Zuradelli M , Pensotti V , De Vecchi G , Volorio S , Verderio P , Pizzamiglio S , Matullo G, Aneli S , Birolo G , Zanardi F, Tondini C, Zambelli A, Livraghi L, Franchi M, Radice P and Peterlongo P. Analysis of Italian BRCA1/2 Pathogenic Variants Identifies a Private Spectrum in the Population from the Bergamo Province in Northern Italy. *Cancers* (2021)

Figueiredo J, Melo S, Carneiro P, Moreira AM, Fernandes MS, Ribeiro AS, Guilford P, Paredes J, Seruca R. Clinical spectrum and pleiotropic nature of *CDH1* germline mutations. *J Med Genet*. 56(4):199-208 (2019)

Finch A, Beiner M, Lubinski J, et al. Salpingo-oophorectomy and therisk of ovarian, fallopian tube, and peritoneal cancers in women with a BRCA1 or BRCA2 Mutation. *JAMA*. 296:185-192 (2006)

Fitzgerald RC, Hardwick R, Huntsman D, Carneiro F, Guilford P, Blair V, Chung DC, Norton J, Ragunath K, Van Krieken JH, Dwerryhouse S, Caldas C, International Gastric Cancer Linkage Consortium. Hereditary diffuse gastric cancer: updated consensus guidelines for clinical management and directions for future research. *J Med Genet*. 47(7):436 (2010)

Foulkes WD, Smith IE, Reis-Filho JS. Triple-negative breast cancer. *N Engl J Med.* 363:1938-48 (2010)

References

Gallagher DJ, Gaudet MM, Pal P, et al. Germline BRCA mutations denote a clinicopathologic subset of prostate cancer. *Clin Cancer Res*. 16:2115-2121 (2010)

George A. UK BRCA mutation testing in patients with ovarian cancer. *Br J Cancer*. 113 Suppl 1(Suppl 1):S17-S21 (2015)

Giri VN, Hegarty SE, Hyatt C, et al. Germline genetic testing forinherited prostate cancer in practice: Implications for genetic testing,precision therapy, and cascade testing. *Prostate*. 79:333-339 (2019)

Glazer AM, Wada Y, Li B, et al. High-throughput reclassification of SCN5A variants. *Am J Hum Genet*. 107(1):111-123 (2020)

Goldberg M, Bell K, Aronson M, Semotiuk K, Pond G, Gallinger S, Zbuk K. Association between the Lynch syndrome gene MSH2 and breast cancer susceptibility in a Canadian familial cancer registry. *J Med Genet.* 54(11):742-746 (2017)

Green RC, Berg JS, Grody WW, et al. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. *Genet Med*.;15:565–74 (2013)

GTEx Consortium. Erratum: Genetic effects on gene expression across human tissues. Nature. 2018 Jan 25;553(7689):530. doi: 10.1038/nature25160. Epub 2017 Dec 20. Erratum for: *Nature*. 550(7675):204-213 (2017)

Hall MJ, Forman AD, Pilarski R, et al. Gene panel testing for inherited cancer risk. *J Natl Compr Canc Netw.* 12:1339-1346 (2014)

Hall MJ, Obeid E, Daly MB. Multigene panels to evaluate hereditarycancer risk: reckless or relevant? *J Clin Oncol*. 34:4186-4187 (2016)

Hegde M, Ferber M, Mao R, Samowitz W, Ganguly A, Committee WGotACoMGaGALQA. ACMG technical standards and guidelines for genetic testing for inherited colorectal cancer (Lynch syndrome, familial adenomatous polyposis, and MYH-associated polyposis). *Genet Med*. 16(1):101-116 (2014)

Heikkinen K, Rapakko K, Karppinen SM, et al. Association of common ATM polymorphism with bilateral breast cancer. *Int J Cancer*. 116:69-72 (2005)

Hitchins MP, Wong JJ, Suthers G, et al. Inheritance of a cancerassociated MLH1 germ-line epimutation. *N Engl J Med*. 356(7): 697-705 (2007)

Holter S, Borgida A, Dodd A, et al. Germline BRCA mutations in alarge clinic-based cohort of patients with pancreatic adenocarcinoma. *JClin Oncol*. 33:3124-3129 (2015)

References

Holter S, Hall MJ, Hampel H, et al. Risk assessment and genetic counseling for Lynch syndrome—practice resource of the National Society of Genetic Counselors and the Collaborative Group of the Americas on Inherited Gastrointestinal Cancer. *J Genet Couns.* Published online ahead of print, (2022)

Hong YC, Liu HM, Chen PS, et al. Hair follicle: a reliable source of recipient origin after allogeneic hematopoietic stem cell transplantation. *Bone Marrow Transplant*. 40:871-874 (2007)

Hu C, Hart SN, Polley EC, et al. Association between inherited germline mutations in cancer predisposition genes and risk of pancreatic cancer. *JAMA*. 319:2401-2409 (2018)

Hu C, Polley EC, Yadav S, Lilyquist J, Shimelis H, Na J, Hart SN, Goldgar DE, Shah S, Pesaran T, Dolinsky JS, LaDuca H, Couch FJ. The Contribution of Germline Predisposition Gene Mutations to Clinical Subtypes of Invasive Breast Cancer From a Clinical Genetic Testing Cohort. J Natl Cancer Inst. 14;112(12):1231-1241 (2020)

Imai K, Yamamoto H, Carcinogenesis and microsatellite instability: the interrelationship between genetics and epigenetics. *Carcinogenesis*. Volume 29, Issue 4, Pages 673–680 (2008)

Issa JP. The Two-Hit Hypothesis Meets Eigenetics. *Cancer Research*. 82 (7): 1167–1169 (2022)

Jerzak KJ, Mancuso T, Eisen A. Ataxia-telangiectasia gene (*ATM*) mutation heterozygosity in breast cancer: a narrative review. *Curr Oncol*. Apr;25(2):e176-e180 (2018)

Jin TF, Zhang WT, Zhou ZF. The 6q25.1 rs2046210 polymorphism is associated with an elevated susceptibility to breast cancer: A meta-analysis of 261,703 subjects. *Mol Genet Genomic Med*. 7(3):e553 (2019)

Karami F, Parvin Mehdipour P, A comprehensive focus on global spectrum of *BRCA1* and *BRCA2* mutations in breast cancer. *Biomed Res Int.* (2013)

Krepischi ACV, Achatz MIW, Santos, EMM, Costa SS, Lisboa BCG, Brentani H, Santos TM, Gonçalves A, Nóbrega AF, Pearson PL, Morgante AMV, Carraro DM, Brentani RR & Rosenberg C. Germline DNA copy number variation in familial and early-onset breast cancer. *Breast Cancer Research* volume 14, Article number: R24 (2012)

Krepischi ACV, Pearson PL , Rosenberg C. Germline copy number variations and cancer predisposition. *Future Oncol* (2012)

References

Kuchenbaecker KB, Hopper JL, Barnes DR, et al. Risks of breast, ovarian, and contralateral breast cancer for BRCA1 and BRCA2 mutation carriers. *JAMA*. 317:2402-2416 (2017)

Kumaran M, Cass CE, Mackey JR, Hubaux R, Lam W, Yasui Y, Damaraju S. Germline copy number variations are associated with breast cancer risk and prognosis. *Scientific Reports* volume 7, Article number: 14621 (2017)

Kurian AW, Hare EE, Mills MA, et al. Clinical evaluation of a multiple- gene sequencing panel for hereditary cancer risk assessment. *J Clin Oncol*. 32:2001-2009 (2014)

Kurian AW, Ward KC, Hamilton AS, et al. Uptake, results, and outcomes of germline multiple-gene sequencing after diagnosis of breast cancer. *JAMA Oncol*. 4:1066-1072 (2018)

Kwong A, Chen J, Shin VY, Ho JC, Law FB, Au CH, Chan TL, Ma ES, Ford JM. The importance of analysis of long-range rearrangement of BRCA1 and BRCA2 in genetic diagnosis of familial breast cancer. *Cancer Genet*. 208(9):448-54 (2015)

LaDuca H, Stuenkel AJ, Dolinsky JS, et al. Utilization of multigene panels in hereditary cancer predisposition testing: analysis of more than2,000 patients. *Genet Med*. 16:830-837 (2014)

LaDuca H, Polley EC, Yussuf A, et al. A clinical guide to hereditary cancer panel testing: evaluation of gene-specific cancer associations and sensitivity of genetic testing criteria in a cohort of 165,000 high-risk patients. *Genet Med*. 22:407-415 (2020)

Landrum MJ, Lee JM, Benson M, et al. ClinVar: improving access to variant interpretations and supporting evidence. Nucleic Acids Res. 46(D1):D1062-D1067 (2018)

Lebo MS, Zakoor KR, Chun K, et al. Data sharing as a national quality improvement program: reporting on BRCA1 and BRCA2 variantinterpretation comparisons through the Canadian Open Genetics Repository (COGR). *Genet Med*. 20(3):294-302 (2018)

Li J, Dai H, Feng Y, et al. A comprehensive strategy for accurate mutation detection of the highly homologous PMS2. *J Mol Diagn*. 17(5):545-553 (2015)

Lincoln SE, Yang S, Cline MS, et al. Consistency of BRCA1 and BRCA2 variant classifications among clinical diagnostic laboratories. *JCOPrecis Oncol*. 1 (2017)

Lips EH, Michaut M, Hoogstraat M, et al. Next generation sequencing of triple negative breast cancer to find predictors for chemotherapy response. *Breast Cancer Res*. 17: 134 (2015)

References

Lang SH, Swift SL, White H, et al. A systematic review of the prevalence of DNA damage response gene mutations in prostate cancer. *Int J Oncol*. 55:597-616 (2019)

Li GM. Mechanisms and functions of DNA mismatch repair. *Cell Research.* 18, 85 - 98 (2008)

Li X, You R, Wang X, et al. Effectiveness of prophylactic surgeries in BRCA1 or BRCA2 mutation carriers: a meta-analysis and systematic review. *Clin Cancer Res*. 22(15):3971-3981 (2016)

Liu C, Wang QS, Wang YJ. The CHEK2 I157T variant and colorectal cancer susceptibility: a systematic review and meta-analysis. *Asian Pac J Cancer Prev*. 13(5):2051-5 (2012)

Lucas AL, Shakya R, Lipsyc MD, et al. High prevalence of BRCA1 and BRCA2 germline mutations with loss of heterozygosity in a series ofresected pancreatic adenocarcinoma and other neoplastic lesions. *Clin Cancer Res*. 19:3396-3403 (2013)

Macklin S, Durand N, Atwal P, Hines S. Observed frequency and challenges of variant reclassification in a hereditary cancer clinic. *Genet Med*. 20(3):346-350 (2018)

Manchanda R, Patel S, Gordeev VS, et al. Cost-effectiveness of population-based BRCA1, BRCA2, RAD51C, RAD51D, BRIP1, PALB2 mutation testing in unselected general population women. *J Natl CancerInst.* 110:714-725 (2018)

Mavaddat N, Peock S, Frost D, et al. Cancer risks for BRCA1 and BRCA2 mutation carriers: results from prospective analysis of EMBRACE.*J Natl Cancer Inst*.;105:812-822 (2013)

Mersch J, Brown N, Pirzadeh-Miller S, et al. Prevalence of variant reclassification following hereditary cancer genetic testing. JAMA. 320(12):1266-1274 (2018)

Meyer P, Landgraf K, Hogel B, et al. BRCA2 mutations and triple-negative breast cancer. *PLoS One*. 7:e38361 (2012)

Mighton C, Lerner-Ellis JP. Principles of molecular testing for hereditary cancer. *Genes Chromosomes Cancer*. Jun;61(6):356-381 (2022)

Mishima S, Taniguchi H, Akagi K, et al. Japan Society of Clinical Oncology provisional clinical opinion for the diagnosis and use of immunotherapy in patients with deficient DNA mismatch repair tumors, cooperated by Japanese Society of Medical Oncology, first edition. *Int J Clin Oncol*. 25(2):217-239 (2020)

Na R, Zheng SL, Han M, et al. Germline mutations in ATM and BRCA1/2 distinguish risk for lethal and indolent prostate cancer and areassociated with early age at death. *Eur Urol*. 71:740-747 (2017)

References

Nagy R, Sweet K, Eng C. Highly penetrant hereditary cancer syndromes. *Oncogene*. 23(38):6445-6470 (2004)

National Comprehensive Cancer Network (NCCN). NCCN Clinical Practice Guidelines in Oncology. Genetic/Familial High-Risk Assessment: Breast, Ovarian, and Pancreatic. Version 1.2022. National Comprehensive Cancer Network (NCCN); 2021

Newman S, Hermetz KE, Weckselblatt B, Rudd MK. Nextgeneration sequencing of duplication CNVs reveals that most are tandem and some create fusion genes at breakpoints. *Am J Hum Genet*. 96(2):208-220 (2015)

Norquist BM, Harrell MI, Brady MF, et al. Inherited mutations in women with ovarian carcinoma. *JAMA Oncol.* 2:482-490 (2016)

Oh M, Alkhushaym N, Fallatah S, Althagafi A, Aljadeed R, Alsowaida Y, Jeter J, Martin JR, Babiker HM, McBride A, Abraham I. The association of BRCA1 and BRCA2 mutations with prostate cancer risk, frequency, and mortality: A meta-analysis. *The Prostate*. 79(8):880-895 (2019)

Pal T, Agnese D, Daly M, et al. Points to consider: is there evidence to support BRCA1/2 and other inherited breast cancer genetic testing for all breast cancer patients? A statement of the American College of Medical Genetics and Genomics (ACMG). *Genet Med*. 22:681-685 (2020)

Pan W, Lu K, Wang W, Yao J, Hou Y. PALB2 as a potential prognostic biomarker for colorectal cancer. Comput Biol Chem. 20;87:107289 (2020)

Patruno M, De Summa S, Resta N, Caputo M, Costanzo S, Digennaro M, Pilato B, Bagnulo R, Pantaleo A, Simone C, Natalicchio MI, De Matteis E, Tarantino P, Tommasi S, Paradiso A. Spectrum of Germline Pathogenic Variants in BRCA1/2 Genes in the Apulian Southern Italy Population: Geographic Distribution and Evidence for Targeted Genetic Testing. *Cancers (Basel)*. 21;13(18):4714 (2021)

Peleg Hasson S, Menes T, Sonnenblick A. Comparison of Patient Susceptibility Genes Across Breast Cancer: Implications for Prognosis and Therapeutic Outcomes. *Pharmgenomics Pers Med*. Jul 27;13:227-238 (2020)

Petrucelli N, Daly MB, Bars Culver JO, Feldman GL. BRCA1 and BRCA2 hereditary breast/ovarian cancer. *GeneReviews*. (2011)

Petrucelli N, Daly MB, Pal T. BRCA1- and BRCA2-Associated Hereditary Breast and Ovarian Cancer. *Gene Reviews*. (2022)

Piccinin C, Panchal S, Watkins N, Kim RH. An update on genetic risk assessment and prevention: the role of genetic testing panels in breast cancer. *Expert Rev Anticancer Ther*. 19(9):787-801 (2019)

References

Pilarski R. Cowden syndrome: a critical review of the clinical literature. *J Genet Couns*. 18:13-27 (2009)

Pineda M, Mur P, Iniesta MD, et al. MLH1 methylation screening is effective in identifying epimutation carriers. *Eur J Hum Genet*. 20(12):1256-1264 (2012)

Piri L. Welcsh, Mary-Claire King, *BRCA1* and *BRCA2* and the genetics of breast and ovarian cancer, *Human Molecular Genetics*. Vol 10, Issue 7, 705–713 (2001)

Potrony M, Badenas C, Aguilera P, Puig-Butille JA, Carrera C, Malvehy J, Puig S. *Update in genetic susceptibility in melanoma*. Vol 3, No 15 (2015)

Pritchard CC, Mateo J, Walsh MF, et al. Inherited DNA-repair genemutations in men with metastatic prostate cancer. *N Engl J Med*. 375:443-453 (2016)

Rainville IR, Rana HQ. Next-generation sequencing for inherited breast cancer risk: counseling through the complexity. *Curr Oncol Rep*. 16:371 (2014)

Rennert G, Lejbkowicz F, Cohen I, Pinchev M, Rennert HS, Barnett-Griness O. MutYH mutation carriers have increased breast cancer risk. *Cancer*. 118(8):1989-93 (2012)

Ribeiro Moura Brasil Arnaut J, Isabella dos Santos Guimarães, Anna Cláudia Evangelista dos Santos, Flora de Moraes Lino da Silva, Jorge Ricardo Machado, Andreia Cristina de Melo. Molecular landscape of Hereditary Melanoma. *Critical Reviews in Oncology/Hematology*.Volume 164 (2021)

Robson ME, Bradbury AR, Arun B, et al. American Society of Clinical Oncology Policy statement update: genetic and genomic testing for cancersusceptibility. *J Clin Oncol*. 33:3660-3667 (2015)

Rumilla K, Schowalter KV, Lindor NM, et al. Frequency of deletions ofEPCAM (TACSTD1) in MSH2-associated Lynch syndrome cases. *J Mol Diagn.* 13:93-99 (2011)

Salo-Mullen EE, O'Reilly EM, Kelsen DP, et al. Identification of germline genetic mutations in patients with pancreatic cancer. *Cancer*. 121:4382-4388 (2015)

Schaub MA, Boyle AP, Kundaje A, Batzoglou S, Snyder M. Linking disease associations with regulatory information in the human genome. *Genome Res*. 22(9):1748-59 (2012)

Schneider KA, Garber J. Li-Fraumeni syndrome. *GeneReviews*. (2013)

Sheehan M, Heald B, Yanda C, Kelly ED, Grobmyer S, Eng C, Kalady M, Pederson H. Investigating the Link between Lynch Syndrome and Breast Cancer. *Eur J Breast Health*. 1;16(2):106-109 (2020)

References

Shimelis H, LaDuca H, Hu C, et al. Triple-negative breast cancer riskgenes identified by multigene hereditary cancer panel testing. *J Natl Cancer Inst*. 110:855-862 (2018)

Schouten JP, McElgunn CJ, Waaijer R, Zwijnenburg D, Diepvens F, Pals G. Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res.* 30:e57 (2002)

Schrader KA, Hurlburt J, Kalloger SE, et al. Germline BRCA1 andBRCA2 mutations in ovarian cancer: utility of a histology-based referral strategy. *Obstet Gynecol*. 120:235-240 (2012)

Schuler N, Palm J, Schmitz S, Lorat Y, Rube CE. Increasing genomic instability during cancer therapy in a patient with Li-Fraumeni syndrome. *Clin Transl Radiat Oncol*. 7:71–8 (2017)

Siegel R, Ward E, Brawley O, Jemal A. Cancer statistics, 2011: the impact of eliminating socioeconomic and racial disparities on premature cancer deaths. *CA Cancer J Clin*. 61:212-236 (2011)

Siegel RL, Miller KD, Jemal A. Cancer statistics. *CA Cancer JClin*. 69:7-34 (2019)

Slavin TP, Van Tongeren LR, Behrendt CE, et al. Prospective Study ofCancer Genetic Variants: Variation in Rate of Reclassification by Ancestry. *J Natl Cancer Inst*. 110:1059-1066 (2018)

Smith TM, Lee MK, Szabo CI, Jerome N, McEuen M, Taylor M, Hood L, and King MC. Complete genomic sequence and analysis of 117 kb of human DNA containing the gene BRCA1. *Genome Research.* 6: 1029-1049 (1996)

So MK, Jeong TD, Lim W, et al. Reinterpretation of BRCA1 and BRCA2 variants of uncertain significance in patients with hereditary breast/ovarian cancer using the ACMG/AMP 2015 guidelines. Breast Cancer. 26(4):510-519 (2019)

Song H, Cicek MS, Dicks E, et al. The contribution of deleterious germline mutations in BRCA1, BRCA2 and the mismatch repair genes toovarian cancer in the population. *Hum Mol Genet*. 23:4703-4709 (2014)

Stevens KN, Vachon CM, Couch FJ. Genetic susceptibility to triple-negative breast cancer. *Cancer Res.* 73: 2025-2030 (2013)

Stoffel EM, Boland CR. Genetics and Genetic Testing in Hereditary Colorectal Cancer. *Gastroenterology*. 149(5):1191-1203.e2 (2015)

Talseth-Palmer B.A. The genetic basis of colonic adenomatous polyposis syndromes. *Hered Cancer Clin Pract.* 15, 5 (2017)

Tanakaya K. Current clinical topics of Lynch syndrome. *Int J Clin Oncol*. 24(9):1013-1019 (2019)

References

Thompson D, Antoniou AC, Jenkins M, et al. Two ATM variants andbreast cancer risk. *Hum Mutat*. 25:594-595 (2005)

Tischkowitz M, Balmaña J, Foulkes WD, et al. Management of individuals with germline variants in PALB2: a clinical practice resource of the American College of Medical Genetics and Genomics (ACMG). *Genet Med*. 23(8):1416-1423 (2021)

Tommiska J, Jansen L, Kilpivaara O, et al. ATM variants and cancer risk in breast cancer patients from Southern Finland. *BMC Cancer*. 6:209 (2006)

Tran SD, Pillemer SR, Dutra A, et al. Differentiation of human bonemarrow-derived cells into buccal epithelial cells in vivo: a molecular analytical study. *Lancet*. 361:1084-1088 (2003)

Tsai GJ, Rañola JMO, Smith C, et al. Outcomes of 92 patient-driven family studies for reclassification of variants of uncertain significance. *Genet Med*. 21(6):1435-1442 (2019)

Tsaousis GN, Papadopoulou E, Apessos A, Agiannitopoulos K, Pepe G, Kampouri S, Diamantopoulos N, Floros T, Iosifidou R, Katopodi O, Koumarianou A, Markopoulos C, Papazisis K, Venizelos V, Xanthakis I, Xepapadakis G, Banu E, Eniu DT, Negru S, Stanculeanu DL, Ungureanu A, Ozmen V, Tansan S, Tekinel M, Yalcin S, Nasioulas G. Analysis of hereditary cancer syndromes by using a panel of genes: novel and multiple pathogenic mutations. *BMC Cancer*. 19(1):535. (2019)

Tung N, Domchek SM, Stadler Z, et al. Counselling framework for moderate-penetrance cancer-susceptibility mutations. *Nat Rev Clin Oncol*. 13:581-588 (2016)

Turner SA, Rao SK, Morgan RH, Vnencak-Jones CL, Wiesner GL. The impact of variant classification on the clinical management of hereditary cancer syndromes. *Genet Med*. 21:426-430 (2018)

van den Broek AJ, van 't Veer LJ, Hooning MJ, et al. Impact of age atprimary breast cancer on contralateral breast cancer risk in BRCA1/2 mutation carriers. *J Clin Oncol.* 34:409-418 (2016)

van Marcke C, De Leener A, Berliere M, et al. Routine use of genepanel testing in hereditary breast cancer should be performed with caution. *Crit Rev Oncol Hematol*. 108:33-39 (2016)

Vail PJ, Morris B, van Kan A, et al. Comparison of locus-specific databases for BRCA1 and BRCA2 variants reveals disparity in variantclassification within and among databases. *J Community Genet*. 6:351-359 (2015)

Valdez JM, Nichols KE, Kesserwan C. Li-Fraumeni syndrome: a paradigm for the understanding of hereditary cancer predisposition. *Br J Haematol*. 176:539–52 (2017)

References

Vaz-Drago R, Custódio N, Carmo-Fonseca M. Deep intronic mutations and human disease. *Hum Genet*. 136(9):1093-1111 (2017)

Venkitaraman AR. Cancer susceptibility and the functions of BRCA1and BRCA2. *Cell*. 108:171-182 (2002)

Walsh T, Casadei S, Coats KH, et al. Spectrum of mutations in BRCA1, BRCA2, CHEK2, and TP53 in families at high risk of breastcancer. *JAMA*. 295:1379-1388 (2006)

Walsh T, Lee MK, Casadei S, et al. Detection of inherited mutations forbreast and ovarian cancer using genomic capture and massively parallel sequencing. *Proc Natl Acad Sci U S A*.107:12629-12633 (2010)

Wang, Y., Bernhardy, A.J., Nacson, J. *et al. BRCA1* intronic *Alu* elements drive gene rearrangements and PARP inhibitor resistance. *Nat Commun.* 10, 5661 (2019)

Weitzel JN, Lagos VI, Cullinane CA, et al. Limited family structure andBRCA gene mutation status in single cases of breast cancer. *JAMA.* 297:2587-2595 (2007)

Win AK, Reece JC, Dowty JG, Buchanan DD, Clendenning M, Rosty C, Southey MC, Young JP, Cleary SP, Kim H, Cotterchio M, Macrae FA, Tucker KM, Baron JA, Burnett T, Le Marchand L, Casey G, Haile RW, Newcomb PA, Thibodeau SN, Hopper JL, Gallinger S, Winship IM, Lindor NM, Jenkins MA. Risk of extracolonic cancers for people with biallelic and monoallelic mutations in MUTYH. *Int J Cancer*. 139(7):1557-63 (2016)

Xiang HP, Geng XP, Ge WW, Li H. Meta-analysis of CHEK2 1100delC variant and colorectal cancer susceptibility. Eur J Cancer. 47(17):2546-51 (2011)

Young SR, Pilarski RT, Donenberg T, et al. The prevalence of BRCA1 mutations among young women with triple-negative breast cancer. *BMC Cancer*. 9:86 (2009)

Zerdoumi Y, Lanos R, Raad S, Flaman JM, Bougeard G, Frebourg T, Tournier I. Germline TP53 mutations result into a constitutive defect of p53 DNA binding and transcriptional response to DNA damage. *Hum Mol Genet*. 26:2591–602 (2017)

Zhang B, Beeghly-Fadiel A, Long J, Zheng W. Genetic variants associated with breast-cancer risk: comprehensive research synopsis, meta-analysis, and epidemiological evidence. *Lancet Oncol*. 12(5):477-488 (2011)

Zhen DB, Rabe KG, Gallinger S, et al. BRCA1, BRCA2, PALB2, and CDKN2A mutations in familial pancreatic cancer: a PACGENE study. *Genet Med*. 17:569-577 (2015)

Zhu M, Chen X, Zhang H, Xiao N, Zhu C, He Q, Guo W, Cai Z, Shen H, Wang Y. AluYb8 insertion in the MUTYH gene and risk of early-onset breast and gastric cancers in the Chinese population. Asian Pac J Cancer Prev. 12(6):1451-5 (2011)

References

Zoure AA, Slaoui M, Bambara HA, Compaoré TR, Ouédraogo NLM, Mzibri ME, Attaleb M, Traoré SS, Simpore J, Bakri Y. BRCA1 c.68_69delAG (exon2), c.181T>G (exon5), c.798_799delTT and 943ins10 (exon11) mutations in Burkina Faso. *Journal of Public Health in Africa* (2018)

9. Web References

VarSome-The Human Genetics Community-2021:

https://varsome.com

American Cancer Society | Information and Resources about for Cancer: Breast, Colon, Lung, Prostate, Skin:

https://www.cancer.org

National Comprehensive Cancer Network:

https://www.nccn.org/professionals/physician_gls/default.aspx

GeneCards, The human gene database-2021:

https://www.genecards.org

Genetics Home Reference. Familial adenomatous polyposis:

http://ghr.nlm.nih.gov/condition/familial-adenomatous-polyposis

Clustal Omega:

https://www.ebi.ac.uk/Tools/msa/clustalo/

Translation of DNA sequence in protein sequence:

https://web.expasy.org/translate/

10. Acknowledgements

I would like to express my deepest gratitude and appreciation to my supervisor Prof. Nadia G. Ranzani for her helpful guidance through these three PhD years, for her continuous encouragement and interest in my work. At many stages in the course of this research project I benefited from her advice, particularly so when discussing my ideas and data. Her positive outlook and confidence in my research inspired me and gave me confidence. Her careful editing contributed enormously to the production of this thesis. Her enthusiasm and support over this period have been invaluable.

I want to acknowledge and extend my warmest thanks to the responsible of Clinical Genomics Unit, Dr. Loris Bernard, for giving me the opportunity to work in his laboratory and undertake this research, providing me with additional support and assistance. His guidance and advise carried me through all the stages of my project.

I want to thank Dr. Elena Belloni for her constant scientific support and ideas.

Many thanks go also to all my lab colleagues for the scientific interaction.

The most important help I have received throughout my life has been the love and encouragement of my parents, to whom this thesis is dedicated.