SAPIENZA

UNIVERSITA' DI ROMA

DOTTORATO DI RICERCA IN BIOLOGIA UMANA E GENETICA MEDICA

XXXV CICLO

DNA repair genes in cancer predisposition: detection of germline pathogenic variants by multigene panel testing

DOTTORANDO
Dott.ssa Soha Sadeghi

DOCENTE GUIDA
Prof. Antonio Pizzuti

CO-DOCENTE GUIDA
Dott.ssa Maria Piane

COORDINATORE DEL DOTTORATO
Prof.ssa Laura Stronati

ANNO ACCADEMICO 2022-2023

ABSTRACT

The 5 to 10% diagnosed cancers are linked to an inherited faulty gene. Mutations in distinct DNA repair systems elevate the susceptibility to various cancer types and germline pathogenic (P) variants in DNA damage repair (DDR) genes BRCA1 and BRCA2 explain only 10-20% of these cases. Currently, new DDR genes have been related to of Breast, Ovarian, colorectal and endometrial cancer, but the prevalence of pathogenic variants remains to be explored.

The purpose of this study was to investigate the spectrum and the prevalence of pathogenic variants in DDR pathway genes other than BRCA1/2 and to correlate the genotype with the clinical phenotype.

A cohort of 416 patients (298 cases were non-BRCA) was analyzed by next-generation sequencing using a multigene panel of the 28 DDR pathways genes related to Breast, Ovarian, colorectal and endometrial cancer. 41 of 416 affected individual were diagnosed with Lynch syndrome. 213 unique variants in 27 of 28 analyzed genes were found, 37 classified as likely pathogenic/ pathogenic and 177 as variants of unknown significance. 10 of 37 LP/P variants were discovered in 10 patients with Lynch syndrome.

It was observed a high incidence of deleterious variants in the *ATM*, *MUTYH*, *CHEK2* and *MSH6* gene. These results support the clinical utility of multigene panel to increase the detection of P/LP carriers and to identify new actionable pathogenic gene variants useful for preventive and therapeutic approaches.

Keywords: Cancer; Lynch syndrome; DDR; BRCA1/2; HBOC

b

LIST OF PUBLICATIONS

1. Genomic Breakpoints' Characterization of a Large CHEK2 Duplication in an Italian Family with Hereditary Breast Cancer. Aldo Germani,Daniele Guadagnolo, Valentina Salvati, Caterina Micolonghi, Rita Mancini,Gioia Mastromoro, **Soha Sadeghi**, Simona Petrucci, Antonio Pizzuti and Maria Piane. Diagnostics 2022; https://doi.org/10.3390/diagnostics12071520 (PMID: 35885426).

2. Heterozygous pathogenic nonsense ATM variant resulting in unusually high gastric cancer (GC) susceptibility. Daniele Guadagnolo, Gioia Mastromoro, Enrica Marchionni, Aldo Germani, Fabio Libi, **Soha Sadeghi**, Camilla Savio, Simona Petrucci, Maria Piane, Antonio Pizzuti https://doi.org/10.21203/rs.3.rs-1003666/v1 (researchsquare, 28, 10, 2021)

3. A new SMAD4 splice site variant in a three-generation Italian family with juvenile polyposis syndrome. Caterina Micolonghi , Maria Piane , Aldo Germani , Soha Sadeghi , Fabio Libi , Camilla Savio, Marco Fabiani, Rita Mancini, Danilo Ranieri, Vito Domenico Corleto, Pasquale Parisi, Vincenzo Visco, Giovanni Di Nardo, Simona Petrucci, Diagnostics 2022; (Accepted)

4. Long QTc in Hypertrophic Cardiomyopathy: a consequence of myocardial hypertrophy or a distinct genetic disease? C.Micolonghi, F.Cava, M.B. Musumeci, C. Savio, M. Fabiani, G.Tini, L.Alesi, L. Angileri, I, De Santis, A. Germani, F. Libi. C. Rossi, **S, Sadeghi**, V. Visco, A, Pizzuti, C. Autore, M.Piane, S.D. Rubattu, S.Petrucci. XXV CONGRESSO NAZIONALE SIGU 2022

5. Genotype-phenotype correlation in early and late -onest Hypertrophic Cardiomyopathy Italian patients. M. Fabiani, G. Tini, C.Micolonghi, M.B. Musumeci, C. Savio, F. Aurora, L. Alesi, L. Angileri, I, De Santis, A. Germani, F. Libi. C. Rossi, **S, Sadeghi**, V. Visco, A, Pizzuti, C. Autore, S.Petrucci. S.D. Rubattu, M.Piane, XXV CONGRESSO NAZIONALE SIGU 2022

TABLE OF CONTENTS

Chapter1

Chapter 7

List of the Figure

List of the Tables

ACRONYMS and ABBREVIATIONS

- AQ: Alignment Quality
- BAI: BAM index
- BAM: BinaryAlignment/MAP
- BER: Base ExcissionRepair
- HBOC: Hereditary Breast and Ovarian Cancer
- HR: HomologousRecombination
- MMR: Mismatch Repair
- NER: Nucleotide ExcissionRepair
- NGS: Next Generation Sequencing
- NHEJ: Non Homologous-End Joining
- PGM: Personal Genome Machine
- TN: Triplo Negativo
- TNBC: Triple Negative Breast Cancer
- TNM: TumorNoduleMetastasis
- TS: Torrent Suite
- VCF: Variant Call Format
- VUS: Variant of UnknownSignificance

Chapter 1

Introduction

Chapter 1

Introduction

1.1 DNA Repair Pathways and Mechanisms

DNA repair is a phenomenal multi-pathway, the multi-enzyme system required to ensure the integrity of the cellular genome from the inherent instability of DNA, the natural limitations on the accuracy of DNA synthesis, and the challenge of the environment(1-3).

Continuously, the DNA repair machinery scans the genome and maintains genome integrity by mending or removing any detected damage (2). regarding the type and the severity of DNA damage and also, the cell cycle status, the DNA repair machinery utilizes several different pathways to restitute the genome to its original state. When DNA cannot be repaired due to harm or other circumstances, the repair machinery attempts to minimize the damage to ensure cell viability(1, 3) (Figure 1-1).

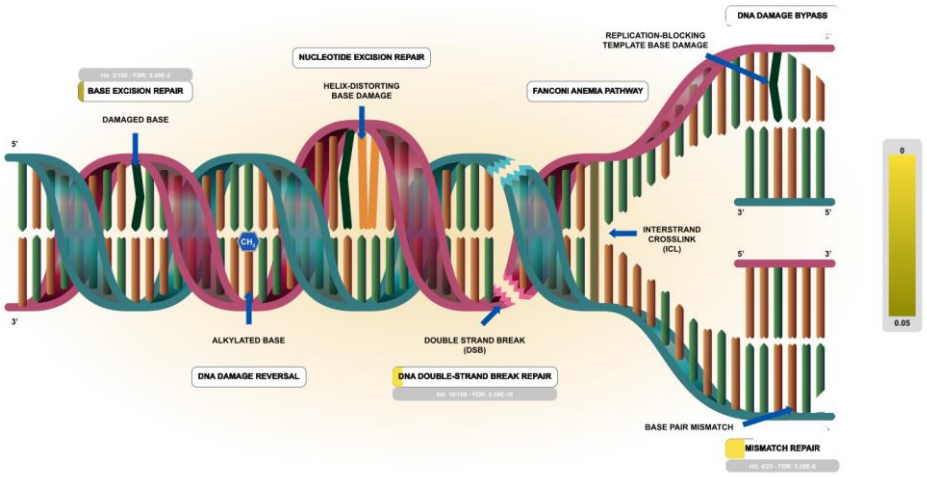

Figure 1-1 : DNA repair pathway: DNA damage bypass, DNA damage reversal, base excision repair, nucleotide excision repair, mismatch repair, repair of double-strand breaks, and repair of interstrand crosslinks (Fanconi anemia pathway)[1].

Accumulation of DNA alterations that are the result of cumulative its damage due to a reduction in mammalian DNA-repair capacity is associated with cellular senescence, aging, and cancer. In addition, the underlying cause of a large number of familial cancer syndromes, such as Fanconi anemia, xeroderma pigmentosum, Nijmegen breakage syndrome, and Lynch syndrome, is germline mutations in DNA repair genes(1, 4-6).

[1]- Https://reactome.org/

cancer-causing genetic changes resulting in alterations, loss, amplification, or changes in expression of genes important for normal cellular functions and growth properties, including proto-oncogenes and tumor suppressor genes, accumulate slowly as a person ages, leading to a higher risk of cancer later in life (7).

The cancer cells are be able to tolerate increased amounts of unrepaired DNA damage associated with genomic instability, and inactivate DNA damage–inducible signaling and checkpoint pathways. In humans, it has been estimated that up to 100,000 spontaneous DNA lesions are generated daily per cell by environmental agents and endogenous processes, such as transcription and replication of the DNA, and these mutations can lead to the development of cancer. Therefore, DNA repair process and the DNA damage response are pivotal for the basic processes of transcription and replication necessary for cellular survival, maintaining genomic stability and avoiding the development of malignancies(7, 8) (figure 1-2).

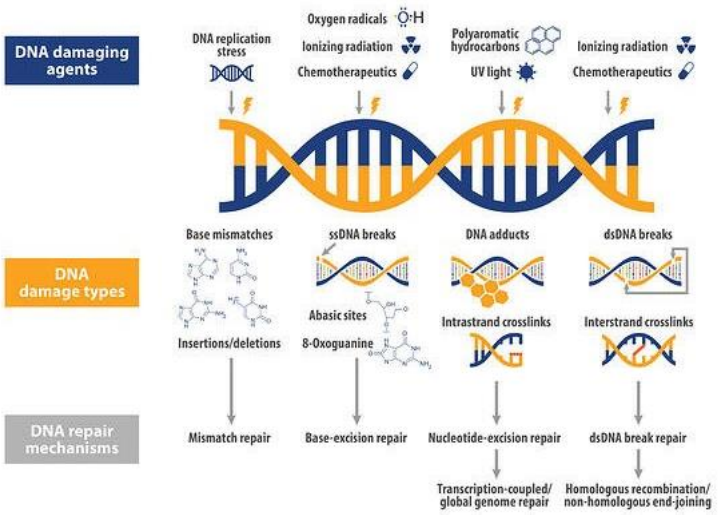

Figure 1-2: Cells constantly deal with damage to their DNA that can originate from endogenous processes, such as DNA replication stress, or exogenous exposures such as ionizing radiation and chemotherapy drugs. DNA damaging agents give rise to different types of DNA damage, and failure to repair the damage can result in a number of possible devastating consequences to the cell, including genetic instability and accumulation of mutations promoting tumorigenesis.[1]

[1] - Https://blog.crownbio.com/dna-damage-response

Chapter 1

A complex set of cellular surveillance and repair mechanisms has evolved to reverse or limit potentially deleterious DNA damage. Some of these DNA repair systems are so important that life cannot be sustained without them. Cells present different mechanisms of DNA repair depending on the type of DNA lesion involved(9-11).

Generally, there are seven main pathways employed in human DNA repair included :

1-DNA damage bypass[1]

2-DNA damage reversal[2]

3-base excision repair[3]

4-nucleotide excision repair[4]

5-mismatch repair[5]

6-repair of double-strand breaks [6]

7-repair of interstrand crosslinks (Fanconi anemia pathway) (figure 1-3).

DNA repair pathways are intimately associated with other cellular processes such as DNA replication, DNA recombination, cell cycle checkpoint arrest, and apoptosis. The choice of the repair mechanism used by the cell to repair its DNA is determined by the type of lesion and the position of the cell in the cell cycle(9-11).

Single-strand breaks are repaired by BER (9), bulkier single-strand lesions that distort the DNA helical structure are repaired by NER (10), and misincorporation of nucleotides resulting in mismatches in the DNA sequence are repaired by MMR (11). DSBR can be restored by two different pathways: non homologous end joining[7] that promotes the potentially inaccurate relegation of DSBs, and homologous recombination [8] that provides an error free mechanism of repair(9-11) (figure 1-3).

[1] - DDB
[2] - DDR
[3] - BER
[4] - NER
[5] - MMR
[6] - DSBR
[7]- NHEJ
[8] - HR

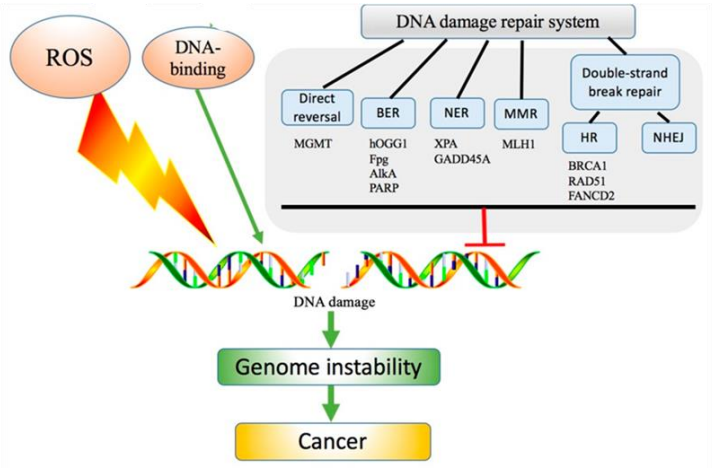

Figure 1-3: DNA repair mechanisms. Environmental DNA-damaging agents include the ultraviolet[1] component of sunlight, which generates cyclobutane pyrimidine dimers and oxidative base damage; ionizing radiation, which produces clusters of ROS that create double-strand DNA breaks; and base-damaging chemicals such as aflatoxins, benzo(a)pyrene, methyl chloride, and nitrosamines, which alter or destroy base-pairing capacity.[2]

An increasing number of human hereditary diseases that are characterized by severe developmental problems or a predisposition to cancer have been linked to deficiencies in DNA repair, as example: xeroderma pigmentosum, Cockayne syndrome, trichothiodystrophy, Werner syndrome, and possibly Bloom syndrome. Premature aging in these syndromes may be linked to the aging process through the corresponding genes' roles in transcription(12, 13); it may also be the result of defective NER, attributed to impaired interaction of the gene products with other DNA-repair or -replication proteins, or their failure to prevent the stalling of replication forks at sites of DNA damage(14, 15). In the case of Werner syndrome, the gene product's putative role as a MMR exonuclease could also serve as a link to aging(15). Altered function of BER genes has also been implicated in accelerated aging.

1.1.1 DNA damage bypass pathway

Because DNA damage has the potential to inhibit and/or alter the fidelity of replication and transcription, there is an essential need for diverse and highly precise repair processes(16). In addition, bypass mechanisms allow unrepaired damage to be tolerated if encountered during replication. The DNA damage bypass pathway does not eliminate the damage but instead allows translesion DNA synthesis[3]

[1] - UV
[2] - https://blog.crownbio.com/dna-damage-response
[3] - TLS

using a damaged template strand. Translesion synthesis authorizes cells to complete DNA replication, postponing the repair until cell division is finished. DNA polymerases that partake in translesion synthesis are error-prone, mostly introducing base substitutions and/or small insertions and deletions. there is significant overlap between the various repair and bypass pathways in terms of the cognate lesions that each can deal with. This functional redundancy is partially a reflection of the very high load of endogenous DNA damage and underscores the importance of these pathways in the maintenance of genome stability(2, 16).

1.1.2 Base Excision Repair (BER)

The hydrolysis and exposure to reactive metabolites that cause oxidation and alkylation of DNA is a pivotal source of DNA damage to cellular genomes. The repair system mainly has a role in in identifying and removing such lesions, as well as in dealing with the spontaneous loss of purines from DNA, via the base excision repair[1] pathway. Therefore, the BER pathway includes a number of DNA glycosylases that cleave a vast array of damaged bases from the DNA sugar-phosphate backbone.

BER is initiated by one of a set of lesion-specific glycosylases that recognize the altered or inappropriate base. DNA glycosylases produce a DNA strand with an abasic site[2]. The abasic site is processed by DNA endonucleases, DNA polymerases, and DNA ligases, the choice of which depends on the cell cycle stage, the identity of the participating DNA glycosylase, and the presence of any additional damage. Base excision repair yields error-free DNA molecules(17, 18).

In some cases, bifunctional glycosylases, which also cleave the phosphodiester bond adjacent to the damaged base, preclude the need for endonuclease activity of APE1[3] and instead require polynucleotide kinase 3′-phosphatase for single-strand break end-processing. In the short-patch pathway, DNA polymerase (POLB) displaces the AP site and adds a nucleotide. Last, ligase (LIG3) forms a phosphodiester bond to complete repair. In the long-patch pathway, DNA polymerase (POLB, POLD, or POLE) displaces and adds more than one nucleotide, flap structure-specific endonuclease (FEN1) removes the displaced nucleotides, and ligase (LIG1) completes repair. Proliferating cell nuclear antigen (PCNA) is

[1] - BER
[2] - AP site (apurinic/apyrimidinic site)
[3] - AP-endonuclease

required for POLD function and acts as a scaffolding protein. Poly (ADP-ribose) polymerase (PARP) binds to and recruits essential mediators to single-strand break intermediates. The most direct evidence for a role for BER in cancer comes from the discovery that germline mutations in the *MUTYH* gene, involved in processing 8-oxoG lesions, is associated with recessive inheritance of a predisposition to develop multiple colorectal adenomas (polyposis) and colon cancers(17, 18) (figure1-4).

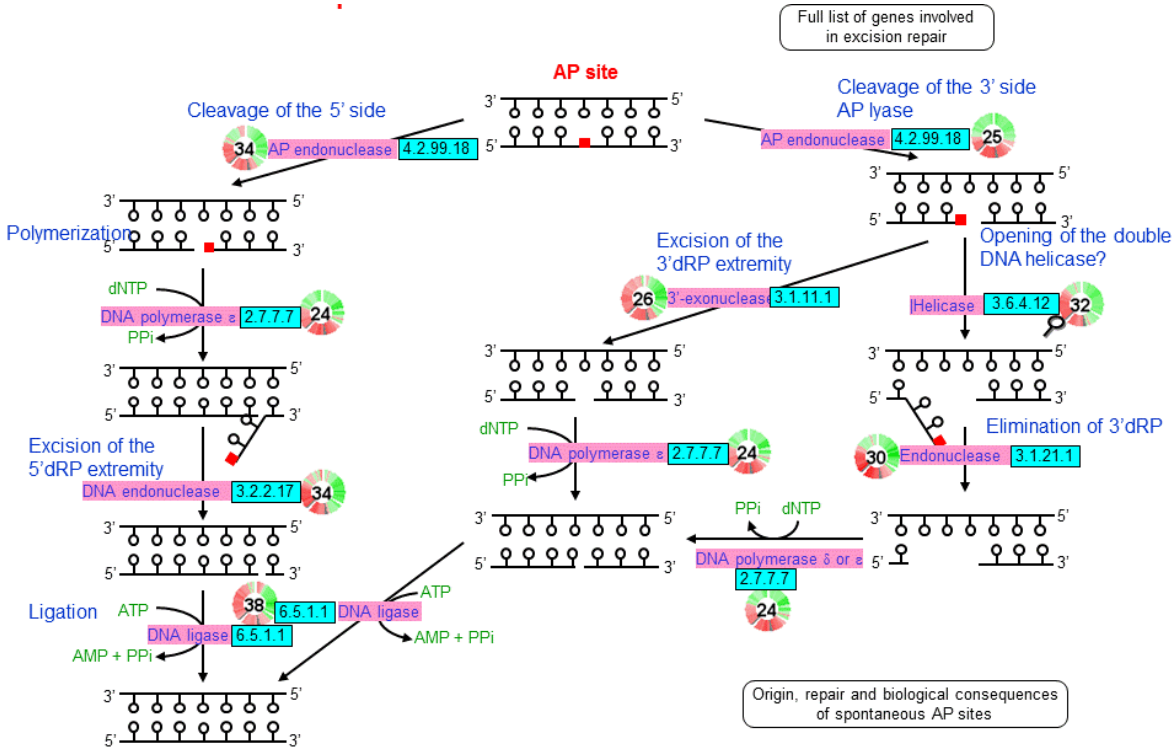

Figure 1-4: Base excision repair of AP site(18).

1.1.3 Nucleotide Excision Repair[1]

One of the most important mechanisms of DNA repair is in which the damaged or incorrect part of a DNA strand is Recognized of the damage leads to the removal of a short single-stranded DNA segment that contains the lesion. The undamaged single-stranded DNA remains and DNA polymerase uses it as a template to synthesize a short complementary sequence. The nucleotide excision repair pathway is involved in the removal of bulky lesions that cause distortion of the DNA double helix. Final ligation to complete NER and form a double-stranded DNA is carried out by DNA ligase. NER can be divided into two sub pathways: global genomic NER (GG-NER or GGR) and transcription-coupled

[1]- NER

NER (TC-NER or TCR). The two sub pathways differ in how they recognize DNA damage but they share the same process for lesion incision, repair, and ligation(19, 20).

NER functions to remove a range of types of lesions, including bulky base adducts of chemical carcinogens, intrastrain cross-links[1], and UV-induced cyclobutene pyrimidine dimers[2] and 6-4 photoproducts. NER proteins excise the oligonucleotide that includes the lesion from the affected DNA strand, which is followed by gap-filling DNA synthesis and ligation of the repaired DNA molecule(19, 20).

The steps for NER in order are:

1. recognition of the damaged site
2. incision of the damaged DNA strand near the site of the defect
3. removal of a stretch of the affected strand containing the lesion repair replication to replace the excised region with a corresponding stretch of normal nucleotides with use of the complementary strand as a template
4. ligation to join the repair patch at its 3′end to the contiguous parental DNA strand(20).

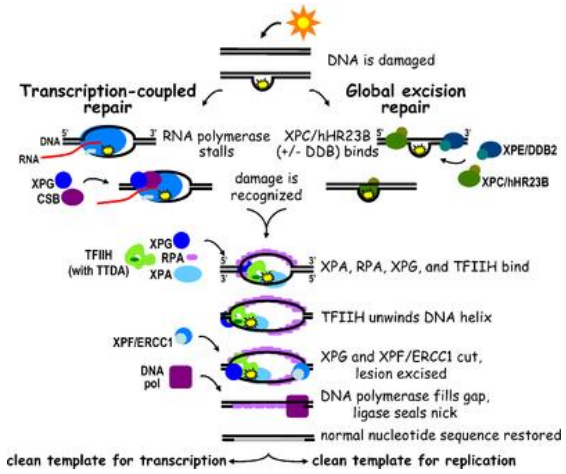

Figure 1-5 nucleotide excision repair (21). Diagram of both the TC-NER and GG-NER pathways. The two pathways differ only in initial DNA damage recognition.

The importance of NER is evidenced by the severe human diseases that result from in-born genetic mutations of NER proteins. Xeroderma pigmentosum and Cockayne's syndrome are two examples of NER-associated diseases.

[1] - ICLs
[2] - CPDs

The majority of human NER genes have been identified and cloned, and The importance of NER was evidenced by the severe human diseases that result from in-born genetic mutations of NER proteins. Xeroderma pigmentosum and Cockayne's syndrome are two examples of NER-associated diseases(21-23).

The p53 gene, central to maintaining genomic stability in human cells, is required for efficient GGR of UV light– and carcinogen–induced DNA damage, and functions as a DNA damage–activated transcription factor that directly regulates the expression of several NER damage recognition genes. Similarly, several other important cancer-related genes have been shown to transcriptionally regulate the DNA damage recognition NER genes XPC and DDB2, including BRCA1 and E2F1(23) (figure1-5).

1.1.4 Double-Strand Break Repair[1]

Double strand breaks (DSBs) in the DNA can be directly generated by some DNA damaging agents, such as X-rays and reactive oxygen species[2]. DSBs can be repaired via a highly accurate homologous recombination repair[3] pathway, or through error-prone nonhomologous end joining[4], single strand annealing[5], and microhomology-mediated end joining[6] pathways(24, 25).

Diseases of DNA double-strand break repair[7] are caused by mutations in genes involved in the repair of double-strand breaks, one of the most cytotoxic types of DNA damage. Even a single occurrence in the entire genome is sufficient to signal cell cycle checkpoints that unrepaired DSBs lead to cell death, cellular senescence, or malignant transformation(26, 27).

DSBs also cause some problems during mitosis due to intact chromosomes being a prerequisite for proper chromosome segregation during cell division. Regarding this, these lesions often induce multifold sorts of chromosomal aberrations, including deletions, aneuploidy, and chromosomal translocations all of which are associated with carcinogenesis. Chromosomal instability also is characteristic of breast tumors that harbor BRCA2 mutations, probably because of defective recombination-mediated DSB repair(27).

[1] - DSBs
[2] - ROS
[3] - HRR
[4] - NHEJ
[5] - SSA
[6] - MMEJ
[7] - DSBR

Single strand breaks (SSBs) in the DNA, generated either by DNA damaging agents or as intermediates of DNA repair pathways such as BER, are converted into DSBs if the repair is not complete prior to DNA replication. Simultaneous inhibition of DSB repair and BER through cancer mutations and anti-cancer drugs, respectively, is synthetic lethal in at least some cancer settings and is a promising new therapeutic strategy(27).

Recent research has shown some of the protein kinases involved in signaling cellular processes in response to DSBs. Most of these have been identified to be defective in cancer-prone disorders exhibiting genomic instability, such as the ATM protein[40] and the CHEK2 protein kinase(26, 27).

The p53 tumor suppressor gene is the most important aim of these kinase activities. when the aforementioned protein phosphorylated, is activated and involved in G1 arrest and apoptosis after ionizing radiation and, when found mutated in the germline, causes the Li-Fraumeni cancer susceptibility syndrome. an array of other enzymes involved in the actual DNA transactions necessary for DSB repair have been found in cancer-prone disorders, including MRE11 (AT-like disorder), NBS1 (Nijmegen breakage syndrome), BRCA1 and BRCA2 (breast-ovarian cancer syndrome), and the RecQ-like helicases (Werner, Bloom, and Rothmund-Thomson syndromes) (figure 1-6)(24, 28).

variety of research suggested that BRCA1 and BRCA2 are centrally involved in various aspects of DNA repair and DNA damage response pathways. As an example, BRCA1 is phosphorylated after exposure to DNA-damaging agents by ATM, ATR, and Chek2; associates with an array of DNA repair proteins including ATM, RAD51, and the RAD50-MRE11-NBS1 protein complex following DNA damage; and localizes to nuclear foci with these proteins after treatment with ionizing radiation(29). The association of BRCA1 with RAD51, an enzyme involved in the coordination of recombination, suggests its involvement in DSB repair and implicates BRCA1 in HR(30).

Other evidneces suggested that BRCA1 may regulate cellular processes via transcriptional coactivation. BRCA1 through transcription regulate the NER genes XPC and DDB2 and affect GGR[1] of UV- and cisplatin-induced DNA damage and

[1] - Global Genome Repair

to transcriptionally regulate BER genes and affect BER of oxidative DNA damage(31).

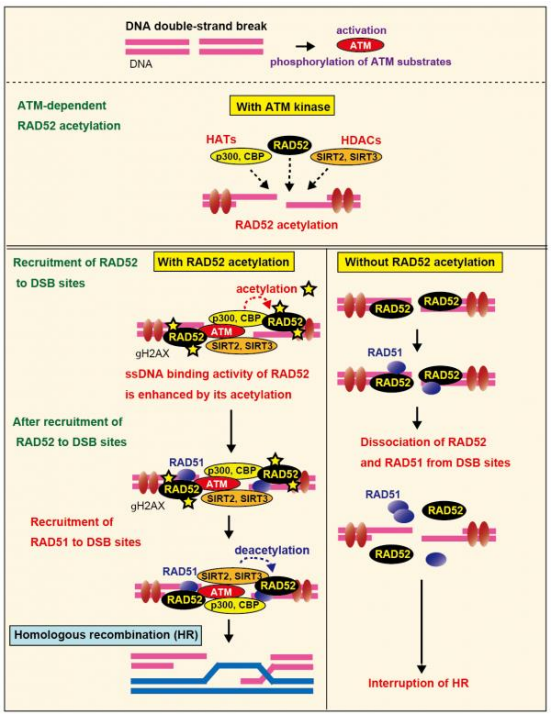

Figure 1-6 : double-strand break repair pathways in mammalian cell(24). Homologous recombination (HR) is a DSB repair system in which a central player, RAD51, functions with several proteins, including RAD52. DSBs activate the DNA damage response signaling network, in which the ataxia telangiectasia mutated (ATM) protein plays a chief role, by phosphorylating numerous target proteins. As compared to phosphorylated proteins, relatively few acetylated proteins have been functionally characterized in DNA repair.

1.1.5 Mismatch Repair[1]

MMR is another example of an excision repair mechanism that uses a similar strategy to maintain genomics. Mismatch repair proteins recognize mismatched base pairs or small insertion or deletion loops during DNA replication and correct erroneous base pairing by excising mismatched nucleotides exclusively from the nascent DNA strand, leaving the template strand intact(2).

In addition, this repair mechanism is able to deal with small loops of single-stranded DNA at sites of insertions or deletions in the duplex DNA structure. The importance of this mechanism in maintaining genetic stability is shown by the observation that its absence results in significant growth in the frequency of spontaneously occurring mutations, particularly in microsatellite sequences of highly repetitive DNA. Some of these spontaneous mutations arise from mistakes introduced during DNA replication, spite the operation of a “proofreading” system that also helps to ensure high fidelity of replication(32).

[1] - MMR

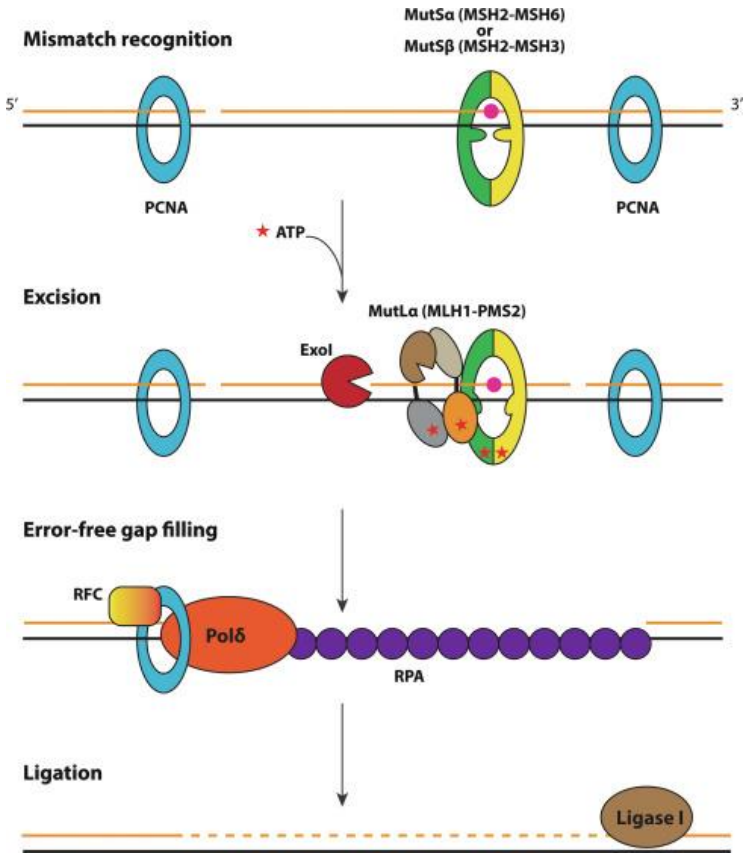

Figure 1-7: DNA mismatch repair in mammals(33)

Defects in mammalian DNA mismatch repair[1] genes (*MLH1, PMS2, MSH2, and MSH6*) result in microsatellite instability[2] and reduced fidelity during replication and repair steps. Defective variants of MMR genes are associated with sporadic cancers with hypermutation phenotypes as well as hereditary cancer syndromes such as Lynch syndrome (hereditary non-polyposis colorectal cancer) and constitutional mismatch repair deficiency syndrome[3].

MSI is an important predictor of sensitivity to cancer immunotherapy as the high mutational burden renders MSI tumors immunogenic and sensitive to programmed cell death-1[4] immune checkpoint inhibitors (34-37).

As with other mechanisms of excision repair, four principal steps are required for MMR (figure 1-8):

(1) mismatch recognition

(2) recruitment of additional MMR factors

[1] - MMR
[2] - MSI
[3] - CMMRD
[4] - PD-1

(3) identification of the newly synthesized DNA strand containing the mismatched nucleotides, followed by their excision

(4) resynthesis of the excised tract and ligation.

In humans, heterodimers of the MSH2/6 proteins recognize single base-pair mismatches and short insertion-deletion loops, whereas MSH2/3 dimers recognize longer loops. Heterodimeric complexes of MLH1/PMS2 and MLH1/PMS1 interact with the MSH complexes and replication factors (figure 1-7), for strand discrimination and DNA excision. Similar to NER and BER, additional proteins are then recruited for repair replication based on the original DNA template (35, 37)(figure 1-8).

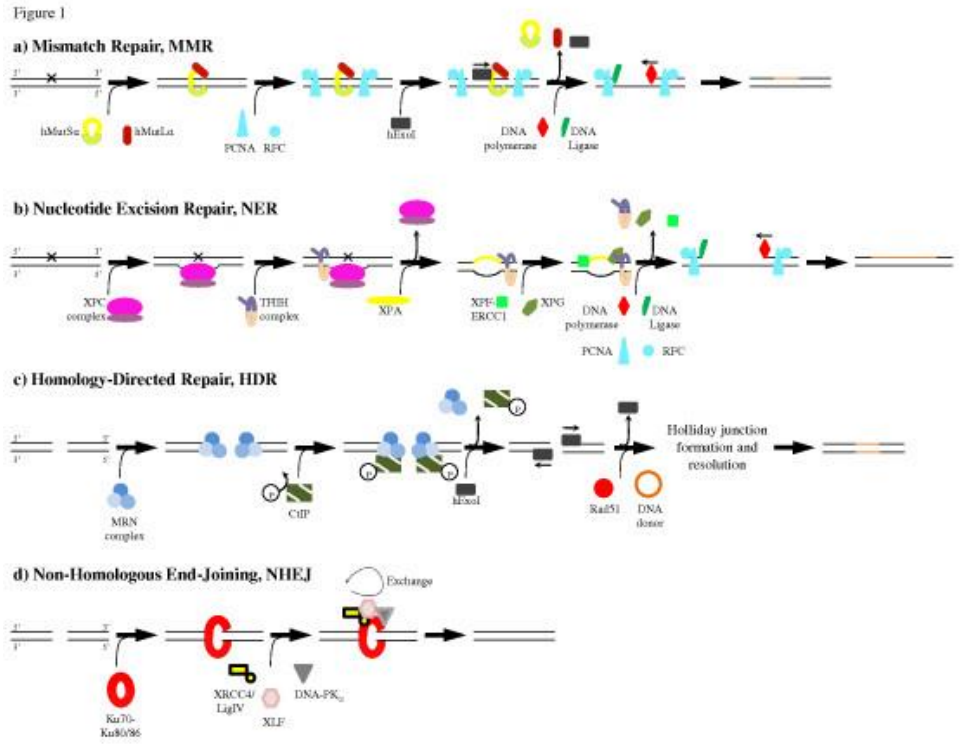

Figure 1-8 : Components involved in mammalian repair pathways. A: In mismatch repair (MMR), hMutSα recognizes the DNA damage whereby hMutLα is recruited resulting in nicks on either side of the mismatch. Human exonuclease I (hExoI, 5'→3' activity) excises the mismatch and its flanking sequences after which DNA polymerase (3'→5' activity), along with PCNA and RFC, re-synthesizes a new DNA strand. B: In nucleotide excision repair (NER), the XPC complex recognizes the DNA damage causing the recruitment of the TFIIH complex, which unwinds the DNA to an open complex. XPA binds the damaged DNA strand after which endonucleases, XPG and XPF-ERCC1, excise the mismatch and DNA polymerase, with PCNA and RFC re-synthesizes the DNA strand. C: In homology-directed repair (HDR), the DSB is bound by the MRN complex recruiting CtIP and hExo, the latter of which excise nucleotides surrounding the break. Rad51 initiates homology search and when a homologous DNA donor is found, the DSB is repaired through Holliday junction formation and resolution. D: In non-homologous end-joining (NHEJ), the Ku complex recognizes the DSB leading to a simultaneous recruitment of DNA-PKCS, XRCC4:LigIV and XLF. The exchange of these factors drives the ligation of the non-homologous ends. Artemis nuclease, DNA polymerases μ and λ and other protein factors can be involved if the DNA ends are not directly compatible. See text for further details(34).

1.2 DNA repair Genes

The development of cancer is the result of accumulated mutations in oncogenes and tumor suppressor genes. These genes are involved in cell proliferation, growth, differentiation, survival, apoptosis, and DNA repair(figure1-9) (38, 39). The dysregulation of repair genes can be associated with significant, damaging health effects, which can include an increased prevalence of birth defects, an enhancement of cancer risk, and a rate of ageing. Increased knowledge about the human genome and advances in genotyping technology have made possible genome-wide association studies[1] of human diseases(38, 39).

well over 150 genes directly involved in DNA repair have now been identified in humans, and their cDNA sequence established. These genes function in various sets of pathways that involve the recognition and removal of DNA lesions, tolerance to DNA damage, and protection from errors of incorporation made during DNA replication or DNA repair (aforementioned). Additional genes indirectly affect DNA repair, by regulating the cell cycle, ostensibly to provide an opportunity for repair or to direct the cell to apoptosis. Recently investigation of the individual genes' mutation in the appropriate clinical setting where there is a high index of suspicion for a specific mutated gene or syndrome, is recommended(38, 39)(figure1-9).

Next-generation sequencing[2] offers a new venue for risk assessment. At the present time, there are clear clinical guidelines European Society for Medical Oncology[3] and National Comprehensive Cancer Network[4] for individuals with a mutation in a high-penetrance gene, the current guidelines provide concrete direction for the management of these patients(38, 39).

A great number of familial cancers is due to highly penetrant, but rare genes. In recent years, additional rare, moderate-penetrance genes and common, low-penetrance alleles have also been identified by the increase of knowledge in GWAS.

[1] - GWAS
[2] - NGS
[3] - ESMO
[4] - NCCN

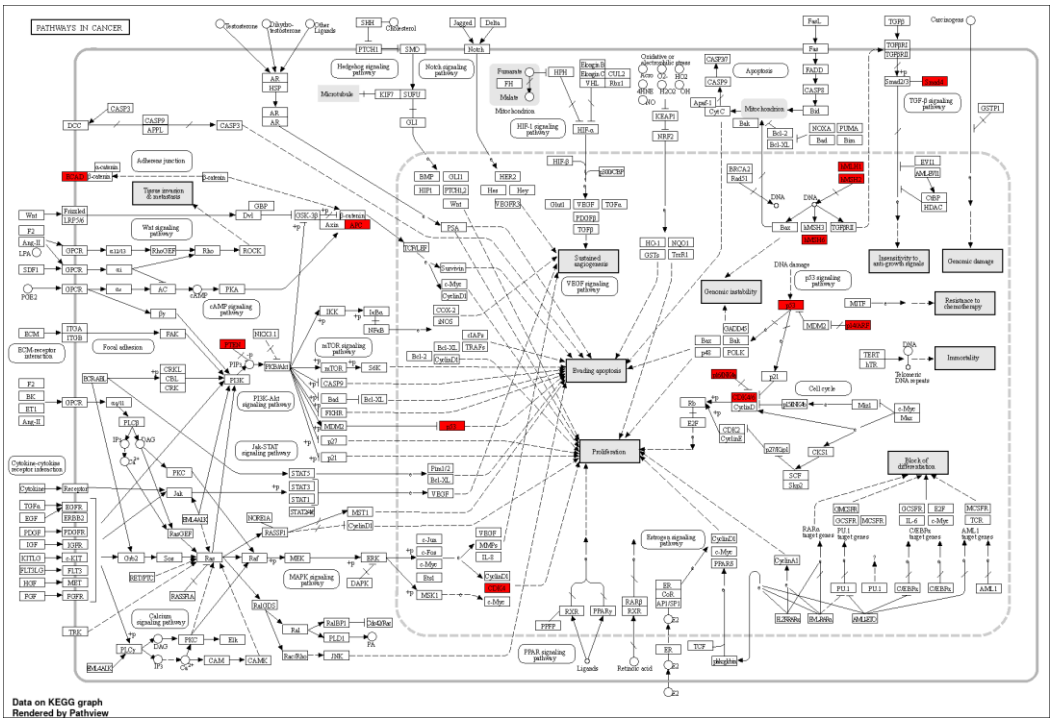

Figure 1-9 : KEGG pathway in cancer . The disgram shows the role of the DNA repair genes in cancer[1].

Depending on the relative risk that these genes give to tumor development, the following are distinguished(figure1-9):

A. high penetrance genes: the *BRCA1*, *BRCA2*, *TP53*, *PTEN*, *STK11*, and *CDH1* genes, whose mutations are associated with a relative risk ≥ 5;
B. genes with moderate penetrance: *CHEK2*, *ATM*, *PALB2*, and *BRIP1* genes, whose mutations are associated with a relative risk ≥ 1.5 and <5;
C. low penetrance genes: *CASP8*, *FGFR2*, *MAP3K1*, *LSP1*, *TNRC9*, *H19* and others, whose mutations are associated with a relative risk ≥ 1.01 and <1.5

Prediction models indicated that there are unlikely to be added yet-to-be-identified high-penetrance genes. Investigation of common, low-penetrance alleles contributing to risk in a polygenic fashion has yielded a small number of suggestive single-nucleotide polymorphisms[2], but the contributive risk of an individual SNP is quite small(38, 39).

While a small number of low/moderate -penetrance alleles may contribute to risk in a polygenic fashion, this is likely to be relevant to a majority of cases and their identification should be considered routine practice.

[1] - Http://bioinformatics.sdstate.edu/go
[2] - SNPs

A precise and comprehensive family history of cancer is essential to identify individuals who may have a risk for inherited cancer and should include 3-generation family history with information on both maternal and paternal lineages. The focus should be on both the individuals with malignancies and also, family members without a personal history of cancer(39).

The contribution of mutations in other genes to the burden of breast or ovarian cancer is shown below:

1.2.1 *BRCA1* and *BRCA2* genes

BRCA1 and *BRCA2* genes were known as tumour suppressor genes. They prevent cells from growing and dividing too rapidly or in an uncontrolled way. *BRCA1* and *BRCA2* work as an important regulator of cell-cycle "checkpoint control" mechanisms involving cell-cycle arrest, cell death (apoptosis) and DNA repair. Levels of *BRCA1* increase during DNA synthesis and mitosis (figure 1-10).

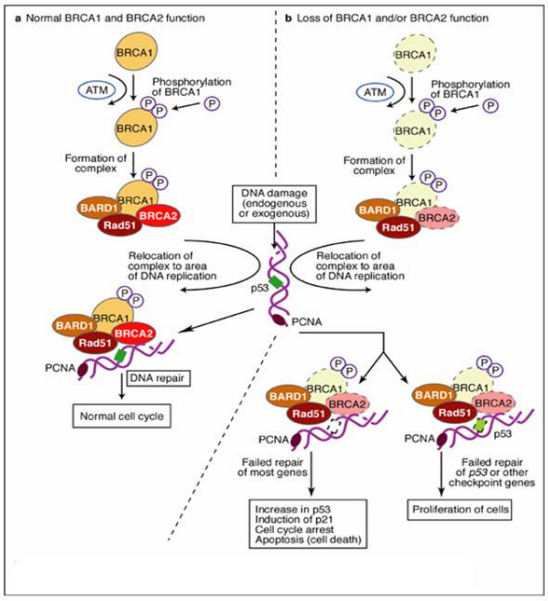

Figure 1-10 : The role of BRCA1 and BRCA2 protein in DNA repair[1]

The term definition of BRCAness is a defect in homologous recombination repair[2], mimicking BRCA1 or BRCA2 loss. It shares with these hereditary cancers. In turn, BRCA-deficient cells utilize error-prone DNA-repair pathways, causing increased genomic instability, which may be responsible for their sensitivity to DNA damaging agents and poly[3]-ribose polymerase inhibitors[4] (40, 41)(figure 1-11).

[1] - https://slideplayer.com/slide/5109761/
[2] - HRR
[3] - ADP
[4] - PARPis

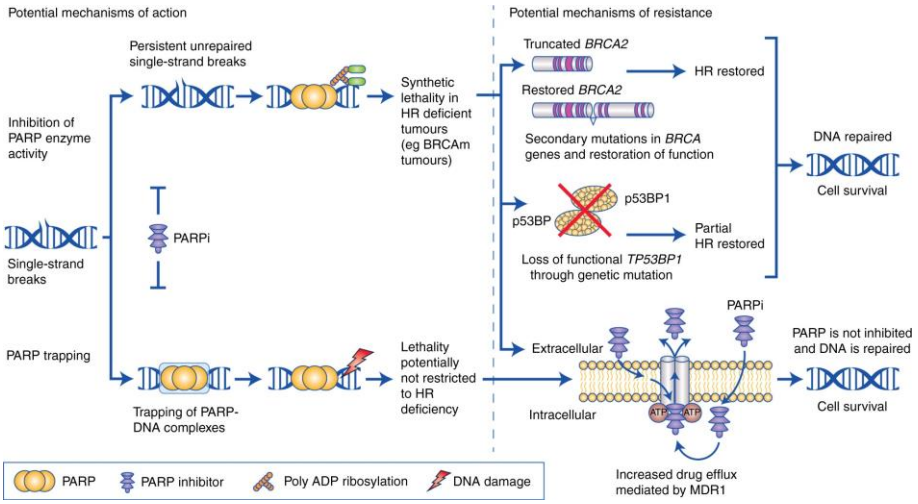

Figure 1-11 : PARP inhibitors: Some possible mechanisms of action and resistance. The left panel illustrates two possible mechanisms of action of PARPi. Upper pathway: Inhibition of PARP enzyme activity or catalytic inhibition interferes with the repair of single-strand breaks, leading to stalled DNA replication forks that requires HR repair. In HR-deficient tumours, such as those with BRCAm, PARP inhibition results in synthetic lethality. Lower pathway: PARP trapping refers to trapping of PARP proteins on DNA, which also leads to replication fork damage, but because this pathway utilises additional repair mechanisms, it is not restricted to tumours with HR deficiency. The right panel illustrates three possible mechanisms of resistance to PARPi. These include: (1) secondary mutations in BRCA genes that restore BRCA function and HR; (2) somatic mutation of TP53BP1, causing partial restoration of HR; and (3) increased PARPi efflux mediated by MDR1/P-glycoprotein 1, preventing the drugs from acting at the appropriate sites. The first two mechanisms of resistance restore HR and apply to PARP catalytic inhibition in HR-deficient tumours; whereas, the third mechanism applies to both mechanisms of action of PARPi. BRCAm BRCA mutation; HR homologous recombination; MDR1 multidrug resistance protein 1; p53BP1 tumour suppressor p53-binding protein 1; PARP poly(ADP-ribose) polymerase; PARPi PARP inhibitor(42).

BRCA1 gene is expressed in some organs, including the breast, ovary and thymus and testis, and also, it has a role in the differentiation of epithelial tissues. It seems the normal function of BRCA1 protein suppress the signalling of mammary epithelial cells by estrogen receptors. In addition, it can play a pivotal role in controlling the sex steroid-regulated pathways inducing breast cancer development(40).

Research suggests that the BRCA1 protein also regulates the activity of other genes and plays a critical role in embryonic development. The BRCA1 protein probably interacts with many other proteins including a protein called RAD51, including other tumour suppressors and proteins that regulate cell division. Like BRCA1, BRCA2 is expressed in numerous tissues. It is involved in the same biochemical processes of BRCA1 that aforementioned(40) (figure 1-10).

Mutations in *BRCA* genes cannot account for all cases of HBOC, indicating that the remaining cases can be attributed to the involvement of constitutive epi-mutations or other cancer susceptibility genes, which include Fanconi anaemia[1] cluster (*PALB2*, *BRIP1* and *RAD51C*), mismatch repair[2] cluster (*MLH1, MSH2, PMS1, PMS2* and *MSH6*),

[1] - FA
[2] - MMR

DNA repair cluster (*ATM, ATR* and *CHEK1/2*), and tumour suppressor cluster (*TP53*, *SKT11* and *PTEN*) have been associated with increased risk of breast and ovarian cancer as part of other cancer syndromes and plays a key role in the pathogenesis of cancer predispo-sition syndromes(40)(figure1-10, figure1-9).

Currently, the BRCA Exchange displays over 20,000 unique *BRCA1* and *BRCA2* variants; more than 6,100 of those variants are displayed with expert classifications from the ENIGMA (https://enigmaconsortium.org). The Consortium and to date approximately 3,700 have been deemed pathogenic (Figure 1-12) have been identified by extensive mutational analysis. Most of the *BRCA1* and *BRCA2* mutations are frameshift or non-sense that give rise to truncated nonfunctioning proteins whereas other mutations are missense substitutions or intronic variants, including those involved in the splicing pro-cess(43).

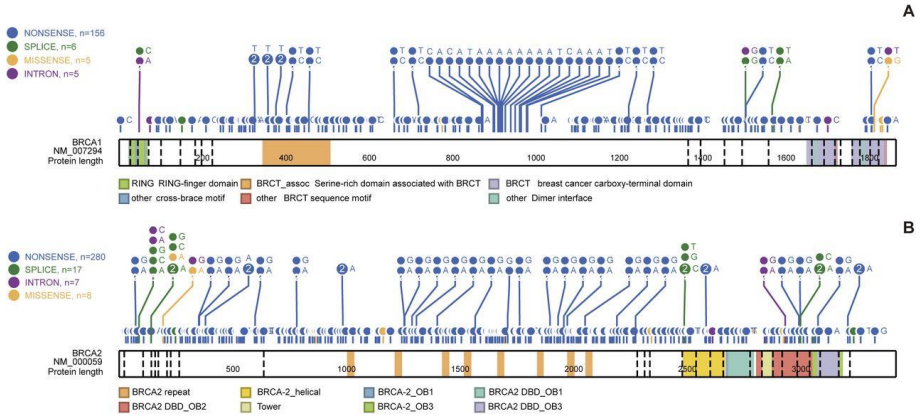

Figure 1-12: Distribution of the shared human BRCA PLP variants. It shows that the shared human BRCA PLP variants were not enriched at specific functional domains. (A) BRCA1. (B) BRCA2. Different colors within coding region indicate the locations of functional domains. Generated by using the ProteinPaint program[1].

1.2.2 Adenomatous polyposis coli[2]

The *APC* gene provides instructions for making the APC protein, which plays a critical role in several cellular processes. The APC protein acts as a tumor sup-pressor. Therefore, it keeps cells from growing and dividing too fast or in an uncontrolled way(44).

Mutations in APC often occur early on in cancers such as colon cancer. *APC* germline variants are associated with familial adenomatous polyposis[3] and at-tenuated FAP, with 95% being nonsense/frameshift mutations leading to prem-

[1] - https://proteinpaint.stjude.org
[2] - APC
[3] - FAP

ature stop codons. The FAP is an autosomal dominant colon cancer predisposition syndrome characterized by hundreds to thousands of colorectal adenomatous polyps, and accounts for ~1% of all colorectal cancer. Inactivation of APC protein was often found in many cancers such as thyroid cancer, breast cancer, and colorectal cancer. Mutations in APC lead to loss of β-catenin regulation, altered cell migration and chromosome instability(44).

1.2.3 ATM

ATM serine/threonine kinase, symbol *ATM*, is a serine/threonine protein kinase that is recruited and activated by DNA double-strand breaks. In 1995, the gene was discovered by Yosef Shiloh(45) who named its product ATM since he found that its mutations are responsible for the disorder ataxia–telangiectasia. In 1998, the Shiloh and Kastan laboratories independently showed that ATM is a protein kinase whose activity is enhanced by DNA damage(46, 47).

ATM Protein phosphorylates several key proteins that initiate activation of the DNA damage checkpoint, leading to cell cycle arrest, DNA repair or apoptosis. Several of these targets, including *p53, CHK2, BRCA1, NBS1* and *H2AX* are tumor suppressors.

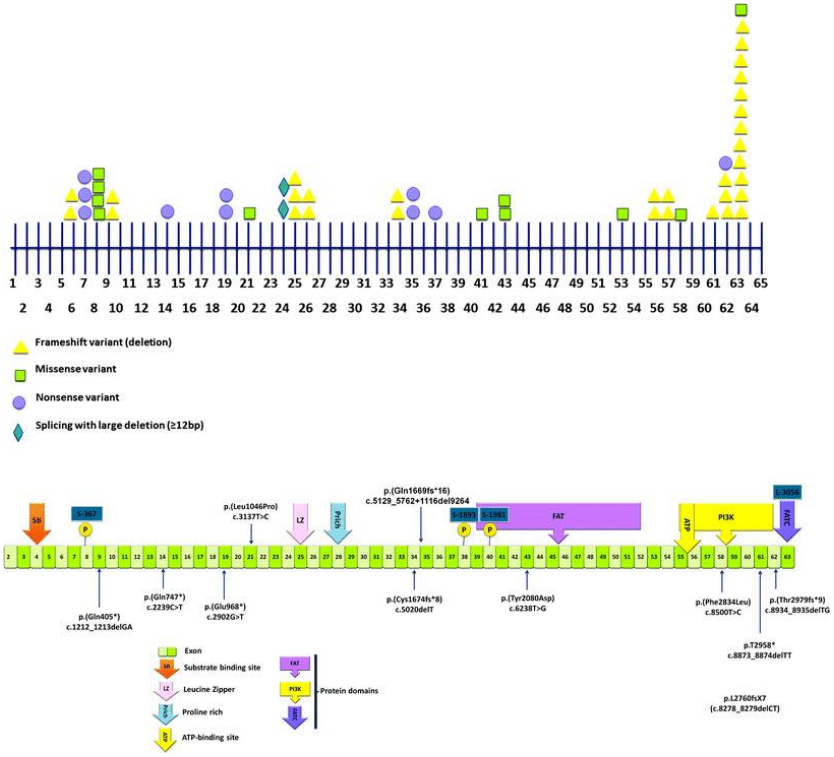

Figure 1-13: Schematic representation of the ATM gene and the variants identified. Top Representation of the 63 exons of the ATM gene and position of the gene variants found in the 54 alleles identified. Indicated symbols express the different types of variants. Bottom Position of the new gene variants and predicted impact on the different protein functional domains(48)

ATM also is associated with the autosomal recessive condition ataxia-telangiectasia. This condition is characterized by progressive cerebellar ataxia with onset between the ages of one and 4 years, telangiectasias of the conjunctivae, oculomotor apraxia, immune defects, and cancer predisposition, particularly leukemia and lymphoma. The ataxia telangiectasia mutated[1] gene is a moderate-risk breast cancer susceptibility gene; germline loss-of-function variants are found in up to 3% of hereditary breast and ovarian cancer[2] families who undergo genetic testing(figure 1-13)(49-51).

1.2.4 BARD1, BRIP1, MRE11A, NBN, RAD50, and RAD51C

BARD1, BRIP1, MRE11A, NBN, RAD50, and *RAD51C* are genes in the Fanconi anemia-BRCA pathway[3]. Variants in these genes are estimated to increase up to a 4-fold increase in the risk for breast cancer. This pathway is also associated with a higher risk of ovarian cancer and, less often, pancreatic cancer. The MRE11-RAD50-NBS1[4] protein complex plays a pivotal role in the sensing and early processing of double-strand breaks, thus maintaining genomic integrity More importantly, individuals carrying biallelic hippomorphic NBN mutations suffer from the Nijmegen breakage syndrome, being susceptible to several types of cancer. Approximately 40% of them will develop a malignancy before the age of 21(52).

1.2.5 SMAD4

In mammals, the *SMAD4* gene contains 54829 base pairs and is located in the region 21.1 of the chromosome 18.

SMAD4 are genes mutated in juvenile polyposis syndrome[5] and account for 45-60% of cases of JPS. JPS is an autosomal dominant disorder that predisposes to the development of polyps in the gastrointestinal tract. People with JPS are considered to be at an increased risk for stomach, colorectal, small intestine, and pancreatic cancers. The overall estimated cancer risk associated with JPS is between 9% to 50%, but the risks for each specific type of cancer have not been determined(53).

[1] - ATM
[2] - HBOC
[3] - (FA)-BRCA pathway
[4] - MRN
[5] - JPS

1.2.6 CHEK2

Checkpoint Kinase 2[1] gene is located on the long (q) arm of chromosome 22 at position 12.1. CHEK2 is a Protein Coding gene. Diseases associated with CHEK2 include breast, prostate, colon, thyroid or kidney Cancer, and Li-Fraumeni Syndrome 2. Among its related pathways are Homology Directed Repair and DNA Damage(54).

CHEK2 regulates cell division by preventing cells from entering mitosis or arresting cell cycle in the gap 1 phase (G1), in response to DNA damage. *CHEK2* regulates the function of BRCA1 protein in DNA repair and has been associated with familial breast cancers. In carriers with no affected relative, the risk is approximately 20%, and it increases up to 44% when both first- and second-degree relatives are affected(54).

1.2.7 CDH1

The human *CDH1* gene has16 exons and is 100 kb. *CDH1* germline variants code the cell-cell adhesion protein E-cadherin were first identified in families with hereditary diffuse gastric cancer[2]. In addition, some research was shown that CDH1 germline variants have been associated with lobular breast cancer[3] and with hereditary diffuse gastric cancer. Over 30 germline pathologic mutations have been reported in families with HDGC. The identified mutations are distributed throughout the gene and are truncating mutations, caused by frameshift mutations, exon/intron splice site mutations, point mutations, and missense mutations. It was reported that the cumulative risk of gastric cancer for CDH1 mutation carriers by age 80 years is 67% for men and 83% for women. CDH1 mutations indicated a significant association with a lifetime risk of 39-52% of lobular breast cancer(55).

1.2.8 EPCAM, MLH1, MSH2, MSH6 and PMS2

Inherited defects in *EPCAM, MLH1, MSH2*, *MSH6* and *PMS2* genes encoding components of the major post replication DNA mismatch repair[4] system have been found to underlie many cases of Lynch Syndrome (hereditary nonpolyposis colon cancer[5]). The range of mutations identified in these genes includes missense,

[1] - CHEK2
[2] - HDGC
[3] - LBC
[4] - MMR
[5] - HNPCC

nonsense, frameshift, splice site mutations and deletion mutations. It seems that there are some polymorphisms in MMR genes that may cause an increased risk of developing cancer(figure1-14).

Lynch syndrome causes the growth of risk of cancer including colon cancer (60-80% lifetime risk), uterine/endometrial cancer (20-60% lifetime risk), gastric cancer (11-19% lifetime risk), ovarian cancer (4-13% lifetime risk), and also the small intestine cancer, hepatobiliary tract cancer, upper urinary tract and brain cancer(56).

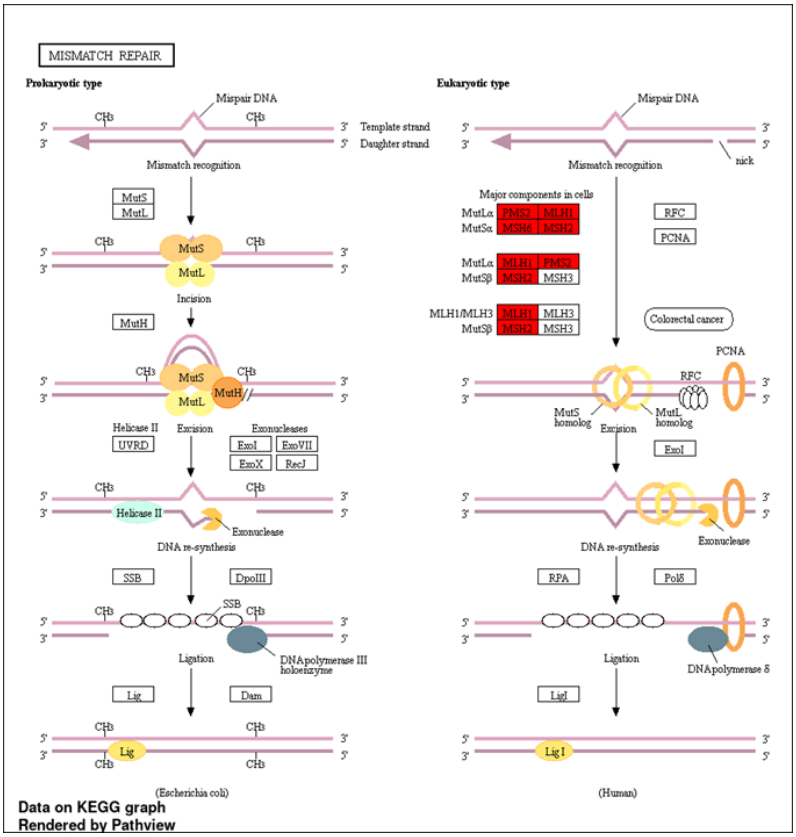

Figure 1-14: Role of the DNA repair gene in Mistmach repair pathway

1.2.9 MUTYH

MUTYH gene has its locus on the short (p) arm of chromosome 1 (1p34. 1). The gene has 16 exons.

MUTYH is involved with base-excision repair of DNA damaged by oxidative stress. Failure of BER results in mispairing of this nucleotide with adenine and resultant somatic CG–AT transversions in multiple genes, such as the APC and KRAS genes. *MUTYH* mutations are inherited in an autosomal recessive way, and approximately 10%-20% of classical familial adenomatous polyposis[1]

[1] - FAP

cases without an APC mutation and 30% of attenuated familial adenomatous polyposis[1] cases included(57).

MUTYH-associated polyposis[2] is an autosomal recessive polyposis syndrome caused by biallelic pathogenic germline variants in the MUTYH gene. Patients with MAP may be homozygous or compound heterozygous for pathogenic germline variants in the MUTYH gene(58).

1.2.10 PALB2[3]

PALB2 gene is located on chromosome 16, and it works together with *BRCA2* to repair damaged DNA. Partner and localizer of BRCA2 germline variants have been associated with an increased risk of pancreatic and breast cancer. Also, Familial pancreatic and breast cancer due to *PALB2* mutations is inherited in an autosomal dominant pattern. PALB2 absence prevents the recruitment of BRCA2 and RAD51 to the DNA double-strand breaks[4] site. *PALB2* operates as a tumor suppressor and is pivotal for efficient DNA repair by HR[5] in response to DSBs (figure 1-15). Via this way, it contributes to maintaining genome integrity(59).

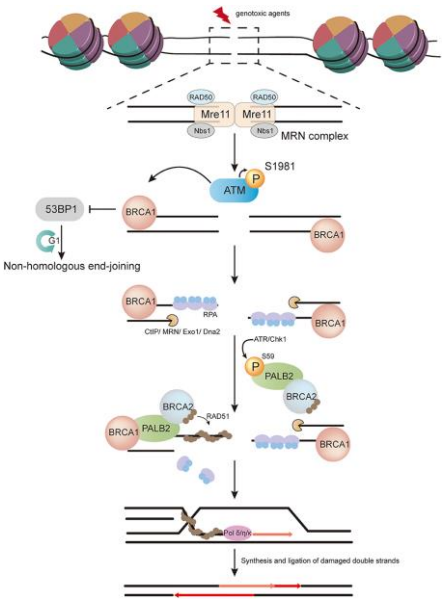

Figure 1-15 : The role of PALB2 HR. In response to DSBs in the S/G2 phase, the Mre11–RAD50–Nbs1 (MRN) complex is recruited to DSBs and promotes ATM recruitment. The inactive ATM dimer dissociates into active monomers via autophosphorylation at serine (S). PALB2 and BRCA2 further promote RPA removal and RAD51 loading. As a result, RAD51-ssDNA filament attacks the intact sister chromatid and extends the strand with the contribution of DNA polymerase. Finally, further restoration and ligation of double strands are done(59).

[1] - AFAP
[2] - MAP
[3] - Partner and Localizer of BRCA2
[4] - DSB
[5] - homologous recombination

1.2.11 PTEN[1]

Phosphatase and TENsin homolog[2] gene encodes a phosphatase called PTEN. It is a tumor suppressor gene located at chromosome 10q23. 31. The *PTEN* structure reveals a phosphatase domain and also has an enlarged active site important for the accommodation of the phosphoinositide substrate(60).

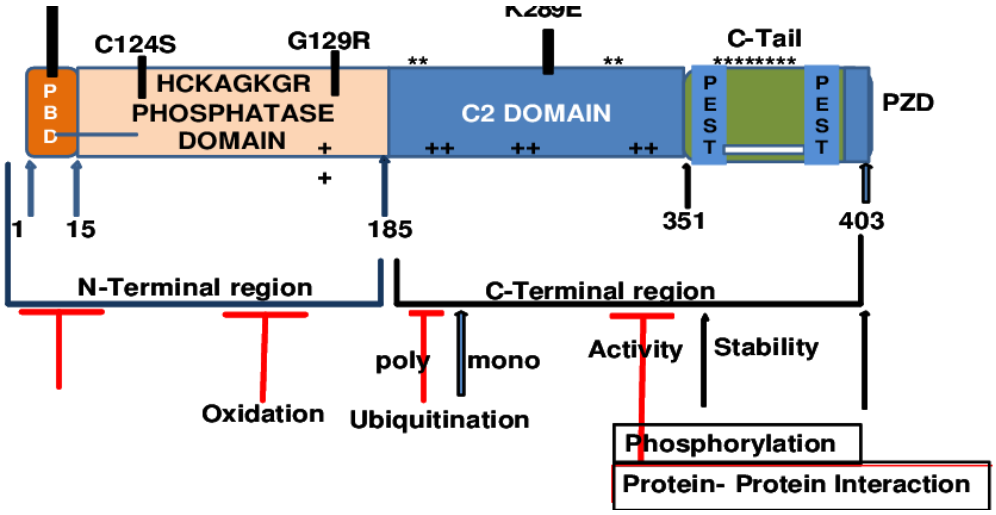

Figure 1-16 : Structure of PTEN (60)

PTEN has a C2 domain. The phosphatase and C2 domains associate across an extensive interface. The PTEN C2 domain bound phospholipid membranes (figure 1-16). mutation on it could reduce the membrane affinity of PTEN and its ability to suppress the growth of glioblastoma tumor cells. PTEN variants that have been associated with PTEN hamartoma tumour syndrome include Cowden syndrome (CS), Bannayan-Riley-Ruvalcaba syndrome and Proteus syndrome. CS is characterized by a high risk of developing tumours of the breast, thyroid and endometrium(60).

1.2.12 RECQL

RECQL[3] gene codes a Protein from a member of the RecQ DNA helicase family and located on 12 Chromosome. The biological function of the RecQ Like Helicase has not yet been determined. DNA helicases involve in various types of DNA repair, such as MMR, NER and direct repair. there are Two alternatively spliced transcripts of this gene, that encode the same isoform but differ in their

[1] - Phosphatase and tensin homolog
[2] - PTEN
[3] - RecQ Like Helicase

5' and 3' UTRs. Mutation on the RECQL gene causes inherited Cancer-Predisposing Syndrome and Tumor Predisposition Syndrome. Frequency of pathogenic RECQL variations in high-risk breast cancer patients who have previously tested negative for BRCA1 and BRCA2 mutations was approximately 1-2%(61).

1.2.13 STK11

The STK11 gene (also called LKB1) located on chromosome 19 encodes instructions in order to make an enzyme called serine/threonine kinase 11. STK11 enzyme is a tumor suppressor and also promotes apoptosis. In addition, serine/threonine kinase 11 function appears to be required for normal development before birth. Mutations in the STK11 gene associate with Peutz-Jegher syndrome[1] causes an increase in 57-81% risk gastrointestinal and breast cancer by age 70(62).

1.2.14 TP53

The *TP53* gene is located on chromosome 17 and contains 11 exons. The gene encodes a tumor suppressor protein that has 53 kDa nuclear phosphoprotein of 393 amino acid residues. This tumor suppressor protein has transcriptional activation, DNA binding, and oligomerization domains.

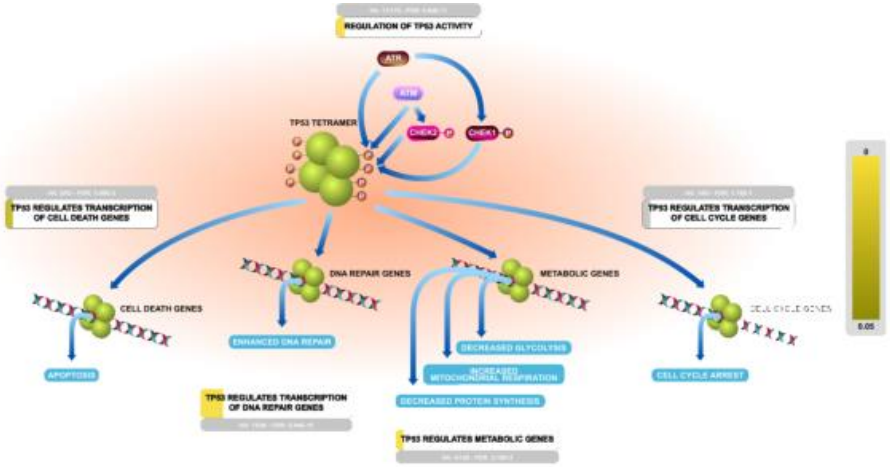

Figure 1-17: Transcriptional Regulation by TP53. TP53 regulates the transcription of multi genes that have role a variety of cellular processes, such as cellular metabolism, survival, senescence, apoptosis and DNA damage response[2].

it regulates expression of target genes via thereby inducing cell cycle arrest, apoptosis, senescence, DNA repair, or changes in metabolism (figure1-17). Mutations in TheTP53 gene are associated with an array of human cancers, including hereditary cancers such as Li-Fraumeni syndrome. Individuals with these

[1] - PJS
[2] _ Http://Reactome.com

mutations showed a 50% risk of developing any of the associated cancers by age 30 and a lifetime risk up to 90%, including sarcomas, breast cancer, brain tumors and adrenal gland cancer(63).

1.2.15 NF1[1]

The *NF1* gene is located on the long (q) arm of chromosome 17. *NF1* encodes a negative regulator in the RAS signal transduction pathway. Over 1000 pathogenic allelic variants of the *NF1* gene have been identified. Carrier Women of mutation *NF1* have a substantial risk of breast cancer and poor survival. *NF1* variations are associated with neurofibromatosis type 1, juvenile myelomonocytic leukemia, and Watson syndrome(64).

1.2.16 RAD51D

The *RAD51D* (RAD51 Paralog D) gene is cytogenetically located on chromosome 17. *RAD51D* are involved in the homologous recombination and repair of DNA. This gene codes a protein from member of the RAD51 protein family. The mutations in *RAD51D* have been associated with familial breast and ovarian cancer and Hereditary Breast Ovarian Cancer Syndrome. *RAD51D* pathogenic mutations are generally rare, contributing to approximately 0.5%–0.9% of breast-ovarian patients of BRCA1 and BRCA2 negative families(65).

1.2.17 CDK4[2]

CDK4 is located on 12choromosome. The protein encoded by this gene is a member of the Ser/Thr protein kinase family.

CDK family is one of the most critical protein families in cell division regulation. *CDK4* accompany with *CDK6* has key role cell cycle G1 phase progression. CDK4/6 interacts with cyclin D to form the cyclin D-CDK4/6 complex, which phosphorylates retinoblastoma[3] in G1 phase of cell cycle. Mutations in this gene as well as in its related proteins including D-type cyclins, p16(INK4a) and Rb cause dysregulation of cyclin D1-CDK4/6-Rb signaling cascade was observed and promoted unchecked cell proliferation and also are associated with a variety of cancers, particularly cutaneous melanoma and breast cancer(66).

[1] - Neurofibromatosis Type 1
[2] - Cyclin Dependent Kinase 4
[3] - Rb

1.2.18 CDKN2A[1]

The Cyclin-Dependent Kinase Inhibitor 2A gene provides instructions for making several tumor suppressor proteins through their involvement in 2 cell cycle regulatory pathways: the p53 pathway and the RB1 pathway. The gene is located on chromosome 9. It is ubiquitously expressed in many tissues and cell types. Variants or deletions in CDKN2A are frequently found in different kind of tumor cells. Mutations in CDKN2A have been associated with risk of melanoma, along with pancreatic, central nervous system cancers and breast cancer. This gene produces several transcript variants which differ in their first exons. Therefore, these collected genetic variants will be useful in developing diagnostic, preventive, and treatment approaches(67).

1.3 Hereditary Breast, and Ovarian cancers[2]

Hereditary breast and ovarian cancer syndrome refer to a predisposition to developing cancer that is transferred through the generations of a family. HBOC syndrome is an autosomal dominantly inherited disease characterized by:

a) a young age of onset
b) more than one synchronous or metachronous tumor
c) a family history of first and second-degree relatives with similar cancers.

Breast cancer is the most commonly diagnosed cancer among women after non-melanoma skin cancer[3] and is the second leading cause of cancer death after lung cancer. HBOC is inherited in the form of a harmful mutation, or change, in a gene. Different studies indicated that mutations in the genes *BRCA1* and *BRCA2* are important causes associated with HBOC, but harmful changes in other genes also can increase the risk of cancer. Having one of these mutations shows the growth in the chances of developing ovarian cancer. Most breast and ovarian cancers are sporadic (80-85%), but hereditary pre-disposition accounts for 10–15% of the cases (Figure 1), and carriers of an inherited genetic mutation and epigenetic aberrations in the tumor suppressor genes have an increased risk of the high probability of cancer developing cancer(68, 69).

[1] - Cyclin-Dependent Kinase Inhibitor 2a
[2] - HBOC
[3] - NMSC

Also, research showed Cumulative Breast Cancer risk for BRCA1 and BRCA2 mutation carriers at 70 years of age is about 57%, respectively, while cumulative OC risk is approximately 40% for BRCA1 and 18% for BRCA2 mutation carriers(68, 69).

The risk of getting breast cancer grows with age, the chance of developing breast cancer up to 49 years (1 in 40 women) is 2.4%, in 50 until 69 years (1 woman in 20) is approximately 5.5% and 4.7% between 70 and 84 (1 woman in 25)(70, 71).

Male breast cancer represents 1% of all breast cancer. In Italy, the rate of new female breast cancers diagnosed has been reported overall 53,000 at every year and the incidence of male breast cancer is about 1 case every 100,000 individuals, with a diagnosis in men aged 58-63 years.

Multiple factors are associated with an increased risk of developing breast cancer:

a) age
b) family history
c) exposure to reproductive hormones
d) dietary factors
e) benign breast diseases and environmental factors.

Women with HBOC syndrome are diagnosed at a young age which resulted in increased lifetime risk for developing breast, ovarian and other cancer. The presence of both breast and ovarian cancer in the family increases the likelihood that a cancer-predisposing mutation is present(69).

The genes involved in the majority of familial cancers are still unknown, Familial breast cancer constitutes a small fraction of rate of the breast cancer. The family features that suggest hereditary breast cancer predisposition include the following:

a) multiple cases of breast and ovarian cancer in different generations;

b) an early onset age at diagnosis of breast cancer;

c) two or more primary cancers in the same individual. These could be multiple primary cancers of the same type (e.g., bilateral breast cancer) or primary cancer of different types (e.g., breast and ovarian cancer in the same individual);

d) male breast cancer(69).

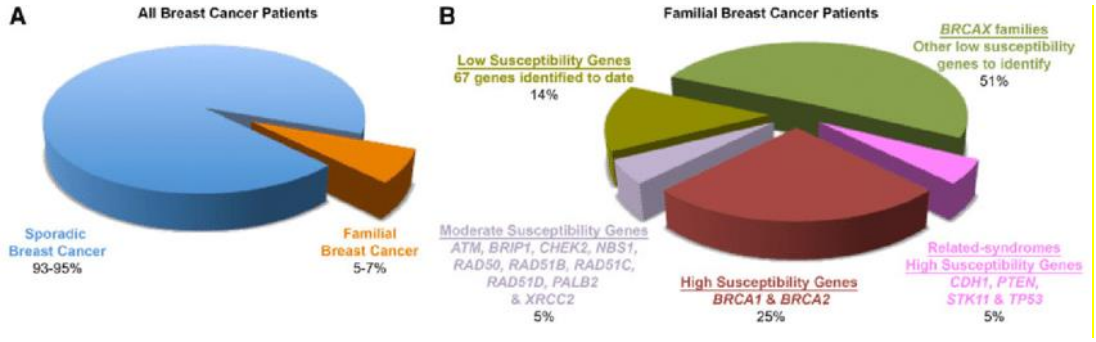

Figure1 Distribution of breast cancer cases: a) Familial breast cancer included a minor percentage of all breast cancer patients. b) Proportion of familial breast cancer patients due to germline mutations in high, moderate, and low penetrance cancer genes(72).

1.4 HEREDITARY COLORECTAL CANCER

Approximately, one third of all Colorectal cancers CRCs is thought to be related to heritable factors(73). The Hereditary CRC syndromes can be subclassified into polyposis-associated syndromes and non-polyposis syndromes.

The non-polyposis syndrome included Lynch Syndrome[1] that mutations in MMR genes is causes it. The adenomatous polyps are a part of the LS phenotype, with much fewer polyps than in the classical polyposis-associated syndromes.

Among the major polyposis-associated syndromes are familial adenomatous polyposis[2], attenuated FAP[3] and MUTYH-associated polyposis[4]. The rarer polyposis-associated syndromes are hyperplastic polyposis and hamartomatous polyposis conditions, including PeutzJeghers syndrome, Cowden syndrome and juvenile polyposis(73).

In MAP, patients have bi-allelic inactivation of the MUTYH-based-excision repair. About 80% of the FAP is due to a germline mutation in the tumor suppressor gene APC or biallelic mutations in the MUTYH(74).

Almost 25% of FAP is due to new mutations in APC. In the AFAP, patients have an inherited mutation of the APC gene(75).

[1] - LS
[2] - FAP
[3] - AFAP
[4] - MAP

1.4.1 Lynch Syndrome

Lynch syndrome is one of the most important hereditary CRC susceptibility syndromes and is caused through the MSI[1] pathway. LS was formerly known as hereditary non-polyposis colorectal cancer[2]. LS is inherited in autosomal dominant and because of the having pathogenic germline mutation in any of the DNA MMR genes *MLH1* on chromosome 3p21, *MSH2* on chromosome 2p16, *MSH6* on chromosome 2p16 or *PMS2* on chromosome 7p22(76). Sometimes germline deletions in the 3′ end of the EPCAM gene located directly upstream of MSH2 are caused LS, these mutations cause methylation-induced transcriptional silencing of MSH2(76).

Genetic heterogeneity and great variation in phenotypic manifestations are the two main factors in the characterization of Lynch syndrome. Both environmental factors and genetic factors are needed for heterogeneity. These two factors complicate the diagnosis and management of patients. Rare skin lesions, including sebaceous lesions and keratoacanthomas, can be indicative of a rare variant of LS called Muir Torres(77). LS is associated with various cancers such as the biliary tract, urinary tract (kidney, renal pelvis, ureter and bladder), ovary, stomach, small bowel, and less frequently brain cancer and pancreas. Cancer risks are generally classified according to gender, tumour type and mutated MMR gene(78).

1.4.1.1 Clinical criteria

There is an extensive variation in age of onset between and within families with LS, and genetic prediction (progressively earlier age at onset in successive generations) has been indicated.

In the 1980s, due to the genetic background of LS was still unknown, the diagnosis of LS was based on family history. In 1990 a group of researchers in Amsterdam tried to standardize clinical criteria for LS. Therefore, they succeed to publish the guidelines referred to as the Amsterdam I Criteria in 1991. It was based on a strong family history of CRC with a younger age of onset(79). These criteria were later several times revised. Amsterdam II Criteria include extracolonic manifestations such as cancers of the endometrium, small bowel or pelvic-

[1] - Microsatellite instability
[2] - HNPCC

ureter system. Increasing molecular knowledge changed and improved guidelines (Table 1-1). The discovery of microsatellite instability prompted the development of the Bethesda Guidelines that based on it, LS is a result of germ line MMR gene mutations(80). Thereafter, the Revised Bethesda Guidelines were edited again.

It includes panels to test microsatellite instability, directions for molecular evaluation of tumors and germline DNA, and a clinical selection of cases (80).

Table 1-1 . Clinical guidelines for identifying Lynch Syndrome

Amsterdam I criteria(79)	Amsterdam I criteria requires at least three relatives with histologically verified CRC, and the following: • Familial Adenomatous Polyposis has been ruled out • One should be first-degree relative to the other two • At least two successive generations are affected • At least one of the affected is diagnosed <50 years of age
Amsterdam II criteria(81)	Amsterdam II criteria requires at least three relatives with a Lynch associated cancer (colorectal, endometrial, small intestine, ureter, renal pelvis) verified by pathologic examination, and the following: • Familial Adenomatous Polyposis has been ruled out in the CRC cases • One should be first-degree relative to the other two • At least two successive generations are affected • At least one of the affected is diagnosed <50 years of age
Revised Bethesda guidelines(82)	Revised Bethesda guidelines for testing of colorectal tumors for microsatellite instability (MSI) in families that meet the Amsterdam criteria: • At least one CRC is diagnosed <50 years of age • Presence of synchronous or metachronous LS- associated

	tumors* regardless of age • CRC with the MSI-H histology** diagnosed in a patient <60 years of age • CRC or LS-associated tumor* diagnosed under 50 years of age, in at least one first-degree relative • CRC or LS-associated tumor* in two first- or second-degree relatives, regardless of age
	* Endometrial, ovarian, gastric, hepatobiliary, or small-bowel cancer or transitional cell carcinoma of the renal pelvis or ureter LS-related tumors including colorectal, endometrial, stomach, ovarian, pancreas, ureter and renal pelvis, biliary tract, and brain (usually glioblastoma as seen in Turcot syndrome) tumors, sebaceous gland adenomas and keratoacanthoma in Muir–Torre syndrome, and carcinoma of the small bowel ** Presence of tumor-infiltrating lymphocytes, Crohn disease-like lymphocytic reaction, mucinous or signet-ring differentiation, or medullary growth pattern.

Nevertheless, the Amsterdam criteria Bethesta guidelines are not solid screening tests for LS. There is a likelihood of chance clustering of cancer within a family as long as only clinical criteria are used to identify LS, because CRC is a relatively common malignancy(83).

1.4.1.2 Founder mutations

Founder mutations originate or are entered via a single ancestor or a single individual respectively, who passes it on to next generations. Some of them are found in only a few families with a single origin, others are commonly identified in specific geographic regions or countries(84). Founder mutations may affect

the prevalence of LS in specific populations. They may lower the cost of molecular diagnosis due to founder mutations are a useful tool in genetic screening.

In the Ashkenazi Jewish population, the majority of LS cases appear to be caused by the mutations c.3959_3962delCAAG and c.3984_3987dupGTCA in MSH6 and the mutation c.1906G>C in MSH2 as an example(85). Also, MSH2 mutation (c.942+3A>T) has been identified in 27% of LS cases in the province of Newfoundland in Canada. However, this mutation is a usual MMR mutation, mutation carriers in Newfoundland share a common haplotype that is not in carriers from England, Japan Hong Kong or Italy. As result, the MSH2 c.942+3A>T offers a founder effect in this population(86).

Identification of germline MMR mutations is important to confirme LS and the clinical management(87). In addition, it has been reported that individuals with MMR deficiency may benefit from immunotherapy (Figure 1-18).

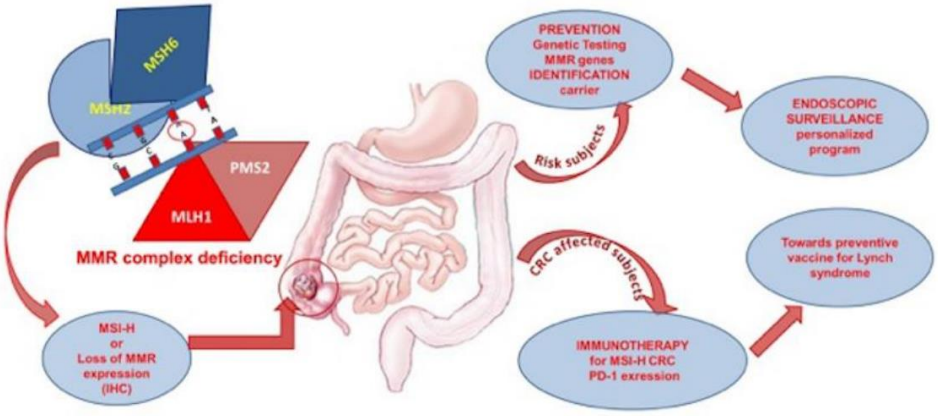

Figure 1-18 : Management of patients with LS(88).

Fortunately, Today the detection of LS has developed and tumor-based screening with germline confirmation has been used, instead of the only clinical 26 criteria. It resulted drop further costs for mutation screening. During the last decade next-generation sequencing of multigene panels for germline testing was applied to syndrome-specific gene. A number of the reported mutations are missense, silent or intronic variants with uncertain pathogenicity[1](89). The InSiGHT Variant Interpretation Committee[2] provides a criterion for classification of the mutations base on definition of variants at https://www.insight-group.org/criteria/.

[1] - VUS
[2] - VIC

In 2014, the MMR variants deposited in the InSiGHT database were reclassify to facilitate the management of families with suspected LS(90).

1.5 Molecular genetic testing in Breast Cancer

Molecular testing to identify genetic and genomic variation has become an integral part of cancer management. For patients with a family history of cancer, genetic testing will be critical to determine whether a hereditary cancer syndrome is present. Tumor genomic profiling, the standard of care for many types of malignancies, is significantly vital in the management of cancer. It has created the need for hereditary cancer mutation testing for all sufferers diagnosed with advanced cancer(91).

Next-generation sequencing allows timely testing of multiple genes. Due to increased knowledge of genes evaluated and also the need for a more comprehensive understanding of clinical management, multiple guidelines for testing have been developed (91).

1.5.1 Next generation sequencing

Next-generation sequencing (NGS) technologies provide the possibility to map whole genomes at affordable costs.

The NGS platforms available are MiS-eq/NextSeq (Illumina) and PGM/S5 Ion-Torrent (Thermo Fischer).

NGS technology are characterized by :

1) The generation of many millions of short reads in parallel,
2) The speed up of sequencing the process compared to the first generation,
3) The low cost of sequencing
4) The sequencing output is directly detected without the need for electrophoresis.

The NGS workflow can be divided into 4 steps following:

1) sample preprocessing,
2) library preparation,
3) sequencing itself
4) bioinformatics (Figure 1-19).

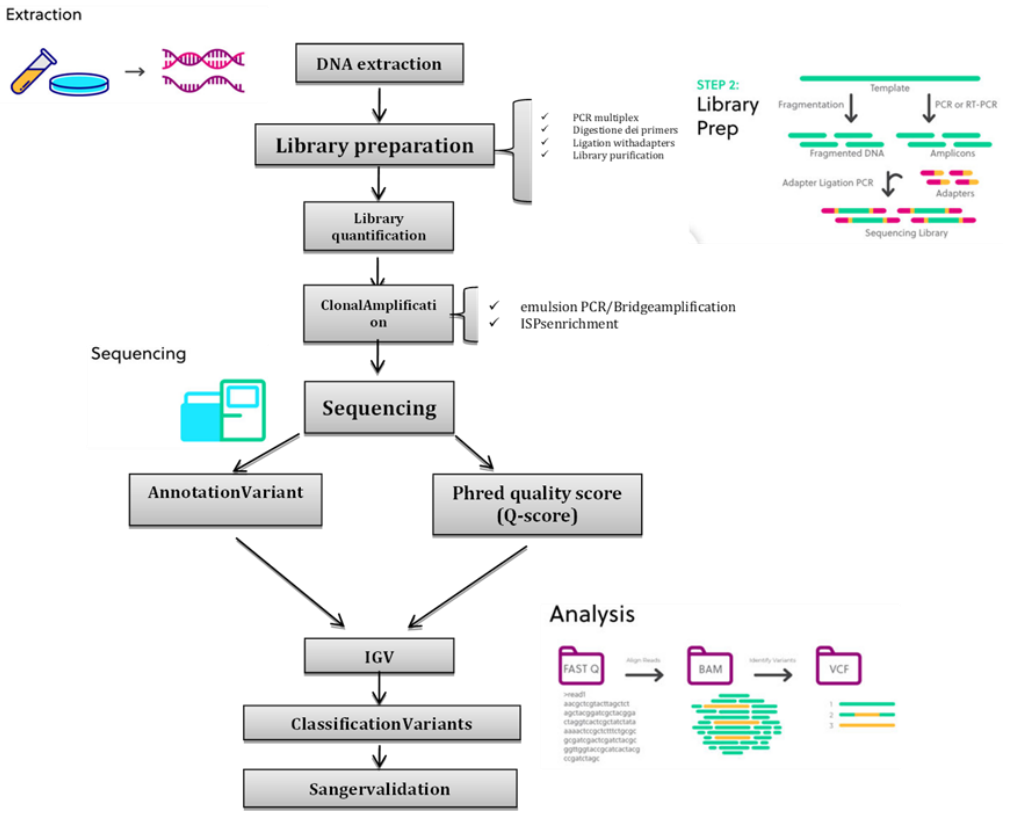

Figure 1-19 : Generic workflow for NGS.

1.5.1.1Ion torrent sequencing

The technology of the Ion Torrent is based on the detection of the hydrogen ion released during the sequencing process. Precisely, Ion Torrent employs a chip, including a set of micro wells and apiece has a bead with several identical fragments.

The hydrogen ion is released by the incorporation of each nucleotide with a fragment in the beads. As a result, the pH of the solution is change. This modification is identified by a sensor linked to the bottom of the micro well and transformed into a voltage signal which is proportional to the count of nucleotides incorporated (figure1-20).

The Ion GeneStudio S5 systems are designed to enable a broad range of targeted next-generation sequencing applications with speed and scalability. Five Ion S5 chips enable a sequencing throughput range of 2M to 130M reads per run (figure1-21).

The workflow has three major steps: library construction, template preparation and sequencing.

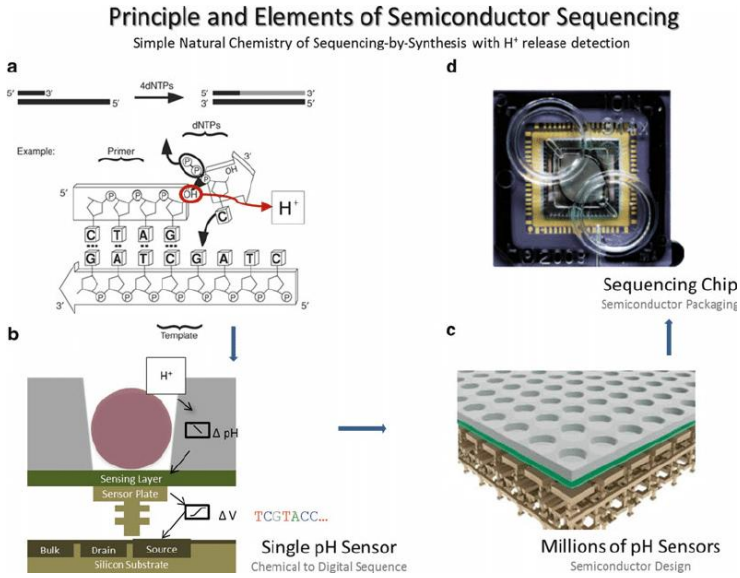

Figure 1-20 : Principles and elements of semiconductor sequencing. Simpler natural chemistry of sequencing by synthesis is implemented in the Ion Torrent platform. As the second strand of DNA is synthesized, the addition of every new nucleotide leads to a release of H + a) which is detected by a silicon pH sensor (b). Severalmillion pH sensors (c) are arranged within a sequencing chip (d). Cross-sectional view (c) shows the ion -sensitive layer in green with the microwells on the top surface (3 µm) and the transistor stack underneath(92).

Figure 1-21 : the Ion GeneStudio™ S5 Systems[1] : Sequencing and analysis in as little as 3 hours with the Ion GeneStudio S5 Prime System

[1] - https://www.thermofisher.com/

1.5.1.1.1 Library Construction

The first step in the workflow is the preparation of the library. For this goal, multiple primer pools to produce overlapping amplicons to cover large target regions is used.

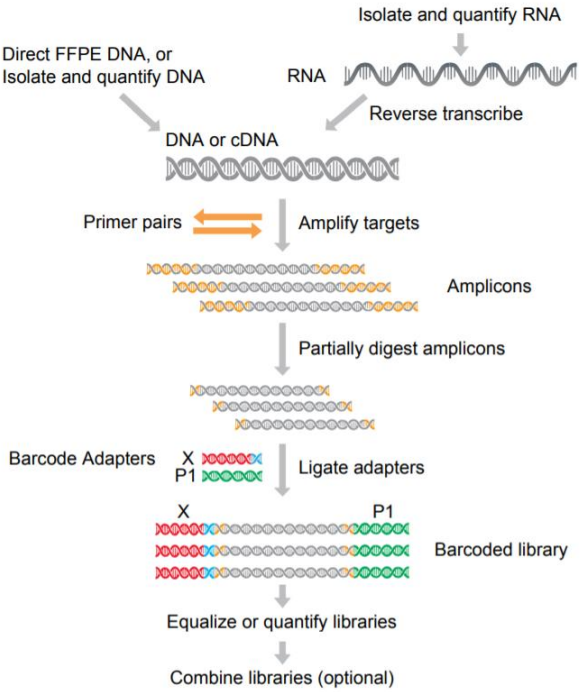

Figure 1-22. Workflow library preparation[1].

After multiplex PCR, the amplicons obtained are treated with FuPa reagent that digests and phosphorylates the ends of amplicons. Two adapters (P1 and X) are tied to the ends of these amplicons, required for emultion PCR step. The P1 adapter binding of the single amplicons to the IPS (Ions Sphere Particle), while the adapter X the binding of the primers. The system allows you to assign an identification code to each DNA sample or barcode (which is associated with adapter A) then, the analysis multiple samples in a single sequencing run (figure 1-22).

The process generally involves taking DNA (or RNA converted to DNA), fragmenting it to a uniform size (generally 200-400b) and then adding sequencing adapters. DNA Fragments are transformed into the library by ligation to adapters of the sequencing including particular sequences designed to interact with the NGS platform.

multiplex sequencing contributes to individual "barcode" sequences added to

[1] - Manual guide Ion AmpliSeq™ Library Kit 2.0 Workflow.

each DNA fragment during NGS library preparation also, each read can be recognized and sorted afore the final data analysis. These barcodes allow multiple samples in a single sequencing run(figure1-22).

1.5.1.1.2 Template Preparation/Amplification by Emulsion PCR

During the library prep, the fragments generated are attached to beads IPS (Ions Sphere Particle) and amplified using emulsion PCR (emPCR). Beads coated with complementary primers are blended with a dilute aqueous solution that includes the fragments to be sequenced with important PCR reagents. Then, the solution is mixed with oil to create an emulsion of micro-droplets. Followed by that clonal amplification of each fragment is done within the microdroplets. After emPCR, the emulsion is 'broken' (commonly by organic extraction and centrifugation) and the amplified beads are enriched in a glycerol gradient (with unamplified beads pelleting at the bottom) via the Enrichment system only the spheres on which each amplicon bound and amplified are chosen. This selection is done with help of the use of streptavidin magnetic beads having an affinity for biotinylated ends of amplicons. A denaturation step with NaOH detaches the streptavidin beads from the amplified ISPs.

1.5.1.1.3 Sequencing

The Ion Torrent system works based on measuring the direct release of H+ (protons) from the reaction. The lack of optics causes The Ion Torrent system can be a relatively inexpensive instrument coupled with disposable chips, and also, they don't have to contend with slow image scans, therefore the sequencing reactions are mostly fast, with 200b reads taking approximately 2 hours. The Ion Torrent sequencers can produce reads lengths of 200 bp, 400 bp and 600 bp with a throughput that can reach 10 Gb for the S5 sequencer. There are three different kinds of chips for the Personal Genome Machine (PGM) - 314, 316, 318 - and five different types for their GeneStudio S5 system (510, 520, 530, 540, 550) (Figure 1-23). The size of the chip is determined by the number of bases needing to be sequenced.

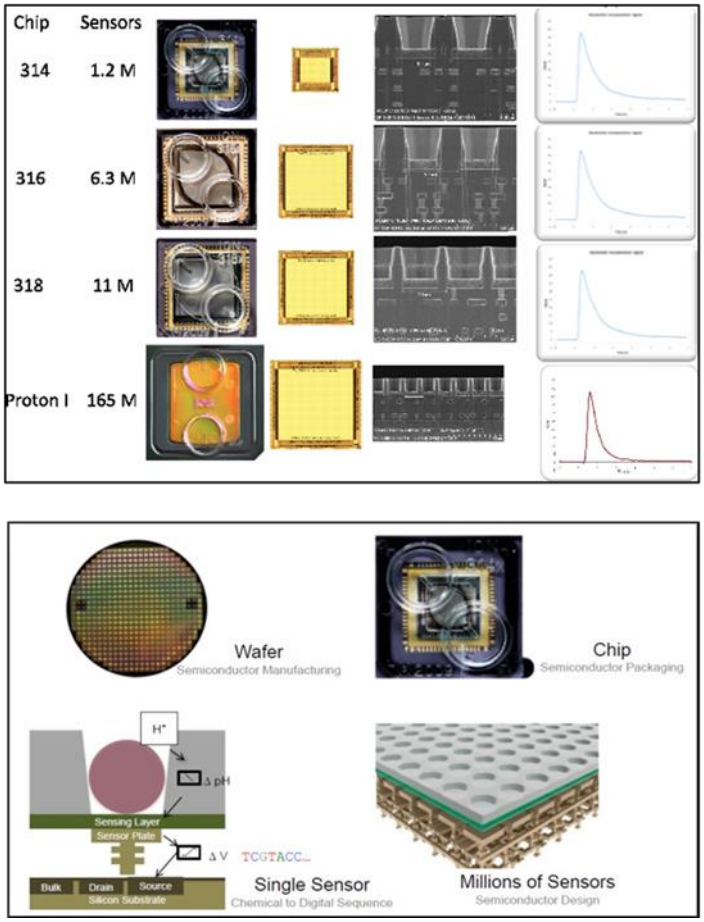

Figure 1-23 . Ion Torrent uses semiconductor CMOS chip architecture. This technology, in the form of a single-use, disposable chip, is the functional core of the Ion Torrent DNA sequencing platform[1].

Each of the chips includes a top layer which is covered by tiny wells. The wells are large enough to fit a single ISP.

then loading ISP solution onto the chip, the chip becomes ready for sequencing. Ion Torrent applies semiconductor sequencing technology. Nucleotides flow upon the chip one at a time. If the nucleotide is mixed, a hydrogen ion is released. This release causes to decrease in the pH of the liquid surrounding the ISP. This pH change when converted to a voltage change and is picked up by the software and recorded as that nucleotide (figure 1-23).

The important advantages of this sequencing technology are:

- read lengths which are longer than other sequencers
- fast sequencing time between 2 and 8 hours.

[1] -http://www.genomics.cn/en/

1.5.1.2 Illumina sequencing

Illumina sequencing technology is based on a synthesis approach and is currently the most widely used technology in the NGS market.

In the first step, the DNA samples are randomly fragmented into sequences and adapters are ligated to both ends of each sequence. Then, these adapters are fixed themselves to the respective complementary adapters, the latter are hooked on a slide with many variants of adapters (complementary) placed on a solid plate.

In the second step, each enclosed sequence to the solid plate is amplified by "PCR bridge amplification that creates a number of identical copies of each sequence; a set of sequences made from the same original sequence is called a cluster. the cluster includes almost one million copies of the same original sequence.

in the third step, nucleotides in the sequences are determined, Illumina uses the sequencing by synthesis approach that employs reversible terminators in which the four modified nucleotides, sequencing primers and DNA polymerases are added as a mix, and the primers are hybridized to the sequences.

Polymerases extend the primers using the modified nucleotides. Each nucleotide is labelled with a fluorescent specific. Clusters are excited by a laser to emit a light signal specific to each nucleotide, which will be detected by a coupled-charge device[1] camera and Computer programs will translate these signals into a nucleotide sequence(figure1-24).

The overall error rate of this sequencing technology is about 1%. Substitutions of nucleotides are the most common type of error in this technology, the main source of error is the bad identification of the incorporated nucleotide.

[1] - CCD

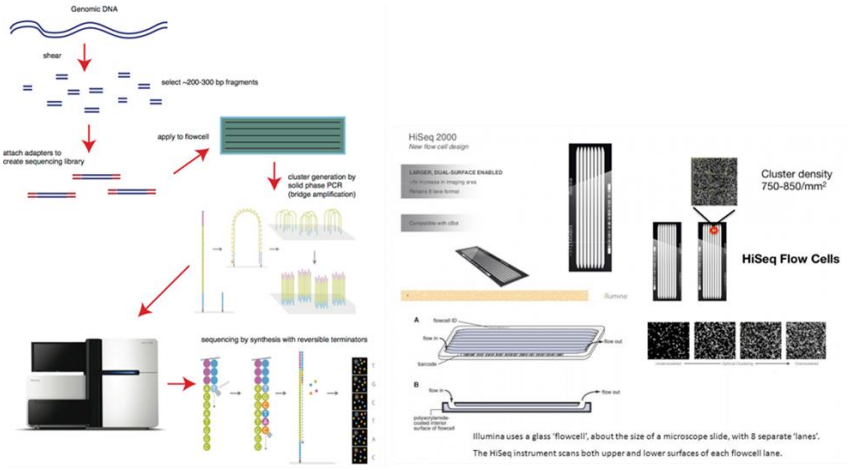

Figure 1-24 ; Illumina workflow[1]:Sequencing by synthesis in the presence of four fluorescently labeled nucleotide.

1.5.2 Bioinformatics Analysis of Target Resequencing

At the end of the work, bioinformatics is used to check and analyze the data. DNA sequence output as files in a FASTQ or unaligned BAM[2] format. With the help of bioinformatics, these data can be aligned. The FASTQ and uBAM file formats store short sequences as plain text with metadata (Figure 1-25).

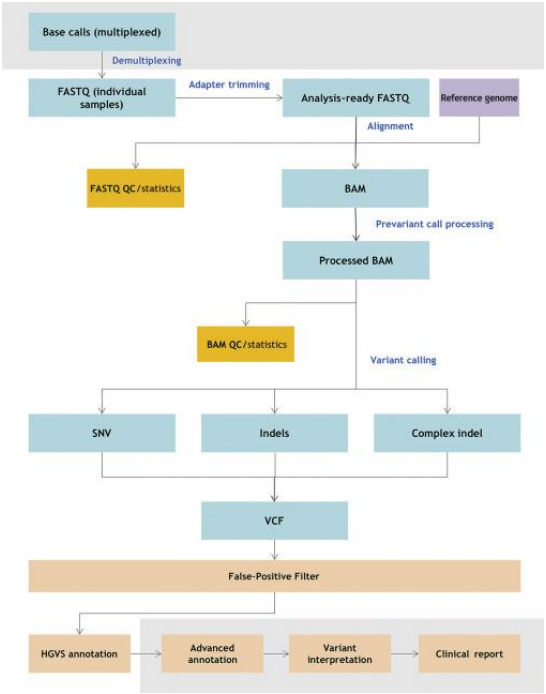

Figure 1-25: Next-generation sequencing (NGS) bioinformatics pipeline(93). The figure shows a bioinformatics pipeline and its components that are generally utilized to process NGS data.

[1] - *https://www.google.com/patents/DE202011003570U1?cl=it*
[2] - uBAM

The process of the sequence alignment assigns a genome positional context to the short reads in the reference genome (h19) and generates various metadata fields, that consists alignment characteristics (matches, mismatches, and gaps). The aligned sequences and the related metadata are stored in a Sequence Alignment Mapping (SAM[1]/BAM[2]) or CRAM file format (figure 1-25).

These files are like a compressed pillar file format storing biological sequences aligned to a reference sequence. Downstream algorithms use of the BAM file to recognize a range of genetic alterations, such as SNV[3], indels[4], and CNV[5] (figure 1-26).

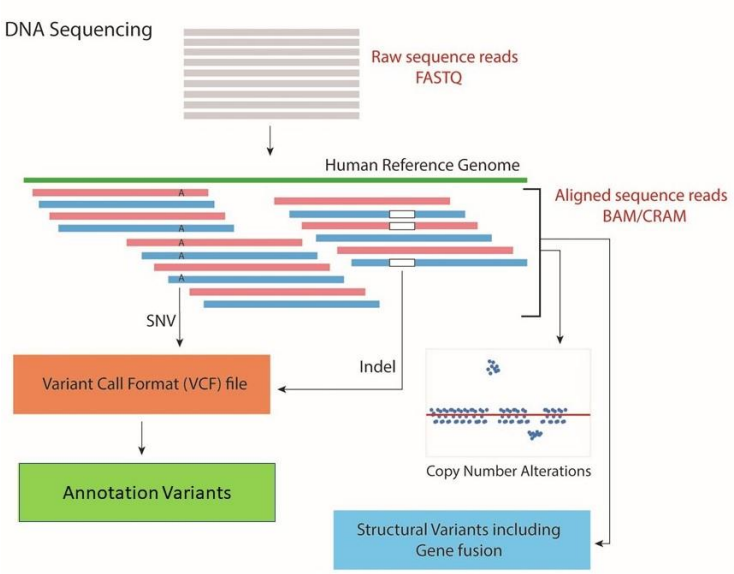

Figure 1-26: Bioinformatics Analysis of Target Resequencing[6]

The results of variant identification are saved in one of the variant call formats[7], such as generic feature format, genome VCF and others. These formats provide a possibility of encoding quantitative information of the variant, including genotype quality, depth of coverage at the variant position, and variant allele fraction(figure1-27).

[1] - Sequence Alignment Map
[2] - Binary Alignment Map
[3] - Single Nucleotide Variants
[4] - Insertions And Deletions
[5] - Copy Number Variation
[6] - Source: Clinical Laboratory News Author: Somakroy Md
[7] - Vcf

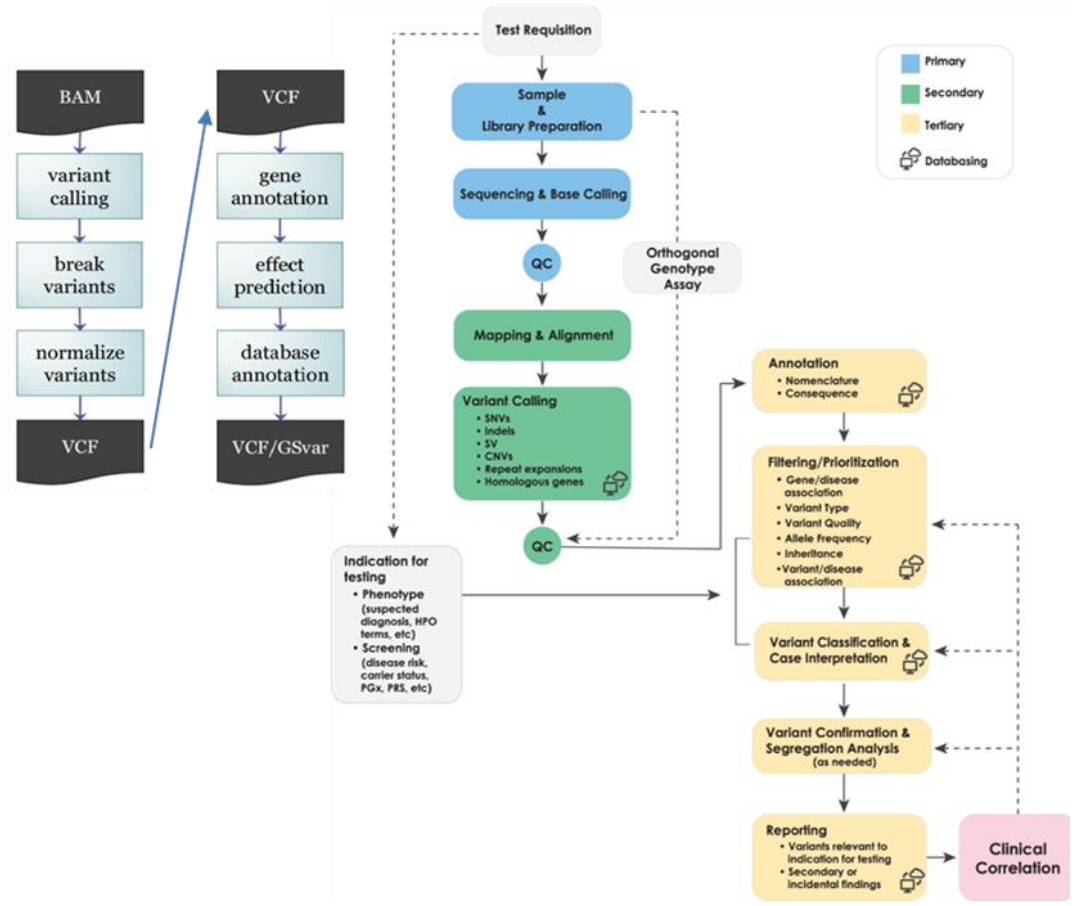

Figure 1-27: The ase shows the main steps for annotation of variants. Variant frequuencies is taken from several database 1000 genome, ExAC, Cosmic.Finally, clinical information from Clinvar, HGMD, Varsome is annotated(94).

The downstream bioinformatics analysis of the DNA sequence variants involves queries across multiple genomic databases to extract meaningful information in field of the gene and variant nomenclature, variant prevalence, functional impact (phyloP, Sift, FATHMM, CADD), and assertion of clinical significance (Figure 1-27).

Databases:

dbSNP*: ttps://www.ncbi.nlm.nih.gov/snp/,*

https://www.ncbi.nlm.nih.gov/clinvar/,

Varsome: *https://varsome.com/, Enigma*

https://brcaexchange.org/variants

A user interface via the IGV[1] software (http://software.broadinstitute.org/software/igv/) will render and visualize annotated DNA sequence variants, CNV, Indel, SNV and other genetic variations (Figure 1-28).

[1] - Integrative Genomics Viewer

Therefore, bioinformatic helps to interpret the clinical significance of the genetic alterations and release a comprehensive molecular report.

Other applications of it in molecular laboratory can be named quality control monitoring of sequencing data across runs, identification of background sequencing noise to reduce false-positive results, validation of upgrades to the bioinformatics pipeline, variant interpretation, the development and validation of novel algorithms in order to process sequence data.

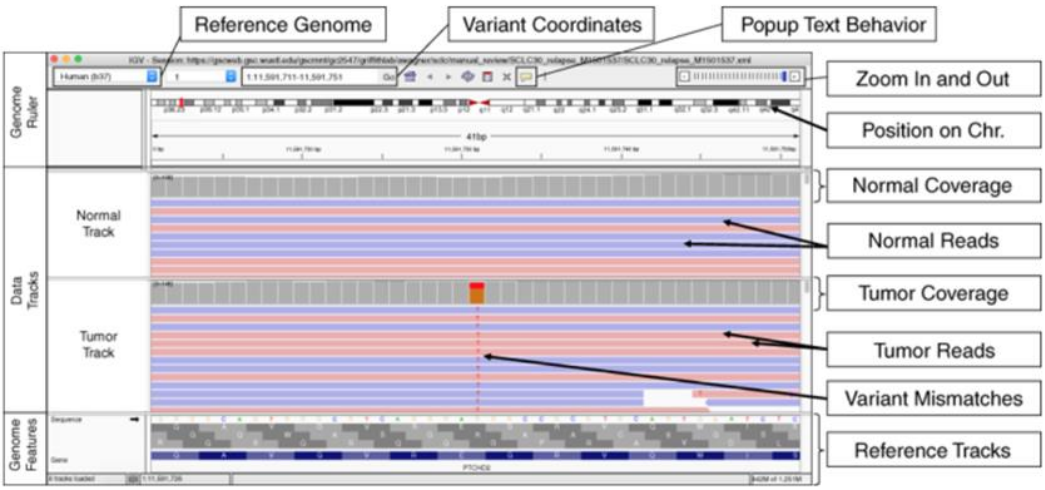

Figure 1-28. Example of IGV interface with associated features relevant to manual review of cancer variants(95).

Variant Nomenclature

Variant nomenclature is an essential part of a clinical report and represents the fundamental element of a molecular test result. The Human Genome Variation Society (HGVS) (https://varnomen.hgvs.org/)variant nomenclature system is the de-facto representation of sequence variants in a clinical report, which is universally accepted as a standard by laboratory accreditation agencies and understood by molecular professionals, clinicians, and medical genetics professionals [47]. The synthesis of this nomenclature for variants identified by NGS testing requires a complex process of conversion of the coordinate system from the reference genome to specific complementary DNA and protein transcripts. The alignment of the transcripts to the forward or the reverse genomic DNA strands and the HGVS 3'rule for variants in repeat sequence regions add additional complexity to the process. Several annotation toolsboth open source and commercial can generate HGVS nomenclature. Clinical reports should include sequence reference(s) to ensure unambiguous naming of the variant at the DNA

level, as well as to provide coding and protein nomenclature to assist in functional interpretations (e.g., “g.” for genomic sequence, “c.” for coding DNA sequence, “p.” for protein, “m.” for mitochondria).

1.5.2.1Variant Classification

The American College of Medical Genetics and Genomics[1] and the Association for Molecular Pathology[2] have published guidelines for interpreting sequence(96). Regarding the guidelines, first, pathogenic variations according to the gathering evidence from various sources have to be interpreted. For instance, the position and the results of family segregation analysis, type of variant, the survey results of many kinds of databases including control populations as well as disease-specific databases, the results of in silico prediction programs and the results of in vitro experiments(96).

These guidelines employed a statistical approach to the classification of variant pathogenicity. When the prevalence of variants in affected cases is notably more than that in the control population (over fivefold based on the odds ratio obtained from the case vs control and the 95% confidence interval around the estimated odds ratio does not include 1.0), this was regarded as evidence supporting the pathogenicity of the variant(96)(table1-2).

Since guidelines are very strict, may result in a larger proportion of variants being categorized as uncertain significance.

There are two sets of criteria to interpret the variants:

1- the criterion for the classification of pathogenic or likely pathogenic variants (Table 1-3)

2- the criterion for the classification of benign or likely benign variants (Table 1-4).

The pathogenic criteria are weighted as very strong (PVS1), strong (PS1–4); moderate (PM1–6), or supporting (PP1–5), and each benign criterion is weighted as stand-alone (BA1), strong (BS1–4), or supporting (BP1–6) (table1-2).

[1] - ACMG
[2] - AMP

Table 1-2: Evidence framework(96). The table organizes each of the criteria base on the type of evidence as well as the strength of the criteria for a benign (left side) or pathogenic (right side) assertion.BS, benign strong; BP, benign supporting; FH, family history; LOF, loss of function; MAF, minor allele frequency; path., pathogenic; PM, pathogenic moderate; PP, pathogenic supporting; PS, pathogenic strong; PVS, pathogenic very strong

	Benign		Pathogenic			
	Strong	**Supporting**	**Supporting**	**Moderate**	**Strong**	**Very strong**
Population data	MAF is too high for disorder BA1/BS1 **OR** observation in controls inconsistent with disease penetrance BS2			Absent in population databases PM2	Prevalence in affecteds statistically increased over controls PS4	
Computational and predictive data		Multiple lines of computational evidence suggest no impact on gene /gene product BP4 Missense in gene where only truncating cause disease BP1 Silent variant with non predicted splice impact BP7 In-frame indels in repeat w/out known function BP3	Multiple lines of computational evidence support a deleterious effect on the gene /gene product PP3	Novel missense change at an amino acid residue where a different pathogenic missense change has been seen before PM5 Protein length changing variant PM4	Same amino acid change as an established pathogenic variant PS1	Predicted null variant in a gene where LOF is a known mechanism of disease PVS1
Functional data	Well-established functional studies show no deleterious effect BS3		Missense in gene with low rate of benign missense variants and path. missenses common PP2	Mutational hot spot or well-studied functional domain without benign variation PM1	Well-established functional studies show a deleterious effect PS3	
Segregation data	Nonsegregation with disease BS4		Cosegregation with disease in multiple affected family members PP1	Increased segregation data →		
De novo data				De novo (without paternity & maternity confirmed) PM6	De novo (paternity and maternity confirmed) PS2	
Allelic data		Observed in *trans* with a dominant variant BP2 Observed in *cis* with a pathogenic variant BP2		For recessive disorders, detected in trans with a pathogenic variant PM3		
Other database		Reputable source w/out shared data = benign BP6	Reputable source = pathogenic PP5			
Other data		Found in case with an alternate cause BP5	Patient's phenotype or FH highly specific for gene PP4			

Table 1-3 Classification of pathogenic or likely pathogenic variants.(96)

Evidence of benign impact	Category
Stand-alone	BA1 Allele frequency is >5% in Exome Sequencing Project, 1000 Genomes Project, or Exome Aggregation Consortium
Strong	BS1 Allele frequency is greater than expected for disorder (see **Table 6**)
	BS2 Observed in a healthy adult individual for a recessive (homozygous), dominant (heterozygous), or X-linked (hemizygous) disorder, with full penetrance expected at an early age
	BS3 Well-established in vitro or in vivo functional studies show no damaging effect on protein function or splicing
	BS4 Lack of segregation in affected members of a family
	Caveat: The presence of phenocopies for common phenotypes (i.e., cancer, epilepsy) can mimic lack of segregation among affected individuals. Also, families may have more than one pathogenic variant contributing to an autosomal dominant disorder, further confounding an apparent lack of segregation.
Supporting	BP1 Missense variant in a gene for which primarily truncating variants are known to cause disease
	BP2 Observed in *trans* with a pathogenic variant for a fully penetrant dominant gene/disorder or observed in *cis* with a pathogenic variant in any inheritance pattern
	BP3 In-frame deletions/insertions in a repetitive region without a known function
	BP4 Multiple lines of computational evidence suggest no impact on gene or gene product (conservation, evolutionary, splicing impact, etc.)
	Caveat: Because many in silico algorithms use the same or very similar input for their predictions, each algorithm cannot be counted as an independent criterion. BP4 can be used only once in any evaluation of a variant.
	BP5 Variant found in a case with an alternate molecular basis for disease
	BP6 Reputable source recently reports variant as benign, but the evidence is not available to the laboratory to perform an independent evaluation
	BP7 A synonymous (silent) variant for which splicing prediction algorithms predict no impact to the splice consensus sequence nor the creation of a new splice site AND the nucleotide is not highly conserved

Table 1-4 , Classification of benign or likely benign variants(96)

Evidence of pathogenicity	Category
Very strong	PVS1 null variant (nonsense, frameshift, canonical ±1 or 2 splice sites, initiation codon, single or multiexon deletion) in a gene where LOF is a known mechanism of disease Caveats: • Beware of genes where LOF is not a known disease mechanism (e.g., *GFAP*, *MYH7*) • Use caution interpreting LOF variants at the extreme 3′ end of a gene • Use caution with splice variants that are predicted to lead to exon skipping but leave the remainder of the protein intact • Use caution in the presence of multiple transcripts
Strong	PS1 Same amino acid change as a previously established pathogenic variant regardless of nucleotide change Example: Val→Leu caused by either G>C or G>T in the same codon Caveat: Beware of changes that impact splicing rather than at the amino acid/protein level PS2 De novo (both maternity and paternity confirmed) in a patient with the disease and no family history Note: Confirmation of paternity only is insufficient. Egg donation, surrogate motherhood, errors in embryo transfer, and so on, can contribute to nonmaternity. PS3 Well-established in vitro or in vivo functional studies supportive of a damaging effect on the gene or gene product Note: Functional studies that have been validated and shown to be reproducible and robust in a clinical diagnostic laboratory setting are considered the most well established. PS4 The prevalence of the variant in affected individuals is significantly increased compared with the prevalence in controls Note 1: Relative risk or OR, as obtained from case–control studies, is >5.0, and the confidence interval around the estimate of relative risk or OR does not include 1.0. See the article for detailed guidance. Note 2: In instances of very rare variants where case–control studies may not reach statistical significance, the prior observation of the variant in multiple unrelated patients with the same phenotype, and its absence in controls, may be used as moderate level of evidence.
Moderate	PM1 Located in a mutational hot spot and/or critical and well-established functional domain (e.g., active site of an enzyme) without benign variation PM2 Absent from controls (or at extremely low frequency if recessive) (Table 6) in Exome Sequencing Project, 1000 Genomes Project, or Exome Aggregation Consortium Caveat: Population data for insertions/deletions may be poorly called by next-generation sequencing. PM3 For recessive disorders, detected in *trans* with a pathogenic variant Note: This requires testing of parents (or offspring) to determine phase. PM4 Protein length changes as a result of in-frame deletions/insertions in a nonrepeat region or stop-loss variants PM5 Novel missense change at an amino acid residue where a different missense change determined to be pathogenic has been seen before Example: Arg156His is pathogenic; now you observe Arg156Cys Caveat: Beware of changes that impact splicing rather than at the amino acid/protein level. PM6 Assumed de novo, but without confirmation of paternity and maternity
Supporting	PP1 Cosegregation with disease in multiple affected family members in a gene definitively known to cause the disease Note: May be used as stronger evidence with increasing segregation data PP2 Missense variant in a gene that has a low rate of benign missense variation and in which missense variants are a common mechanism of disease PP3 Multiple lines of computational evidence support a deleterious effect on the gene or gene product (conservation, evolutionary, splicing impact, etc.) Caveat: Because many in silico algorithms use the same or very similar input for their predictions, each algorithm should not be counted as an independent criterion. PP3 can be used only once in any evaluation of a variant. PP4 Patient's phenotype or family history is highly specific for a disease with a single genetic etiology PP5 Reputable source recently reports variant as pathogenic, but the evidence is not available to the laboratory to perform an independent evaluation

LOF, loss of function; OR, odds ratio.

1.5.2.2 Population databases

When classifying and reporting a variant, clinical laboratories may discover valuable information in databases, as well as in the published literature. However, databases are advantageous to gather information, but also should be used with caution. Population databases (Table 1-5) are commodious for obtaining the frequencies of variants in great populations. Population databases must include healthy and Affected individuals. Therefore, when choosing population databases must determine whether healthy or disease cohorts were used and, if possible, whether more than one individual in a family was included, also the age

range of the subjects.

Table 1-5: Population, disease-specific, and sequence databases(96)

Population databases	
Exome Aggregation Consortium http://exac.broadinstitute.org/	Database of variants found during exome sequencing of 61,486 unrelated individuals sequenced as part of various disease-specific and population genetic studies. Pediatric disease subjects as well as related individuals were excluded.
Exome Variant Server http://evs.gs.washington.edu/EVS	Database of variants found during exome sequencing of several large cohorts of individuals of European and African American ancestry. Includes coverage data to inform the absence of variation.
1000 Genomes Project http://browser.1000genomes.org	Database of variants found during low-coverage and high-coverage genomic and targeted sequencing from 26 populations. Provides more diversity compared to the Exome Variant Server but also contains lower-quality data, and some cohorts contain related individuals.
dbSNP http://www.ncbi.nlm.nih.gov/snp	Database of short genetic variations (typically ≤50 bp) submitted from many sources. May lack details of the originating study and may contain pathogenic variants.
dbVar http://www.ncbi.nlm.nih.gov/dbvar	Database of structural variation (typically >50 bp) submitted from many sources.
Disease databases	
ClinVar http://www.ncbi.nlm.nih.gov/clinvar	Database of assertions about the clinical significance and phenotype relationship of human variations.
OMIM http://www.omim.org	Database of human genes and genetic conditions that also contains a representative sampling of disease-associated genetic variants.
Human Gene Mutation Database http://www.hgmd.org	Database of variant annotations published in the literature. Requires fee-based subscription to access much of the content.
Locus/disease/ethnic/other-specific databases	
Human Genome Variation Society http://www.hgvs.org/dblist/dblist.html Leiden Open Variation Database http://www.lovd.nl	The Human Genome Variation Society site developed a list of thousands of databases that provide variant annotations on specific subsets of human variation. A large percentage of databases are built in the Leiden Open Variation Database system.
DECIPHER http://decipher.sanger.ac.uk	A molecular cytogenetic database for clinicians and researchers linking genomic microarray data with phenotype using the Ensembl genome browser.
Sequence databases	
NCBI Genome http://www.ncbi.nlm.nih.gov/genome	Source of full human genome reference sequences.
RefSeqGene http://www.ncbi.nlm.nih.gov/refseq/rsg Locus Reference Genomic (LRG) http://www.lrg-sequence.org	Medically relevant gene reference sequence resource.
MitoMap http://www.mitomap.org/MITOMAP/HumanMitoSeq	Revised Cambridge reference sequence for human mitochondrial DNA.

In Silico (Computational) Prediction Algorithms

In silico prediction algorithms are often applied the tools to predict whether a nucleotide variation in a gene will change the structure and function of the protein or not (Table 1-6). Although the individual algorithms may vary in their main pathway of risk prediction, they can be divided into two categories:
1- prediction of the intense effect of a missense variant on protein function
2- impact of a sequence variant on splicing(96).

Table 1-6 In silico prediction algorithms more frequently used(96)

Category	Name	Website	Basis
Missense prediction	ConSurf	http://consurftest.tau.ac.il	Evolutionary conservation
	FATHMM	http://fathmm.biocompute.org.uk	Evolutionary conservation
	MutationAssessor	http://mutationassessor.org	Evolutionary conservation
	PANTHER	http://www.pantherdb.org/tools/csnpScoreForm.jsp	Evolutionary conservation
	PhD-SNP	http://snps.biofold.org/phd-snp/phd-snp.html	Evolutionary conservation
	SIFT	http://sift.jcvi.org	Evolutionary conservation
	SNPs&GO	http://snps-and-go.biocomp.unibo.it/snps-and-go	Protein structure/function
	Align GVGD	http://agvgd.iarc.fr/agvgd_input.php	Protein structure/function and evolutionary conservation
	MAPP	http://mendel.stanford.edu/SidowLab/downloads/MAPP/index.html	Protein structure/function and evolutionary conservation
	MutationTaster	http://www.mutationtaster.org	Protein structure/function and evolutionary conservation
	MutPred	http://mutpred.mutdb.org	Protein structure/function and evolutionary conservation
	PolyPhen-2	http://genetics.bwh.harvard.edu/pph2	Protein structure/function and evolutionary conservation
	PROVEAN	http://provean.jcvi.org/index.php	Alignment and measurement of similarity between variant sequence and protein sequence homolog
	nsSNPAnalyzer	http://snpanalyzer.uthsc.edu	Multiple sequence alignment and protein structure analysis
	Condel	http://bg.upf.edu/fannsdb/	Combines SIFT, PolyPhen-2, and MutationAssessor
	CADD	http://cadd.gs.washington.edu	Contrasts annotations of fixed/nearly fixed derived alleles in humans with simulated variants
Splice site prediction	GeneSplicer	http://www.cbcb.umd.edu/software/GeneSplicer/gene_spl.shtml	Markov models
	Human Splicing Finder	http://www.umd.be/HSF/	Position-dependent logic
	MaxEntScan	http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html	Maximum entropy principle
	NetGene2	http://www.cbs.dtu.dk/services/NetGene2	Neural networks
	NNSplice	http://www.fruitfly.org/seq_tools/splice.html	Neural networks
	FSPLICE	http://www.softberry.com/berry.phtml?topic=fsplice&group=programs&subgroup=gfind	Species-specific predictor for splice sites based on weight matrices model
Nucleotide conservation prediction	GERP	http://mendel.stanford.edu/sidowlab/downloads/gerp/index.html	Genomic evolutionary rate profiling
	PhastCons	http://compgen.bscb.cornell.edu/phast/	Conservation scoring and identification of conserved elements
	PhyloP	http://compgen.bscb.cornell.edu/phast/	
		http://compgen.bscb.cornell.edu/phast/help-pages/phyloP.txt	Alignment and phylogenetic trees: Computation of *P* values for conservation or acceleration, either lineage-specific or across all branches

In silico tools/software prediction programs used for sequence variant interpretation.

Commonly, the accurate of the majority algorithms to predict missense variant are 65–80% for examining known disease variants (97).

The in-silico tools often used for interpretation of missense variant in clinical laboratories include PolyPhen2, SIFT, and Mutation Taster (98).

Using of numerous software programs to interoperate sequence variant is also recommended because each programs have their own strengths and weaknesses, depending on the algorithm. The bioinformatic tools help to only prediction and it is recommended to use other source to make clinical assertion.

1.6 Target therapy in Breast Cancer and ovarian cancer

Personalized medicine has tried to utilize targeted therapies with increased choosing and efficiency in pre-determined sufferer cohorts. Personalized cancer

medicine comes from studies of human genes and the genes in varied cancers. Throughout the past decade, a substantial the research was done in the field of oncology has been helped with the development and use of targeted therapy. These studies have contributed scientist design further effective treatments. They have utilized genetic data to develop tests for malignancy and ways to prevent it as well(99).

Personalized cancer medicine helped to understand which patients are appropriate candidates for specific treatments (immunotherapy and PARP inhibitors are good examples). These days, an array of trials is underway to identify targeted therapies for patients with chemotherapy-resistant and/or advanced-stage disease(99).

Breast cancer is one of the first malignancies in which targeted therapies have been used successfully. Personalized cancer medicine has fewer side effects than other kinds of therapy because it is designed to be more specific and also personalized treatment may affect healthy cells less than cancer cells.

Regarding to the histological and molecular characteristics of breast cancer distinguished four subtypes of breast cancer: luminal A and luminal B (expressing the estrogen receptor[1]), basal-like and human epidermal growth factor receptor 2[2] - enriched (with-out ER expression)(99).

Now, our understanding of resistance to targeted therapy is becoming more nuanced. Despite these advances in the management of HR-positive and HER2-over expressing tumors, relapse and the development of metastatic disease are still pivotal problems, and the final common pathway for patients with HR-positive and HER2- positive metastatic breast cancer is the development of resistance to targeted therapy and continued progression of the malady(100).

The resistance to endocrine therapy or anti-HER2 therapy can be either intrinsic or new, that the tumor never responds to the therapy, or—more often—acquired resistance, in which the response sunsets over time and cancer eventually progresses.

Advanced (metastatic) breast cancer is a treatable disease. The most important aim of therapy of it is to prolong survival and control symptoms with low treat-

[1] - ER
[2] - HER2

ment-associated toxicity to maintain or improve quality of life (improved quality-adjusted life expectancy). In result investigate further targets, could specifically eliminate malignant cells with minor effect on nonmalignant cell.

1.6.1 PARP Inhibitors

PARP[1] a protein that is involved in the repair of single-strand breaks[2] of DNA. Different studies have shown that HR-deficient cells (e.g. those have BRCA mutations) are acutely sensitive to pharmacological inhibition of Poly (ADP-ribose) polymerase, that results in stalled and collapsed replication forks.

When PARP is inhibited, unrepaired SSBs can be a cause double-strand break[3]. DSBs induced by PARP-inhibition can result in lethal DNA damages in HR-deficient cells due to DSBs of DNA is repaired usually by homologous recombination pathway[4](figure 1-29).

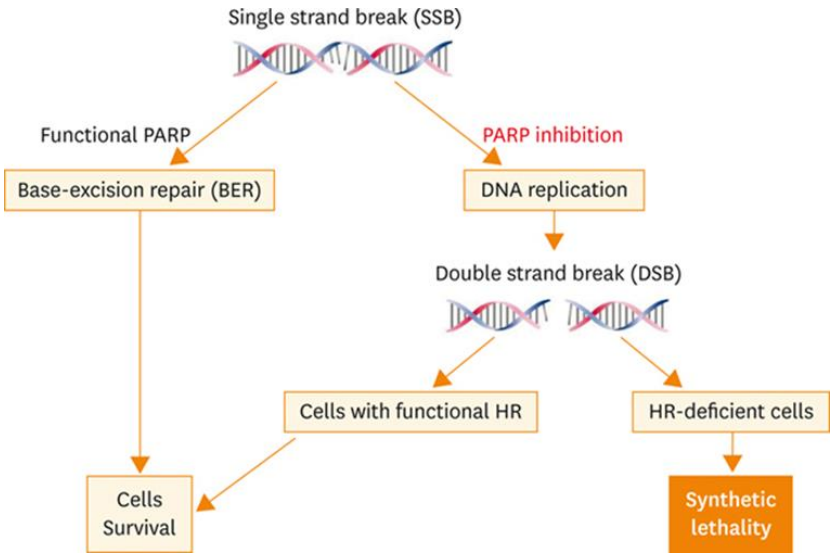

Figure 1-29. Principles of poly (ADP-ribose) polymerase (PARP) inhibition in cancer treatment. HR, homologous recombination(101)

Initial mechanism underlying the SL[5] interaction was that PARP inhibition caused persistent SSBs[6], when encountered by a replication fork sometimes resulted in the collapse of the fork, creating a DSB(102).

Modified initial model showed that some PARPi (especially rucaparib, Olaparib,

[1] - Poly (ADP-ribose) polymerases
[2] - SSBs
[3] - DSBs
[4] - HR
[5] - synthetic lethal
[6] - single-strand breaks

niraparib and talizumab) "trap" PARP1 on DNA, preventing auto Arylation and PARP1 release from the site of damage and in a result interfering with the catalytic cycle of PARP1(103).

This trapped PARP1 protein has been suggested to have a situation analogous to the mechanism of action of cancer drugs that inhibit Topoisomerase II, which also "trap" a DNA repair protein on the double helix (figure1-30).

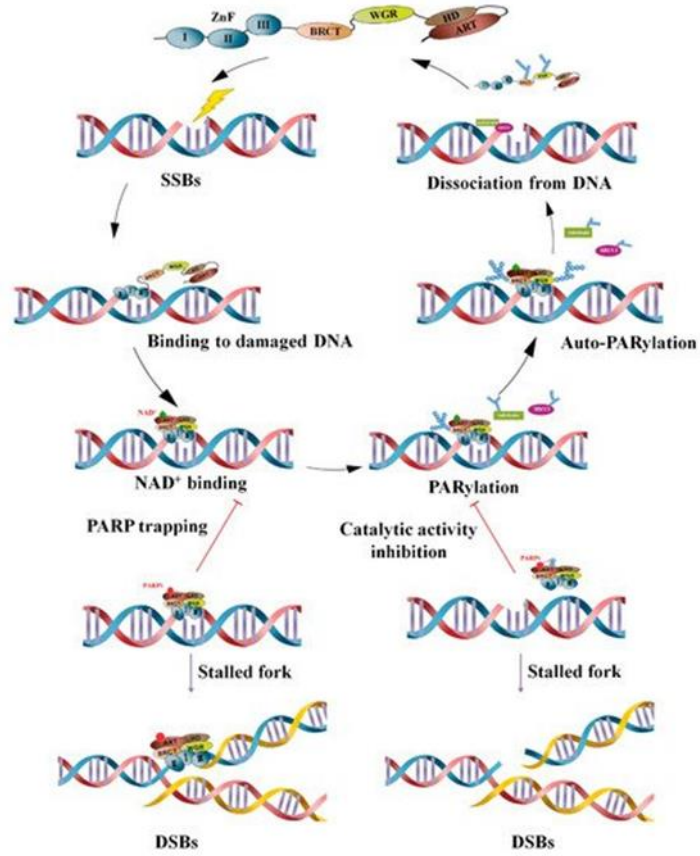

Figure 1-30: A model of PARP inhibitor synthetic lethality. trapped PARP1/DNA nucleoprotein complexes destroy the progression of replication forks. The replication fork is eradicated by trapped PARP1(104).

The inhibition of PARP destroys the repair of SSBs via disturbance of the BER pathway. PARP1 trapping conducts to the accumulation of SSB as well, which leads to DSBs at the replication fork and as result, it causes the death of homologous recombination deficient cells(105).

Large-scale cancer genome sequencing projects showed that germline mutations in HRR genes happen in a wide spectrum of cancers(106), including high-grade serous ovarian cancer[1](107), advanced prostate cancer(108) and pancreatic cancer(109).

[1] - HGS-OVCa

The cancers with HRR mutations are good candidates to test PARPi efficacy. Currently, a variety of PARPi is in clinical development such as Olaparib, Veliparib, Niraparib, Rucaparib and Talazoparib. These PARPi are similar in that they all inhibit PARP1 and PARP2 catalytic activity but have differing power in PARP trapping (figure1-31)(110). various research is being done on the benefit of PARPi in other malignancies, including prostate and pancreas cancers (Table 1-7).

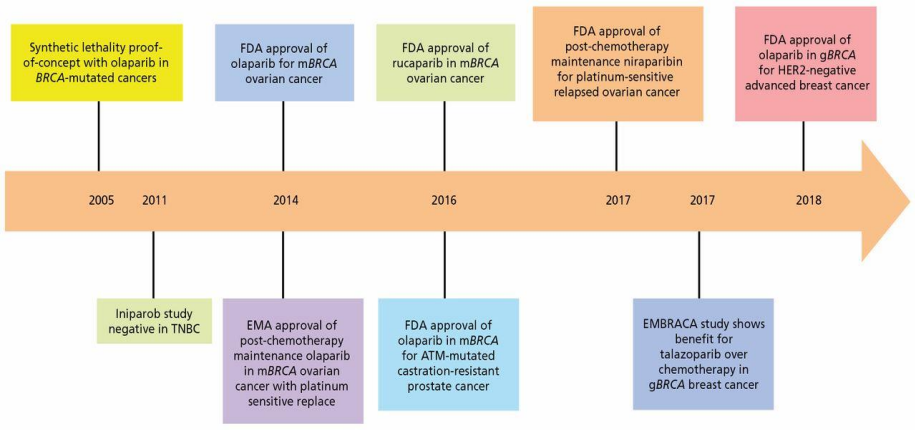

Figure 1-31: Timeline of PARP inhibitor clinical milestones. Abbreviations: ATM, ataxia-telangiectasia mutated; EMA, European Medicines Agency; gBRCA, germline BRCA; mBRCA, BRCA mutated; TNBC, triple-negative breast cancer[1].

Table 1-7. Parp Inhibitors in Phase III Trials.

Drug	Trial Identifier	Setting	Design	Population	Status
Olaparib	OlympiAD[20] NCT02000622	Advanced/Metastatic	Olaparib vs PCT	Advanced/Metastatic g*BRCA*, ≤2 prior lines	Resulted
	OlympiA NCT02032823	Adjuvant	Olaparib vs placebo	Early-stage g*BRCA*, post completion SOC adjuvant therapy	Recruiting
Veliparib	BROCADE 3 NCT02163694	Advanced/Metastatic	C + P + veliparib vs C + P + placebo	Metastatic g*BRCA*, 0–2 lines of prior therapy	Recruiting
	BrighTNess[29] NCT02032277	Neoadjuvant	C + P + veliparib → AC vs C + P + placebo → AC vs Placebo + placebo + P → AC	Neoadjuvant TNBC	Resulted
Talazoparib	EMBRACA[23] NCT01945775	Advanced/Metastatic	Talazoparib vs PCT	Advanced/Metastatic g*BRCA*, ≤3 prior lines	Resulted
Niraparib	BRAVO NCT01905592	Advanced/Metastatic	Niraparib vs PCT	Advanced/Metastatic g*BRCA*, ≤2 prior lines	Not recruiting

Abbrevationg BRCA germline BRCAAC,doxorubicin + cyclophosphamide, C carboplati; P,paclitaxel; SOC, Standartd of care; TNBC Triple-negative breast cancer

[1] - Citation: Journal of the National Comprehensive Cancer Network J Natl ComprCancNetw .

A number of studies have been currently comparing platinum salts to PARPi and evaluating them in combination for advanced breast cancer. PARPis have been widely tested for ovarian cancer treatment in different situations. one half of high-grade serous carcinoma[1], the most common subtype of ovarian cancer, displays defects in the HR DNA repair pathway, with mutations identified in BRCA1/2 in ~22% of the cases with ~15% germline and ~7% of tumoral mutations(108).

Mutations in other HR genes are less usual and are found in approximately 3% of the cases. Also, Sporadic tumors exhibit HR defects as BRCA mutants (the BRCAness phenotype), and in result a higher response rate to platinum-based chemotherapy and PARPi. A majority of the sufferers with advanced-stage ovarian carcinoma are initially treated with platinum-based chemotherapy, but the most of them will ultimately relapse. Longer treatment-free intervals and improved overall survival rates in this group are associated with their inability to DNA damage repairing(111).

olaparib was approved by the Food and Drug Administration[2] in 2014 for the maintenance treatment of BRCA1/2-mutated ovarian cancer. Then it was extended in 2018 to all platinum-sensitive patients, regardless of BRCA1/2 status, due to it was realized that the benefit extended to all HRD tumours. Currently, two other PARPi, niraparib and rucaparib, have been approved by the FDA for the treatment of ovarian cancer(112).

Rucaparib was also approved by the FDA for the maintenance treatment of ovarian cancer(113). As a result, the population with potential benefits from PARPi is likely wider than germline BRCA mutation-associated disease. It is known that a portion of carriers of the primary or secondary mutation shows resistance to the treatment. Resistance mechanisms in cancer-targeted therapies represent the main challenge to current studies.

Combining different pathways in order to provide a targeted treatment with maximum destruction of cancer cells and minimizing the side effects of treatment methods is a promising result of this research. For this reason, biomarkers are in development to broaden the selection of patients, with the potential clinical benefit from these agents.

[1] - HGSOC
[2] - FDA

Chapter 2

Aim of the study

Aims of the Study

No more than 20% of Hereditary breast, ovarian, endometrial cancer are associated with the germline pathogenic[1] or likely pathogenic[2] variants in the BRCA1 and BRCA2 genes(113). Therefore, discovery of novel genes involved in the susceptibility of hereditary cancers is under investigation. Hence, the identification of P/LP variants in other genes of germline or somatic is pivotal for the future of primary prevention strategies (prophylactic surgery and drug prevention), targeted therapy and surveillance programs.

In this research, the study of hereditary breast/ovarian cancer and lynch syndrome susceptibility genes is crucial.

BRCA1/2 have been primarily surveyed for diagnostic aims due to their variations that demonstrated high penetrance, conferring the 5-fold higher risk of breast cancer in patients with P/LP variant compared to the general population(49).

BRCA1 and *BRCA2* have an important role in the DNA double-strand break repair[3] machinery by contribute homologous recombination[4].

This is a conserved mechanism that interact with variety proteins, such as ATM, a master kinase acting upstream in the genome surveillance pathway, mostly activated by contribute DSBs(114); MRN complex (MRE11, NBN, RAD50), could detect DSBs(115); CHEK2 that allows DNA repair by arresting the cell cycle at the G1/S checkpoint(116); BARD1 and BRIP1, that interact with BRCA1 at N and C-terminal regions, respectively(115); PALB2 and the paralog RAD51C and RAD51D, all involved in the BRCA complexes required for HR(117), and LKB1, encoded by STK11, that colocalizes with ATM and BRCA1 at the sites of the DNA damage(62). Before, All of the DBSR genes have been associated with hereditary breast/ovarian and pancreatic cancers, as well as colon[5] and gastric cancers[6] (118).

Moreover, other genes involved in DDR pathways different from HR and cell cycle control, including *APC, CDH1, CDK4, CDKN2A, PTEN, SMAD4, TP53* (119-121), and the DNA helicase RECQL [19], have been related to high or moderate susceptibility

[1] - P
[2] - LP
[3] - DSBR
[4] - HR
[5] - CC
[6] - GC

to familial breast cancer and other kind of malignancies.

Next-generation sequencing studies have newly illustrated that some genes causing hereditary gastrointestinal cancer syndromes are risk factors for ovarian, breast and pancreatic cancers. Indeed, P/LP variants in the mismatch repair genes, including *MSH2, MSH6, MLH1, PMS2*, and *EPCAM*, classically associated to hereditary colon and endometrial cancers (Lynch Syndrome). They have been identified in the biliary, breast, ovarian and gastric tumors as well(122).

MUTYH is a critical gene in the DDR by base excision repair and responsible for the autosomal recessive form of familial colorectal cancer polyposis, also has been proposed as a risk factor for breast cancer in males(123).

One of the main aims of the oncology clinical survey is identification of P/LP variants in DDR and cycle cell genes. Mutations in these genes are becoming as novel purpose for therapy of various cancers and, particularly, for personalized therapies(124).

However, the prevalence of germline mutations in non-BRCA DDR genes is specially investigated in Breast Cancer, ovarian cancer[1], and pancreatic cancer[2], accessible data of the genetic risk factors in cancer disease are still insignificant.

By applying NGS technologies, we analyzed 28 genes involved in DDR and in the cell cycle control in a cohort of 416 patients with personal and/or family history of Breast Cancer, Ovarian cancer, colorectal, endometrial. This study aimed at

1) To identify new mutations in genes involved in Double Strand Break Repair, mismatch repair and in the cell cycle control,

2) To perform the simultaneous analysis of point mutations and CNVs[3] using a single workflow of analysis on next generation sequencing[4] platform and Multiplex Ligation-dependent Probe Amplification[5] panel,

3) To correlate the genotype to the clinical phenotype,

4) To contribute at clarifying the clinical significance of the identified variants, for a better management of surveillance, specific risk reduction measures and therapeutic options, in families with breast, ovarian cancer and other cancers related to Lynch Syndrome

[1] - OC
[2] - PaC
[3] - Copy number varations
[4] - NGS
[5] - MLPA

Chapter 3

Material and Methods

Chapter 3

Material and Methods

3.1 Patients selection and enrollment

From December 2019 to March 2022, 416 (356 F/60 M) unrelated patients with personal and/or familial history of Breast Cancer[1], Ovarian cancer[2], colorectal[3], endometrial cancer attended the UOC of Medical Genetics and Advanced Cellular Diagnostics of the Department of Clinical and Molecular Medicine (Sant'Andrea University Hospital of Rome) for Multi panel gene (28 DDR gene) molecular testing, according to the American National Comprehensive Cancer Network[4] Guidelines. The Participants with condition following were selected for this study:

1- Individuals satisfied the NCCN testing criteria for the multigene panel (https://www.nccn.org/professionals/physician_gls/default.aspx)

2- If the patients are referred by the doctor for BRCA1/2 test, the result of the test was without P/LP BRCA variants.

3- All participants signed the written informed consent

After conducting dedicated genetic counseling, precise clinical data of participants were collected by medical records as well as personal interviews. The molecular analysis of 28 cancer-related genes (*APC, ATM, BARD1, BMPR1, ABRIP1, CDH1, CDK4, CDKN2A, CHEK2, EPCAM, MLH1, MEN1, MRE11, MSH2, MSH6, MUTYH, NBN, PALB2, PMS2, PTEN, RAD50, RAD51C, RAD51D, RECQL, RET, SMAD4, STK11, TP53*), performed with a multigene cancer panel. The perusal complied with the ethical standards of the Declaration of Helsinki and also was reviewed and approved by the institutional ethics committee. Written informed consent was collected from all participants.

3.2 NGS Sequencing analysis

Genomic DNA of each patient was extracted from peripheral blood using QuickGene DNA whole blood kit S (Kurabo, Japon) and QuickGene-810 (Hereinafter QG-810)'s DNA extraction machine (figure 3-1).

[1] - BC
[2] - OC
[3] - CC
[4] - NCCN

The quantified was calculated by Qubit ds DNA HS Assay Kit on Qubit 2.0 Fluorimeter (Invitrogen, Carlsbad, CA, United States) according to the manufacturer's instructions.

A multigene panel, including 28 genes involved in DNA damage repair pathways as DSBR and MMR, as well as in the cell cycle control was designed by Ion Ampliseq designer software (Version 7.0, Life Technologies, Carls-bad, CA, United States; https://www.ampliseq.com/login/login.action).

Figure 3-1 : QuickGene-810 (Hereinafter QG-810)'s DNA extraction machine

The selection of genes was based on their association with hereditary cancer predisposition. All the selected genes (*APC, ATM, BARD1, BMPR1, ABRIP1, CDH1, CDK4, CDKN2A, CHEK2, EPCAM, MEN1, MRE11, MSH2, MSH6, MUTYH, NBN, PALB2, PMS2, PTEN, RAD50, RAD51C, RAD51D, RECQL, RET, SMAD4, STK11, TP53*) are indeed considered cancer-predisposing genes with high or moderate penetrance based on the relative risk for BC, OC, colorectal, endometrial cancer and other malignancies that their damaging mutations confer in carriers(49, 64) (Table 3-1). The panel contains 610 primer pairs in two pools, covering the exons and exon-intron boundaries (Table S-1).

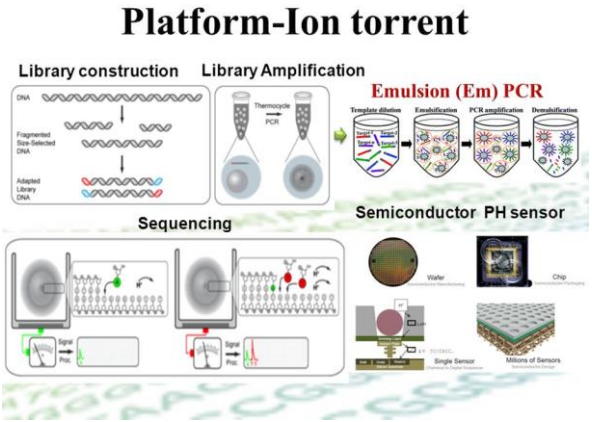

Figure 3-2: Platform-Ion torrent

Table 3-1: DNA damage repair (DDR) genes analyzed in this study

Panel Gene	*Syndrome*	*main pathway*	*Cancer related*	*Reference*
ATM	Ataxia Telangiectasia (AR)	DOUBLE STRAND BREAKS REPAIR (Homologous Recombination)	Breast, ovarian, pancreatic	[2], [8], [12]
PALB2	Fanconi Anemia (AR)			[8], [12], [13]
MRE11A	Nijmegen breakage syndrome-like disorder (AR)		Breast	[9], [13]
RAD50				[9], [87]
BARD1				[10]
NBN	Nijmegen Breakage Syndrome (AR)		Breast, ovarian	[12]
BRIP1	Fanconi Anemia (AR)			[11]
RAD51C				[13]
RAD51D				[13]
STK11	Peutz-Jeghers Syndrome (AD)		Colorectal, breast, pancreatic, gastric, small intestine, cervical, ovarian	[2], [21]
MSH2 *MLH1* *MSH6* *PMS2* *EPCAM*	Lynch Syndrome (AD)	MISMATCH REPAIR	Colorectal, endometrial, ovarian, gastric, urothelial, pancreaticobiliary, cutaneous sebaceous neo- plasms, brain	[2], [12]
MUTYH	MYH-Associated polyposis (AR)	BASE EXCISION REPAIR	Colorectal, duodenal, breast	[84]
RECQL		DNA REPAIR (helicase)	breast	[20]
TP53	Li-Fraumeni Syndrome (AD)	cell cycle control	Breast, sarcoma, brain, adrenocortical, leukemia, gastric	[2], [12], [13]
PTEN	Cowden Syndrome (AD)		Colorectal, breast, endometrial, thyroid, renal	[2], [8], [80]
CHEK2	Li-Fraumeni variant (AD)		Breast ovarian	[2], [12], [73]
CDH1	Hereditary diffuse gastric cancer (AD)		Gastric, breast	[2], [12], [13]
CDK4	Familial melanoma (AD)		Melanoma	[88]
CDKN2A			Melanoma, pancreatic	[2]
SMAD4	Juvenile polyposis Syndrome (AD)		Colorectal, Gastric	[14]
APC	Familial adenomatous polyposis (AD)		Colorectal, small intestine, ampullary, gastric, desmoid, thyroid	[6]

List of the 25 genes selected for this study. AD: autosomal dominant; AR: autosomal recessive.

The Ion GeneStudio™ S5 series is a modern path of benchtop next-generation sequencing[1] systems that contributes to run different projects across multiple research applications, with the simplest sample-to-data NGS workflow and highest speed. These systems offer the chance to run wide-ranging experiments on a single platform by flexibility ability to choose from five Ion Torrent™ chips.

1 - NGS

Figure 3-3: Ion Chef™ Instrument and Ion genestudio s5 system

Together with Ion AmpliSeq™ technology to select target and the Ion Chef™ System for automated library and template preparation, the Ion GeneStudio S5 series helps to streamline targeted NGS workflow(figure 3-3). According to the manufacturer's protocol of emulsion PCR using Ion genestudio S5 system, at first libraries were automatically prepared based on protocol of Ion AmpliSeq™ Kit for Chef DL8 (Thermo Fisher Scientific, Carlsbad, CA, United States) (Figue 3-4) on Ion Chef™ Instrument (Thermo Fisher Scientific, Carlsbad, CA, United States) and 1- or 2-pool Ion AmpliSeq™ primer panels. Then the produced high-quality Ion Sphere™ particles were used in the Ion genestudio S5 system system (Thermo Fisher Scientific, Carlsbad, CA, United States) based on Ion 510™ & Ion 520™ & Ion 530™ Kit – Chef (Thermo Fisher Scientific, Carlsbad, CA, United States)(figue3-5). This kit accommodates 2 sequencing runs per initialization. For 1 sequencing run per initialization (required for 400 base-read applications).

Figure 3-4. Ion AmpliSeq™ Kit for Chef DL8 . Utilizing plug and play, pre-packaged, single-use cartridges and PCR plates for a fully automated workflow. Supplied with barcodes conveniently dried-down in 96-well plates, multiplexing samples is easy and requires no additional pipetting steps.

Figure 3-5. The Ion 510 & Ion 520 & Ion 530 Kit contains pre-packaged single-use template and sequencing reagent cartridges with integrated sample tracking, delivering an automated workflow for sequencing of 8 loaded chips with complete run traceability.

The prepared libraries were sequenced on Ion genestudio S5™ platform (Thermo Fisher Scientific, Carlsbad, CA, USA) using Ion 510™(2-3m reads) , 520™(3-6m reads) and 530™(15-20m reads) Chip (Table 3-2 and 3-4). Sequencing data analysis was performed using Torrent Suite version 5.0.5 and Ion Reporter version 5.6 (Thermo Fisher Scientific, Carlsbad, CA, United States). Five Ion Torrent™ sequencing chips achieve 2–20 M reads per run, the mean read length being up to 275 bp size(Table 3-2).

The average read depth per sample was 11M reads (200bp), with a

mean percentage of reads on target of >99%. Data analysis was performed by Ion Reporter Server System v5.12 (Thermo Fisher Scientific, Carlsbad, CA, USA) and visually confirmed with the Integrative Genomics Viewer (IGV, https://igv.org/, Broad Institute and the Regents of the University of California, CA, USA). All variants reported in this study, identified by NGS technology, were validated by Sanger sequencing and MLPA[1]

Table 3-2. Five Ion Torrent™ sequencing chips achieve 2–130 M reads per run (or 2–260 M reads per day) to enable a broad range of sequencing applications

	Ion 510™ Chip	Ion 520™ Chip	Ion 530™ Chip	Ion 540™ Chip	Ion 550™ Chip
Max. output (reads)	3 M	6 M	20 M	80 M	130 M
Targeted DNA sequencing ** e.g., Ion Torrent™ Oncomine™ Focus Assay	•	•	•	•	•
Small genome sequencing† e.g., Bacterial typing using Ion Xpress™ Plus Fragment Library Kit		•	•		
16S metagenomics sequencing†† e.g., Ion 16S™ Metagenomics Kit		•	•		
Exome sequencing e.g., Ion AmpliSeq™ Exome Panel				•	•
Targeted RNA sequencing e.g., Ion AmpliSeq™ made-to-order RNA panels	•	•	•	•	•
miRNA/small RNA profiling e.g., Ion Total RNA-Seq v2 Kit	•	•	•		
Targeted transcriptome sequencing e.g., Ion AmpliSeq™ Transcriptome Human Gene Expression Kit				•	•
Whole transcriptome sequencing e.g., Ion Total RNA-Seq v2 Kit				•	•
Low-pass whole genome sequencing (PGS) e.g., Ion ReproSeq™ PGS Kit	•	•	•		

** Assumes up to 275 bp insert size. Optimal chip selection based on the size of the panel.
† Assumes 600 bp sequencing only.
†† Assumes 400 bp sequencing only.

[1] - Multiplex Ligation Dependent Probe Amplification

Table 3-3. Ion GeneStudio S5 series specifications

Chip type	Number of reads	Read length (output*)	Ion GeneStudio™ S5 System	Ion GeneStudio™ S5 Plus System	Ion GeneStudio™ S5 Prime System
			Turnaround time (sequencing run** plus analysis time)		
Ion 510 Chip	2–3 million	200 bp (0.3–0.5 Gb)	4.5 hr	3 hr	3 hr
		400 bp (0.6–1 Gb)	10.5 hr	5 hr	5 hr
Ion 520 Chip	4–6 million	200 bp (0.6–1 Gb)	7.5 hr	3.5 hr	3 hr
		400 bp (1.2–2 Gb)	12 hr	5.5 hr	5.5 hr
	3–4 million	600 bp (0.5–1.5 Gb)	12 hr	5.5 hr	5.5 hr
Ion 530 Chip	15–20 million	200 bp (3–4 Gb)	10.5 hr	5 hr	4 hr
		400 bp (6–8 Gb)	21.5 hr	8 hr	6.5 hr
	9–12 million	600 bp (1.5–4.5 Gb)	21 hr	8 hr	7 hr
Ion 540 Chip	60–80 million	200 bp (10–15 Gb)	19 hr	10 hr	6.5 hr
		200 bp (20–30 Gb) 2 runs in 1 day	NA	20 hr	10 hr†
Ion 550 Chip	100–130 million	200 bp (20–25 Gb)	NA	11.5 hr	8.5 hr
		200 bp (40–50 Gb) 2 runs in 1 day	NA	NA	12 hr†

* Expected output with >99% aligned or measured accuracy. Output dependent on read length and application.
** Sequencing run times are between 2.5 and 4 hr.
† Analysis of first run occurs concurrently with the second sequencing run.

3.3 Variants' Classification

The assessment of the identified variants was according to exist evidence of the scientific literature, on gene-specific databases by LSDBs[1] (https://grenada.lumc.nl/LSDB_list/lsdbs) and moreover by consulting ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/), Varsome (https://varsome.com), dbSNP (https://www.ncbi.nlm.nih.gov/snp/), Genome Aggregation Database[2] (https://gnomad.broadinstitute.org), Exome Aggregation Consortium[3].

The clinical classification of the variants was done based on the American College of Medical Genetics and Genomics[4] recommendations with the 5-tier system following(96):

- benign (B),
- likely benign (LB),
- variant of uncertain significance (VUS),
- likely pathogenic (LP),
- pathogenic (P) (Table 3-4)

[1] - Locus-Specific Mutation Databases
[2] - gnomAD
[3] - ExAC
[4] - ACMG

Table 3-4. Scheme For Autosomal Dominant And X-Linked Mendelian Diseases (V2-20-17)

CLASS	AMBRY CLASSIFICATION	CATEGORY	CRITERIA	EXCEPTIONS (NEW BASELINE CLASS)
5	Pathogenic Mutation	A 1 needed	Alterations resulting in premature truncation (e.g.reading frame shift, nonsense)	Truncation in close proximity to 3' terminus (3/4 gene specific)
				LOF has not been established as mechanism of pathogenicity (e.g. MYH7) (3)
			Other ACMG-defined mutation (i.e. initiation codon or gross deletion)	In-frame gross deletion of a single exon not in a known protein functional domain (3), Initiation codon that is not well conserved or possible alternate start (3/4), LOF has not been established as a mechanism of pathogenicity (3)
		B 4 needed	Confirmed *de novo* alteration	Weight is gene and disease specific (2B, 1B, 1C)
			Functionally validated splicing mutation	In-frame skipping of a single exon not in a known protein functional domain. LOF has not been established as a mechanism of pathogenicity. Biologically relevant, naturally occurring, in-frame isoforms with data supporting normal function (3)
			Significant disease association in appropriately sized case-control study(ies)	
			Detected in an individual satisfying established diagnostic critera for classic disease without a clear mutation	
			Last nucleotide of exon	Last nucleotide poorly conserved or non-G, *in silico* not consistent with U2-dependent intron
			Good segregation with disease	1B=(LOD 1.5-3 = 5-9 meioses); 2 or 3B=(LOD >3 = >10 meioses) gene specific
			Deficient protein function in appropriate functional assay(s)	
			Well-characterized mutation at same position	
			Other strong data supporting pathogenic classification	
4	Variant, Likely Pathogenic	3B	Alterations at the canonical donor/acceptor sites (± 1, 2) without other strong (B-level) evidence supporting pathogenicity	LOF has not been established as a mechanism of pathogenicity (3)
		C 4 needed	Rarity in general population databases (dbSNP, ESP, 1000 Genomes)	Dependent on disease penetrance and inheritance pattern
			in silico models in agreement (deleterious) and/or completely conserved position in appropriate species	*in silico* splicing predictions not used as independent line of evidence for last nucleotide of exon
			Moderate segregation with disease (at least 3 informative meioses) for rare diseases	
			Other data supporting pathogenic classification	
			3 of B	
			2 of B and at least 1 of C	
			1 of B and at least 3 of C	
3	VUS		Insufficient or conflicting evidence	
			Gross duplications without strong evidence for pathogenic or benign	
2	Variant, Likely Benign	D 1 needed	Intact protein function observed in appropriate functional assay(s)	
			Intronic alteration with no splicing impact by RT-PCR analysis or other splicing assay	
			Seen *in trans* with a mutation or in homozygous state in individuals without severe disease for that gene	Genes without a defined, severe biallelic phenotype (3)
			Other strong data supporting benign classification	
		E 2 needed	Co-occurence with mutations in the same gene (phase unknown)	Genes without a defined, severe biallelic phenotype (3) When always linked to a the same mutation (can't rule out synergetic effect)
			Co-occurence with mutations in other highly penetrant genes that clearly explain a proband's phenotype	
			Subpopulation frequency in support of benign classification	
			in silico models in agreement (benign)	
			Does not segregate with disease in a family study (genes with incomplete penetrance)	
			No disease association in a small case-control study	
			Other data supporting benign classification	
1	Benign	F 1 needed	General population or subpopulation frequency is too high to be a pathogenic mutation based on disease/syndrome prevalence and penetrance	
			Does not segregate with disease in a family study (genes with complete penetrance)	
			Internal frequency is too high to be a pathogenic mutation based on disease/syndrome prevalence and penetrance	
			No disease association in appropriately sized case-control study(ies)	
			1 of D and at least 2 of E	
			2 or more of D	
			>3 of E w/o conflicting data	
			>4 of E w/conflicting data	

The variant classification scheme is not intended for the interpretation of alterations considered epigenetic factors including genetic modifiers, multifactorial disease, or low-risk disease association alleles and may be limited in the interpretation of alterations confounded by incomplete penetrance, variable expressivity, phenocopies, triallelic or oligogenic inheritance, or skewed X-inactivation.

In addition, the missense prediction programs SIFT (https://sift.bii.a-star.edu.sg) Mutation Taster (http://www.mutationtaster.org), Proven (http://provean.jcvi.org/seq_submit.php), and the splice prediction tool Human Splicing Finder[1] (http://umd.be/Redirect.html) were Independently queried. Gnetic results were considered informative when patients carried LP/P variants, and non-informative[2] when B, LB, or VUS variants were found. Variants were reported by using the Human Genome Variation Society nomenclature guidelines (https://varnomen.hgvs.org/).

[1] - HSF
[2] - NI

3.4 Bioinformatic Analysis

In this research, several bioinformatics software was used to simplify the obtained data. lollipop mutation diagram online software (http://www.bioinformatics.com.cn/) was used to explain and present the results of panel or exome sequencing. Lollipops software generates diffusion-quality, information-dense mutation diagrams for automated pipelines and high-throughput workflows in precision medicine. Automated data combination provides clinical data security, and help to visual discoveries concisely render knowledge with minimal user configuration.

In order to investigation of the chromosome distribution was used SRplot online bioinformatic software (http://www.bioinformatics.com.cn/). Therefore, each chromosome separated rated into many bins, the number of SNPs located within each bin was calculated and plotted.

Also, in order to find the genotypic and phenotypic relationship between the SNPs observed in this study and the interpretation of genetic associations was applied online bioinformatic software '' snpXplorer v2.0 ''(https://snpxplorer.net)(figure3-6).

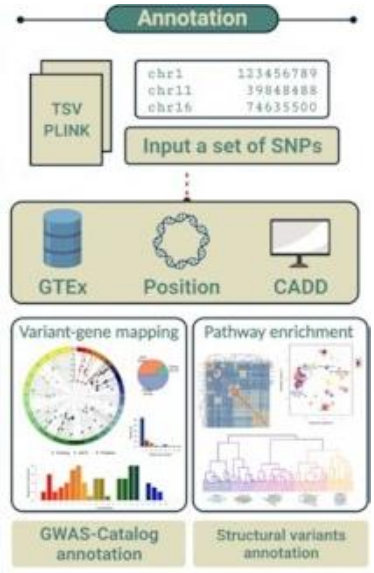

Figure 3-6: functional annotation pipeline. The functional annotation pipeline includes a two-step procedure: firstly, variant-gene mapping (genetic variants are linked to likely affected genes), secondly is gene-pathway mapping (the likely affected genes are tested for pathway enrichment).
(i) associating a variant to a gene when the variant is annotated to be coding by the Combined Annotation Dependent Depletion (CADD, v1.3), (ii) annotating a variant to genes based on found expression-quantitative-trait-loci (eQTL) from GTEx (v8, with the possibility to select the tissue(s) of interest) or (iii) mapping a variant to genes that are within distance d from the variant position, starting with d≤50kb, up d≤500kb , increasing by 50 kb until at least one match is seen[1].

[1] - https://snpxplorer.net.

3.5 Statistical Analysis

A comparison of demographic and clinical variables between groups was done with unpaired t-test and Fisher test for continuous and categorical data, respectively. The p-values lower than or equal to 0.05 were considered statistically significant (figure 3-7).

Figure 3-7 . NGS workflow

Chapter 4

Result

4. Results

From December 2019 to March 2022, a total of 416 (356 F/60 M) eligible patients, with personal and/or familial history of Breast Cancer[1], Ovarian cancer [2], colorectal[3], endometrial cancer who satisfied the National Comprehensive Cancer Network[4] testing criteria for the multigene panel, were included (Table 4-1). 289 patients with breast cancer participating in this study were negative for BRCA 1/2 panel test (table 4-3, Table S2, S4).

38 probands were recommended to do MLPA test to investigate copy number variations. two Probands P220 and P230 were positive for MLPA test; however, P220 with LS and P230 with BC were negative for multi-panel test.

Of 298 patients with BC, 276 were females with BC (~92%) (Mean age range: 53.82 ± 28.5 (27–84)), that 19 of them with bilateral Breast Cancer[5] and 30 of 298 cases suffered at least one other cancer, 22 were males with MBC[6] (~8%) (67.71 ± 21.5.8(49–92). 60 patients were diagnosed with CC (64.23 ± 40.5 (40–119)), 41 had Lynch syndrome[7] (65.38 ± 40.5(38–119), twenty-nine had OC (55.92 ± 21(39–81)), twenty had pancreatic cancer [8] (58.11 ± 21 (34–76)), 12 patients had Gastric cancer[9] (58.61 ± 27.5 (29–84)), eight of 416 patents has Melanoma[10] (49.57 ± 9(42–60)), uterus cancer[11] (60.6 ± 15(43–73)) and Thyroid cancer[12] (53.25 ± 10(43–63)) and two patients were diagnosed for prostatic cancer[13] (54–72). In addition, six of 416 patients had Multiple tumors (such as: Lung cancer, brain cancer, colon cancer, breast cancer and, other malignancy), The remaining 13 cases, had other types of cancers (Mullerian sarcoma, Carcinoid, gastroesophageal cancer, peritoneal carcinosis, leukemia myeloid acuta, hemangiopericytoma) and/or malignancies in their relatives. Overall, 61/416 cases referred a positive family history for different types of malignancies in one or more relatives (Table 4-1). 35 of the patients had 2 types of cancer simultaneously.

[1] - BC
[2] - OC
[3] - CC
[4] -NCCN
[5] - BBC
[6] - Male breast cancer
[7] - LS
[8] - PaC
[9] - GaC
[10] - Mel
[11] - UC
[12] - TC
[13] - PrC

Table 4-1: Main clinical variables of the study sample.

	npts (F/M)			Mean Age Tot ± SD (Min–Max)	Family history 1st, 2nd degree
BC	298	Female BC(276)	BBC (19)	53.82 ± 28.5 (27–84)	53 (BC, GIC, LS, OC, GaC, PaC, CARCINOID, other C)
			fBC (257)		
		Male BC (22)	MBBC (0)	67.71 ± 21.5.8(49–92)	
			MBC (22)		
CC	60 (42/18)		LS : 39 (27/13)	64.23 ± 21.5 (40–83)	11 (BC ,BrC, Liver C, GIC, CC , TC, Kidney, UC, AML, UC, PaC, PrC, LC, bladder C)
OC	29(29/0)		LS : 1 (1/0) (OC,BC)	55.92 ± 21(39–81)	5 (OC , PrC, GaC , BrC , GECa , AML, BC, OC)
PaC	20(11/9)			58.11 ± 21 (34–76)	2(PaC)
GaC	12 (7/5)			58.61 ± 27.5 (29–84)	2(GaC)
MEL	9(9/0)			49.57 ± 9(42–60)	1(bladder C, Kidney C, BC)
UC	10 (9/0)		LS: 1(1/0)	60.6 ± 15(43–73)	4(CC, PrC, LC, Kidney, Bladder C, LS, PaC, BC)
TC	8(6/3)			53.25 ± 10(43–63)	
PrC	2(0/2)			63 ±9(54–72)	1(Kidney, UC, AML)
Multiple tumor	6(4/2)			62.2 ± 19(42–80)	1 (NEOPLASTICS)
OTHER CANCER	17(11/6)			51 ± 30.516–77)	4 (AT, GIC, CC, BC,PaC, Liver C, LC)
Total	416 (356/60)			56.29 ± 51.5 (16–119)	84 (BC, CC, GaC, OC, UC, AT, GIC, LS, Pac, PrC, others)

BC, breast cancer; BBC, bilateral Breast Cancer; MBC, male breast cancer; OC, ovarian cancer; PaC, pancreatic cancer; PrC, prostatic cancer; TC, Thyroid cancer; Mel, Melanoma; GaC, Gastric cancer; CC, Colon Cancer; UC, uterus cancer; LS, lynch syndrome; GIC, gastrointestinal cancer; GECa ,gastroesophageal Ca; PCBA, papillary carcinoma in benign adenoma; BrC, Brain cancer; LC, Lung Cancer; So, Osteosarcoma; Sm, Mullerian's sarcoma; Mel, Melanoma; AM, apocrine metaplasia; dML, diffuse mesenteric leyomiomatosis; AML, acute myeloid leukemia; NET, the neuroendocrine tumor of the pancreas; AT, Ataxia-telangiectasia; C, cancer; pts, patients; F, female; M, male; Fam, familiarity;

4.1 Multi-gene Panel Results

298 of 416 cases who were tested for BRCA1/2 panel being BRCA1/2-negative index (benign/ likely benign) (~89.59%), 31 has VUS for this index (~10.5%).

The result of the multi gene panel testing (Reflex Panel) with NGS assay for 416 sufferers showed that 39 patients (~10%) carrier LP/P variants (35 mutations) (Table 4-2 & table 4-3), 161 were carrier of VUS (145 VUS (~35.03%), 16 VUS+LP/P) (~38.70%) (177 variations) (Table S2, table S3). In two probands were detected two pathogenic big large genomic rearrangement. The remaining 230 (~55%) had no damaging mutations or VUS (1339 variants) (Figure 4-1).

Thirty-five LP/P mutations based on NGS included seven missense (20%), 12 (34.3%) nonsense, 9 (28.57%) Frameshift Deletion, 6 (17.14%) missense, intronic 3 (8.6%), 1 (2.8%) in frame deletion, Synonymous, splice donor, frameshift insertion and Splicing variation (Table 4-2).

The 37 unique variants (35 base on NGS, 2 based on MLPA) (~2.25%) that were

classified as P/LP found in seventeen genes include five were detected in the *ATM* (~14.3%), five in *CHEK2* (~13.5%), four in *MUTYH* (~10.8%), *MSH6* (~10.8%), *MLH1*(~10.8%), three in *RAD51C* (~8.6%), two in *MSH2* (~5.7%) , and one variation was detected in genes *BARD1* (~2.8%), *CDKN2A* (~2.8%), *NBN* (~2.8%), *PALB2* (~2.8%), *PMS2* (~2.8%), APC (~2.8%), *RAD51D* (~2.8%), *RECQL* (~2.8%), *TP53* (~2.8%), *CDH1* (~2.8%)(Table 4-2, 4-3).

The remaining 177 identified variants were VUS (~11.41%), Twenty-five of which found in *ATM* (~14.20%), eighteen in *APC* (10.22%), twelve in *MSH6* (~6.8%), *MLH1*(~6.8%), ten in *Chek2* (~5.6%) and *MSH2* (~5.6%) , *CDH1* (~5.6%), eight in *BRIP1*(~4.5%), *PMS2* (~4.5%), *BARD1* (~4.5%), seven in *MRE11* (~4%), six in *EPCAM* (~3.4%), *RAD50* (~3.4%), *STK11* (~3.4%), five in *PALB2* (~2.8%), four in *MUTYH* (~2.3%), three in *NBN* (~1.7%), *PTEN* (~1.7%), *RAD51D* (~1.7%), *RET* (~1.7%) , two in *CDKN2A* (~1.1%), and one in *BMPR1A* (~0.5%), *MEN* (~0.5%), *RAD51C* (~0.5%), RECQL (~0.5%), *SMAD4* (~0.5%), *TP53* (~0.5%), *CDK4* (~0.5%) were shown in Table S3.

All the 41 probands had malignancies: 11 had lynch syndrome, sixteen had BC, three had GaC, four had OC, three had multiple tumors, and the remaining five including patients with thyroid, papilloma, Mullerian sarcoma, pancreatic cancer, MAP. Some patients suffered from two or more cancer at the same time. No asymptomatic probands were found with damaging mutations. Clinical features and family history of the patients carrying pathogenic variants identified in this study are summarized in Table 4-3.

Table 4-2: P/LP gene variants detected in the sample based on NGS

gene	*# locus*	*transcript*	*dbsnp*	*cDNA (HGVS)*	*protein*	*type of variants*	class	*GnomAD*
APC	chr5:112178000	NM_000038.5	rs768922431	c.6709C>T	p.Arg2237Ter	nonsense	P	/
ATM	chr11:108115727	NM_000051.3	rs747727055	c.875C>T	p.Pro292Leu	missense	LP	0.000008
	chr11:108151895	NM_000051.3	rs587776551	c.3576G>A	p.Lys1192=	synonymous	P	0.000013
	chr11:108183163	NM_000051.3	rs1555111775	c.5944C>T	p.Gln1982Ter	nonsense	P	/
	chr11:108173606	NM_000051.3		c.5347_5350delGAAA	p.Glu1783ThrfsTer9	Frameshift Deletion	P	/
	chr11:108202168	NM_000051.4	rs587781905	c.7517_7520delGAGA	p.Arg2506ThrfsTer3	Frameshift Deletion	P	/
BARD1	chr2:215645393	NM_000465.4	rs796666047	c.1205C>A	p.Ser402Ter	nonsense	P	0.0000066
CDH1	chr16:68845757	NM_004360.5	rs587780784	c.1003C>T	p.Arg335Ter	nonsense	P	0.0000066
CDKN2A	chr9:21994137	NM_058195.4	rs1060501262	c.193+1G>A	p.?	splice donor	P	/
NBN	chr8:90960070	NM_002485.4		c.1896G>A	p.Trp632Ter	nonsense	P	0
CHEK2	chr22:29091788	NM_007194.3	rs200928781	c.1169A>C	p.Tyr390Ser	missense	LP	0.0000066
	chr22:29121087	NM_007194.3	rs17879961	c.470T>C	p.Ile157Thr	missense	LP	0.004042
	chr22:29106048	NM_007194.3	rs730881687	c.793-1G>A	p.?	Splicing variation	LP	/
	chr22:29091856	NM_007194.4	rs555607708	c.1100delC	p.Thr367MetfsTer15	Frameshift Deletion	P	0.001721
PALB2	chr16:23619185	NM_024675.3	rs876659859	c.3350G>A	p.Arg1117Lys	missense	LP	/
PMS2	chr7:6042143	NM_000535.7	rs36038802	c.478C>T	p.Gln160Ter	nonsense	P	0
MLH1	chr3:37089129	NM_000249.3	rs587778949	c.1852_1854delAAG	p.Lys618del	inframe	P	/

	chr3:37090008	NM_000249.3	rs1057520627	c.1897G>A	p.Glu633Lys	Deletion missense	LP	0.000004
	chr3:37042444	NM_000249.4	rs63751642	c.210_213delAGAA	p.Glu71IlefsTer20	Frameshift Deletion	P	/
RAD51D	chr17:33428224	NM_133629.3	rs786202251	c.562delC	|p.Arg188AspfsTer10	Frameshift Deletion	P	/
RAD51C	chr17:56809908	NM_058216.2	rs587781410	c.1026+4_1026+6delAGT	p.?	**Intronic**	P	0.000033
	chr17:56798178	NM_058216.2	rs587782702	c.904+5G>T	p.?	**Intronic**	LP	0.000013
	chr17:56770093	NM_058216.2	rs730881942	c.93delG	p.Phe32SerfsTer8	Frameshift Deletion	P	0.000033
RECQL	chr12:21643163	NM_032941.2	rs1553398842	c.362_363delTA	p.Cys122LeufsTer43	Frameshift Deletion	P	/
MUTYH	chr1:45797371	NM_001128425.1	rs587778536	c.1147delC	p.Ala385ProfsTer23	Frameshift Deletion	P	0.000064
	chr1:45797835	NM_001128425.1	rs587780751	c.933+3A>C	p.?	**Intronic**	LP	0.000112
	chr1:45798475	NM_001128425.1	rs34612342	c.536A>G	p.Tyr179Cys	missense	P	0.001564
	chr1:45799121	NM_001128425.2	rs121908380	c.312C>A	p.Tyr104Ter	nonsense	P	0.0000066
MSH6	chr2:48030639	NM_000179.2	rs267608078	c.3261_3262insC	p.Phe1088fs	Frameshift Insertion	LP	0.00006
	chr2:48027268	NM_000179.2	rs267608058	c.2150_2153delTCAG	p.Val717fs	Frameshift Deletion	P	/
	chr2:48026727	NM_000179.2		c.1605C>G	p.Tyr535Ter	nonsense	P	0.000004
	chr2:48028053	NM_000179.3	rs63750111	c.2931C>A	p.Tyr977Ter	nonsense	P	/
MSH2	chr2:47657020	NM_000251.2	rs63751108	c.1216C>T	p.Arg406Ter	nonsense	P	/
	chr2:47637291	NM_000251.2	rs63750910	c.425C>G	p.Ser142Ter	nonsense	P	/
TP53	chr17:7579529	NM_000546.5	rs876658483	c.158G>A	p.Trp53Ter	nonsense	P	/

Pathogenic, likely pathogenic variants detected by 28 genes of cancer panel among 416 patients with a history familial/personal of cancer. Abbreviations: dbSNP, Single Nucleotide Polymorphism Database (https://www.ncbi.nlm.nih.gov/snp/); rs, reference SNP; HGVS: Human Genome Variation Society (http://www.HGVS.org/varnomen); GnomAD, Genome Aggregation Database (https://gnomad.broadinstitute.org/); ACMG, American College of Medical Genetics and Genomics; P, pathogenic; LP, likely pathogenic. Variants were annotated according to the current HGVS nomenclature; p.? consequence on protein structure unknown.

Table 4-3: Clinical characteristics of probands with P/PL variants

Sample ID	SEX	*gene*	coding	Class	protein	Other variants (vus)	Diagnosis	Age	Family History of Cancer (n# of 1st and 2nddegree affected relatives)	MLPA/ LGR	RESULT BRCA1/ 2 (NGS)
P1	F	*CHEK2*	c.1169A>C	LP	p.Tyr390Ser		Sm	53	CC , BC	n.a	N
P2	F	*CHEK2*	c.470T>C	LP	p.Ile157Thr		multiple tumors (BC , CC, OC, GaC)	47	PrC , RecC	n.a	N
P3	F	*ATM*	c.875C>T	LP	p.Pro292Leu	c.377C>T, p.Pro126Leu	BC , Mel	44	BC	n.a	N
P4	F	*CHEK2*	c.470T>C	LP	p.Ile157Thr	c.1663A>G, p.Ala49Thr	BC	51	BC	n.a	N
P5	F	*CHEK2*	c.793-1G>A	LP	p.?		CT	47	PrC , OTHER	n.a	N
P6	F	*MSH6*	c.3261_3262insC	P	p.Phe1088fs		CC , (LS)	79	CC, TC	n.a	n.a
P7	M	*MSH6*	c.2150_2153delTCAG	P	p.Val717fs		CC , PrC, (LS)	75	Kidney, UC, AML	n.a	n.a
P8	F	*MSH2*	c.1216C>T	P	p.Arg406Ter		CC , (LS)	44	CC, UC, BC, PaC, Kidney, LS	n.a	n.a
P9	F	*MSH2*	c.1216C>T	P	p.Arg406Ter	c.845G>A, p.Arg282Gln	CC , UC , (LS)	69	Mother of P8	n.a	n.a
P11	M	*ATM*	c.3576G>A	P	p.Lys1192=		PaC , GC	66	PaC	n.a	N
P12	M	*RECQL*	c.362_363delTA	P	p.Cys122LeufsTer43		BC , PaC	72	BC, LC	n.a	N
P13	F	*RAD51C*	c.1026+4_1026+6delAGT	P	p.?	c.5365G>C, p.Val1789Leu	OC	53	OC , PrC, GaC , BrC , gastroesophageal Ca , leukemia	n.a	N
P14	F	*CHEK2*	c.1100delC	P	p.Thr367MetfsTer15		BBC	56	BC	n.a	N
P15	F	*CHEK2*	c.1100delC	P	p.Thr367MetfsTer15		BBC , dML	66	BC	n.a	N
P16	F	*RAD51C*	c.904+5G>T	LP	p.?	c.663A>C, p.?	OC	46	OC	n.a	n.a
P17	F	*NBN*	c.1896G>A	LP	p.Trp632Ter	c.244T>C, p.Tyr82His	BC	30	-	n.a	n.a
P18	M	*ATM*	c.5944C>T	P	p.Gln1982Ter		GaC	65	GaC	n.a	n.a
P19	F	*MUTYH*	c.312C>A	P	p.Tyr104Ter		PaC	53	-	n.a	N
P20	M	*MLH1*	c.1852_1854delAAG	P	p.Lys618del		CC , (LS)	58	-	n.a	n.a
P21	F	*MUTYH*	c.1147delC	P	p.Ala385ProfsTer23		BC , TC	60	-	n.a	N
P22	F	*MSH6*	c.1605C>A	P	p.Tyr535Ter		CC , UC , (LS)	82	CC, PrC, LC	n.a	n.a
P23	F	*RAD51C*	c.93delG	P	p.Phe32SerfsTer8	c.1437C>G, p.His479Gln	BC	48	-	n.a	n.a
P24	F	*PALB2*	c.3350G>A	LP	p.Arg1117Lys		BC	50	-	n.a	N
P25	F	*MUTYH*	c.536A>G	P	p.Tyr179Cys	c.6795A>G, p.Gln2265=	MAP	69	MAP	n.a	n.a
P26	F	*APC*	c.6709C>T	P	p.Arg2237Ter		CC , (LS)	55	CC, TC	n.a	n.a
P27	F	*TP53*	c.158G>A	P	p.Trp53Ter		BC	38		n.a	VUS
P28	F	*MLH1*	c.1897G>A	LP	p.Glu633Lys	c.*591A>AT, p.?	BC	54	BC	n.a	n.a

P29	F	*MUTYH*	c.933+3A>C	LP	p.?	c.144+508G>T, p.?	BC	54	BC	n.a	N
P30	F	*ATM*	c.5347_5350delGAAA	P	p.Glu1783ThrfsTer9	c.*591A>AT, p.?	OC	81	-	n.a	n.a
P31	M	*MUTYH*	c.536A>G	P	p.Tyr179Cys		GaC	50	GaC	n.a	n.a
P31	M	*MSH2*	c.425C>G	P	p.Ser142Ter		GaC	50	GaC	n.a	n.a
P33	M	*PMS2*	c.478C>T	P	p.Gln160Ter	c.1277A>T,p.Gln426Leu c.1960C>A, p.Gln654Lys , c.6293T>C, p.Leu2098Pro	CC , Lymphoma, (LS)	66	BC, PaC, Liver C, LC	n.a	n.a
P34	F	*BARD1*	c.1205C>A	P	p.Ser402Ter		BC	52	BC	n.a	N
P35	F	*RAD51D*	c.562delC	P	p.Arg188AspfsTer10	c.2366C>T, p.Ala789Val c.2378A>C, p.Gln793Pro	OC	41	-	n.a	n.a
P36	F	*RAD51C*	c.1026+4_1026+6delAGT	P	p.?		multiple tumors (CC , CT , LC , BrC)	64	-	n.a	n.a
P37	F	*MSH6*	c.2931C>A	P	p.Tyr977Ter		CC, Mel , UC , (LS)	70	CC, Bladder, BC, Kidney	n.a	n.a
P38	F	*MLH1*	c.210_213delAGAA	P	p.Glu71IlefsTer20	c.2770C>T, p.Arg924Trp	CC , (LS)	53	CC, PaC, BrC	n.a	n.a
P39	M	*CDKN2A*	c.193+1G>A	P	p.?		multiple tumors (CC , CT , LC , BrC)	42	-	n.a	N
P40	F	*ATM*	c.7517_7520delGAGA	P	p.Arg2506ThrfsTer3	c.667A>G, p.Ile223Val	BC	48	BC	n.a	n.a
P41	M	*CDH1*	c.1003C>T	P	p.Arg335Ter	c.465-51T>C, p.?	BC	74	CC , BC	n.a	N
P220	f	*MLH1*	ex 9-19 del	P	p.?	-	CC (LS)	51	CC,BC, mel, EC, LC, PaC	MLH1	N
P230	F	*Chek2*	ex 6–13 duplication	P	p.?	-	BC	43	BC, OC	Chek2	n.a

Abbreviations: BC, breast cancer; OC, ovarian cancer; PrC, prostate cancer; LC, lung cancer; Mel, Melanoma; GaC, gastric cancer; BrC, Brain cancer; CC, Colon Cancer; TC, thyroid cancer; PaC, pancreatic cancer; UC, uterine cancer; dML, diffuse mesenteric leyomiomatosis; C, cancer at a not specified site. BBC , bilateral Breat Cancer; GECa , gastroesophageal Carsinoma; So, Osteosarcoma; Sm, Mullerian's sarcoma; AM, apocrine metaplasia; PCBA , papillary carcinoma in benign adenoma; PCa, peritoneal carcinosis; NET, neuroendocrine tumor of the pancreas; GIC, gastrointestinal cancer; EC, endometrial cancer; ex, exon. B , Benign; VUS, Variant of uncertain significance; P, Pathogenic; LP, likely pathogenic variants. LGR ; Large genomic rearrangement, N; negative. N.A, Not analysied; MAP, Mutyh associated polyposis.

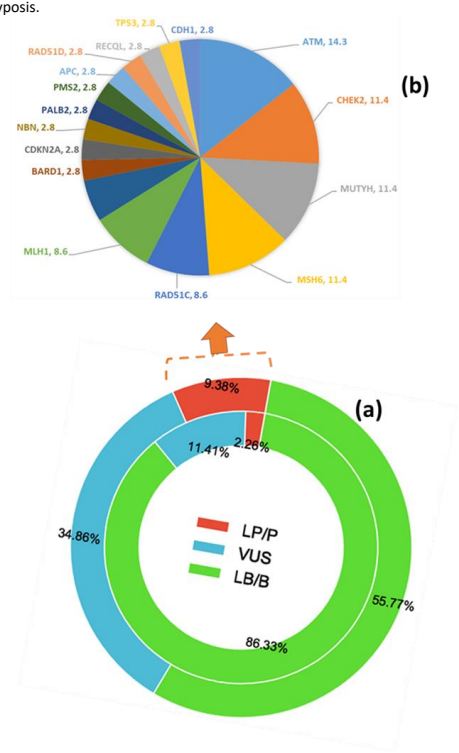

Figure 4-1: Distribution of patients with and without P/LP (P, pathogenic, LP, likely pathogenic variants). (a) The inner circle shows the distribution of variants observed in this study. 35 variants were recognized as LP/P based on NGS. The outer circle is based on the distribution of patients carrying the LP/P, VUS or LB/B pathogenic mutations. 39 out of 416 patients were identified with P/LP variants based on NGS. b: Schematic representation of the percentage of pathogen mutations on the studied genes in the cohort P/LP.[1] P, Pathogenic; LP, Likely Pathogenic; VUS, Variant of uncertain significance; B, benign; LB, Likely benign.

[1] - Two-layer donut plot (a) and SNP density plot were designed by http://www.bioinformatics.com.cn/

4.2 pathogenetic and likely pathogenic mutation

4.2.1 ATM

on ATM gene was identified five LP/P variations in 5 patients (Table 4-2, 4-3)(Figure 4-2):

Figure 4-2: Schematic representation of the ATM Protein and positions of identified P/LP variants. TAN, Tel1/ATM N-terminal motif; FAT, FRAP-ATM-TRRAP domain[1].

The LP ATM variant c.875C>T (p.Pro292Leu) was detected in a 44-year-old woman (P3) (figure 4-3) with a well-differentiated and hormone-responsive in situ ductal carcinoma[2] diagnosed at 43 years, who had removed a melanoma a year before. Her paternal relatives referred to melanoma, osteosarcoma, head-neck, brain, and uterine cancers.

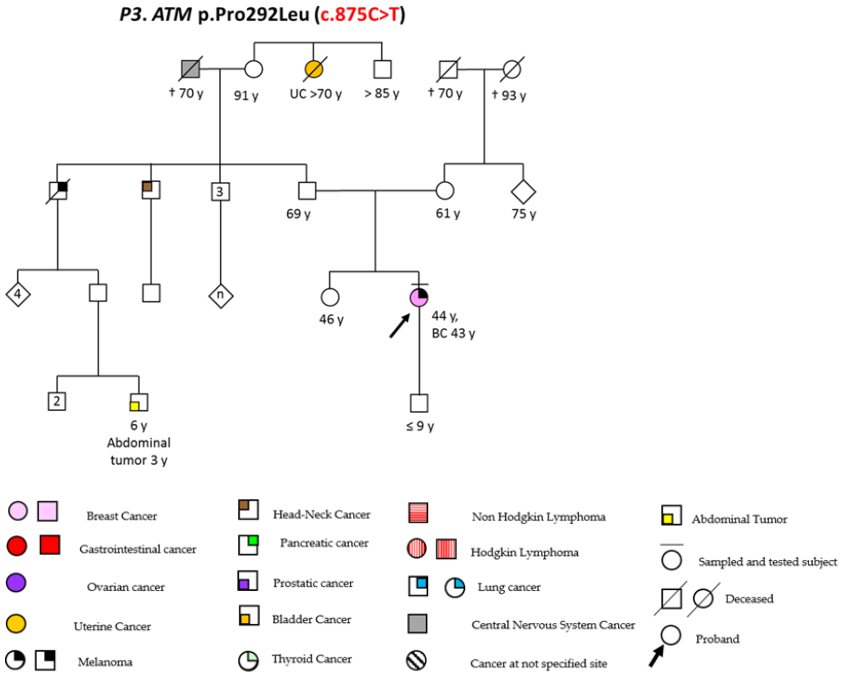

Figure 4-3 : Pedigrees of family with LP c.875C>T variant. Individuals with any cancer are shown as filled circles or square. Arrows show a proband with P/LP variant in ATM family.. The tested subject is indicated with a horizontal line above the circle or square. BC, breast cancer; OC, ovarian cancer CC, colon cancer; UC, uterine cancer; NHL, non-Hodgkin lymphoma; PrC, prostatic cancer; Bil BC, bi-lateral breast cancer.

[1] . lollipop mutation diagram was designed by http://www.bioinformatics.com.cn/
[2] - DCIS

NGS analysis in the patient P3 also detected a VUS c.377C>T, p. (Pro126Leu) in RECQL. The ATM c.875C>T, p. (Pro292Leu), is a rare variant detected in 0.01% of the general population (ExAC[1] source). the mutation is missense and was classified as a single nucleotide variation[2]. The rs747727055 located in coding exon 6 of the ATM gene, results from a C to T substitution at nucleotide position 875 (figure 4-2). This missense variant replaces proline with leucine at codon 292 of the ATM protein. Functional studies have reported instability and low-level expression of the mutant protein, absent or reduced kinase activity, as well as high radiosensitivity after ionizing radiation exposure of cells carrying this variant(125). This alteration has been reported both as the only mutation and in conjunction with a second mutation in ataxia-telangiectasia (A-T) patients(126) .

The patient (P11) carrying the pathogenic *ATM* variant c.3576G>A (p.Lys1192=) had pancreatic cancer at 66 years of age. Twenty-one years earlier, he underwent surgery for a GC. In addition, his brother, his father, and seven of his paternal uncles were deceased from GC (Figure 4-4).

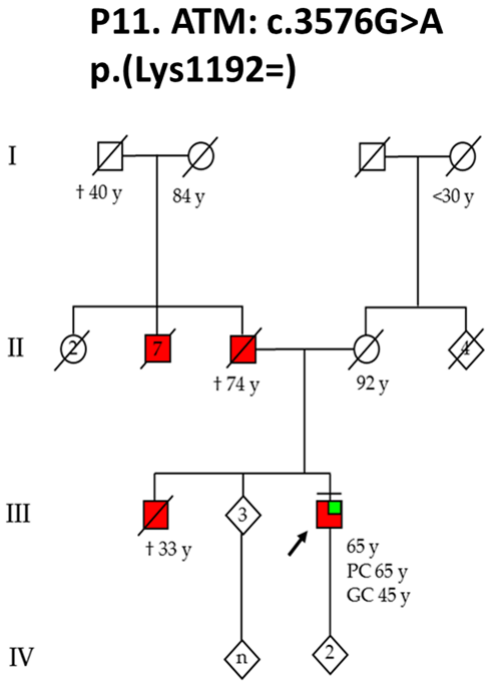

Figure 4-4 : Pedigrees of family with P c.3576G>A variant. Individuals with any cancer are shown as filled circles or square. Arrows show a proband with P variant in ATM family. The tested subject is indicated with a horizontal line above the circle or square. BC, breast cancer; OC, ovarian cancer CC, colon cancer; UC, uterine cancer; NHL, non-Hodgkin lymphoma; PrC, prostatic cancer; Bil BC, bi-lateral breast cancer.

The c.3576G>A substitution is a synonymous variant (SNV)(p.Lys1192=), which,

[1] - Exome Aggregation Consortium
[2] - SNV

located in the last base of exon 26, leads to exon skipping and in-frame deletion (p.Ser1135_Lys1192del) (127, 128). Very rare in the general population, the c.3576G>A has been described in many unrelated Italian AT patients with founder effect(129)(figre4-2).

c.5944C>T (p.Gln1982Ter) variant was detected in a 65-year-old man (P18) with gastric cancer and family history for gastric cancer. 2 of his sisters suffered from the gastric cancer and one of them had ovarian cancer as well.

The c.5944C>T substitution is a synonymous variant (p.Gln1982Ter), which, located in the last base of exon 41 and was classified as SNV.

The pathogenic ATM variant c.5347_5350delGAAA (p.Glu1783ThrfsTer9) was detected in a woman (P30) 81 years old with ovarian cancer. NGS analysis in the patient P30 also identified a VUS c.*591A>AT with unknown protein function in RET. The c.5347_5350delGAAA change creates a premature translational stop signal (p.Glu1783Thrfs*9) in the ATM gene. It is expected to cause an absent or disrupted protein product (figure4-2). The c.5347_5350delGAAA is not present in population databases (ExAC no frequency) and has not been reported in the literature in individuals with ATM-related conditions. Loss-of-function variants in ATM are known to be pathogenic(130).

The ATM variant c.7517_7520delGAGA (p.Arg2506ThrfsTer3) was identified in a woman (P40) 48 years old with Breast cancer as well as a family history of Breast cancer. NGS analysis on the patient P40 identified a VUS c.667A>G, p.Ile223Val in *MUTYH*. This alteration causes a frameshift mutation altering the 3050 amino acid long ATM beginning at position 2506 and leading to a premature termination codon 3 amino acids downstream(figure 4-2). The variant c.7517_7520delGAGA is predicted to result in a truncated or absent protein. Mutation taster predicts the variant to be disease-causing. This alteration was reported in several papers in either compound heterozygosity or homozygosity indicating pathogenicity. The reported patients are predominantly of Italian origin and the variant is considered to be one of the most common Italian AT mutations(129).

4.2.2 *RAD51C*

LP variants in the *RAD51C* gene were detected in a woman with BC, in two women with OC and also a woman with multiple tumors(table4-3) (figure4-5).

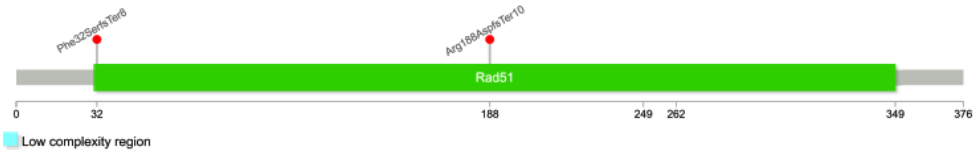

Figure 4-5: Schematic representation of the RAD51C and D Protein and positions of identified P/LP variants[1]

The c.904+5G>T (p.?) splice site variant was found in a patient with OC diagnosed at 40 years (P16). She referred OC also to her two paternal cousins, while her father and her mother had non-Hodgkin lymphoma and thyroid cancer, respectively (Figure 4-6). NGS analysis on the patient P16 identified a VUS c.663A>C, p.Glu221Asp in MSH6. The c.904+5G>T, described in BC and OC patients and families and in 0.002% of the general population (ExAC source), involves a consensus splice site of intron 6(131).

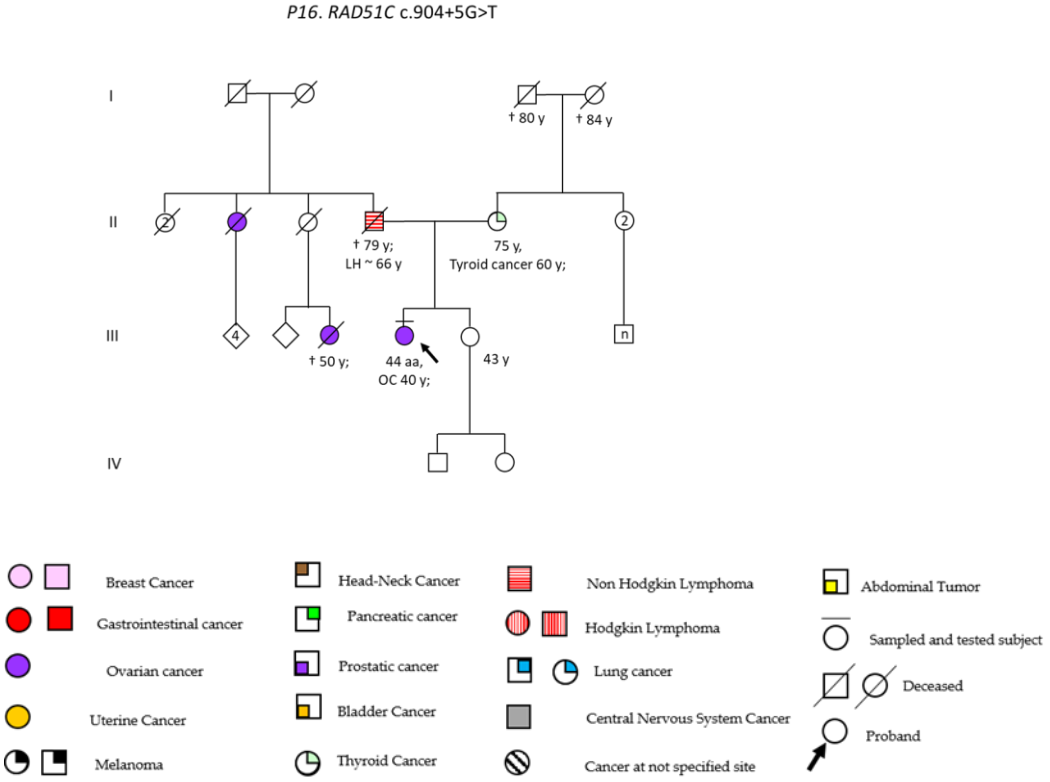

Figure 4-6 Pedigrees of P16

The c.1026+5_1026+7del (P.?) deletion was detected in a high serious grade[2] OC case (P13) with onset at 51 years as well as in a woman 64 years old with multiple

[1] - lollipop mutation diagram was designed by http://www.bioinformatics.com.cn/
[2] - HSG

cancer(P36). The patient P13 referred to a marked maternal familiarity with cancer, including ovarian cancer, Prostate Cancer, esophageal, and gastric cancers in first and second-degree relatives and leukemia and an early-onset brain tumor in cousins (Figure 4-7). This rare variant (ƒ = 1.19 × 10−6, gnomAD data) has been reported in patients with ovarian, breast and uterine cancer. Functional studies have demonstrated that c.1026+5_1026+7del affects a consensus splice site in intron 8 of the RAD51C, leading to the exon 8 skipping and resulting in a frameshift mutation(132).

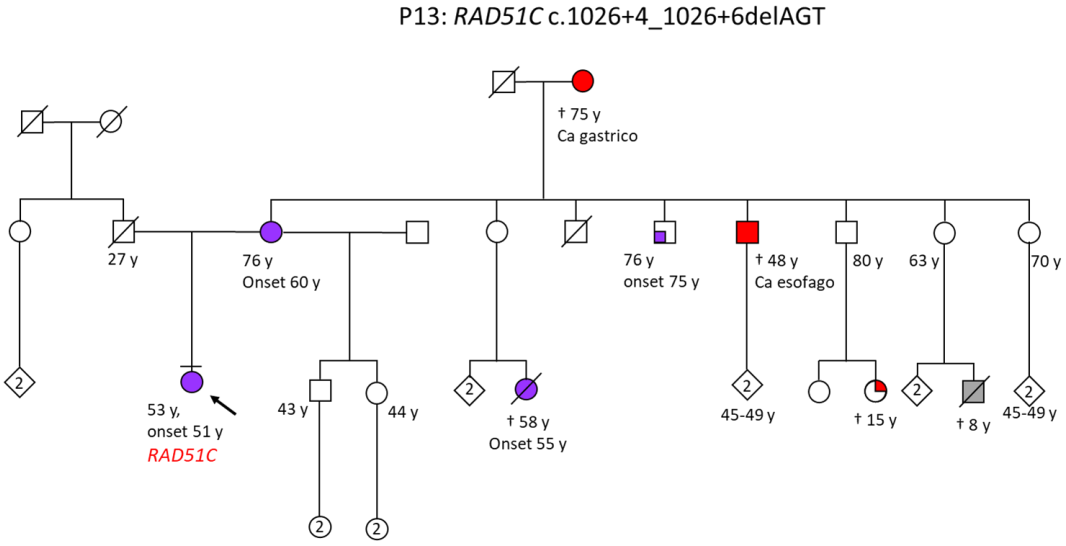

Figure 4-7: Pedigrees of P13. Individuals with any cancer are shown as filled circles or square. Arrows show a proband with P variant in family. The tested subject is indicated with a horizontal line above the circle or square. BC, breast cancer; OC, ovarian cancer CC, colon cancer; UC, uterine cancer; NHL, non-Hodgkin lymphoma; PrC, prostatic cancer; Bil BC, bi-lateral breast cancer.

The RAD51C variant c.93delG (p.Phe32SerfsTer8) was identified in a woman (P23) 48 years old with Breast cancer. NGS analysis on patient P23 detected a VUS c.1437C>G, p.His479Gln in PMS2. The RAD51C c.93delG (p.Phe32SerfsX8) variant causes a premature termination codon, predicted to cause a truncated or absent RAD51C protein due to nonsense-mediated decay, which are commonly known mechanisms for disease(Figure4-5). This variant has also been reported in multiple families and individuals with breast and/or ovarian cancer and based on haplotype analysis authors concluded that it is likely a founder mutation in the Finnish population that is associated with moderate-to-high risk susceptibility for ovarian cancer (gnomAD 5/22298 Finnish European)(133).

4.2.3 RAD51D

RAD51D variant c.562delC (p.Arg188AspfsTer10) was detected in a woman (p36) 41 years old with OC. This sequence change deletes 1 nucleotide from exon 9 of the RAD51D mRNA (c.562delc), causing a frameshift at codon 300. This creates a premature translational stop signal in the last exon of the RAD51D mRNA (p.Arg300Aspfs*10). While this is not anticipated to result in nonsense-mediated decay, it is expected to disrupt the last 29 amino acids of the RAD51D protein. This variant is expected to disrupt amino acid residues Arg300-Thr328 of the RAD51D protein, thereby removing the C-terminus of the ATPase domain(figure4-5). This variant is a rare truncation that likely disrupts an important functional domain in the RAD51D protein(134).

4.2.4 MLH1

Three of the patients (P20, P28 and P38) carry 3 pathogenic alternations (c.1852_1854delAAG, c.1897G>A, c.210_213delAGAA respectively) on the MLH1 gene (figure4-8).

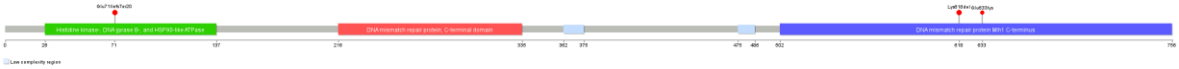

Figure 4-8: Schematic representation of the MLH1 Protein and positions of identified P/LP variants[1]

P20 was a man 58 years old with Lynch syndrome that carrier c.1852_1854delAAG (p.Lys618del). The MLH1 c.1852_1854delAAG; p.Lys618del variant (rs63751247), known as Lys616del, is a common variation in individuals diagnosed with Lynch syndrome and segregates with the disorder. Functional characterization of the variant protein represents reduced mismatch repair activity because of decreased steady-state levels of MLH1 protein and reduced localization with PMS2 and EXO1(figure4-8). Based on available information, this variant is considered pathogenic(135).

P28 was a woman 54 with BC and a family history of breast cancer. She carried an LP variant c.1897G>A (p.Glu633Lys) and c.*591A>A (RET) VUS variant. This pathogenic variant is denoted MLH1 c.1897G>T at the cDNA level and p.Glu633Ter at the protein level. The substitution forms a nonsense variant, which changes a Glutamic Acid to a premature stop codon (GAA>TAA)(figure4-8), and is predicted to

[1] - lollipop mutation diagram was designed by http://www.bioinformatics.com.cn/

cause loss of normal protein function through either protein truncation or nonsense-mediated mRNA decay(135).

P38 was a patient 53 years old with Lynch syndrome. Her father and grandfather (paternal family) had Colon cancer, aunt and uncle (paternal family) had brain and pancreatic cancer respectively. She carried a pathogenic frameshift deletion mutation c.210_213delAGAA. This variant causes p.Glu71IlefsTer20. Also P38 had VUS variant c.2770C>T, p.Arg924Trp on ATM. The c.210_213delAGAA pathogenic mutation, located in coding exon 3 of the MLH1 gene, results from a deletion of 4 nucleotides at nucleotide positions 210 to 213, causing a translational frameshift with a predicted alternate stop codon (p.E71Ifs*20)(figure4-8). This variant has been seen in many families that either met Amsterdam criteria or Bethesda guidelines for testing hereditary nonpolyposis colorectal cancer (HNPCC)/Lynch syndrome(136).

4.2.5 MSH2

Three of the probands (P8, P9 and P31) carry 2 pathogenic single nucleotide variations c.1216C>T and c.425C>G respectively on the *MSH2* gene (figure 4-9).

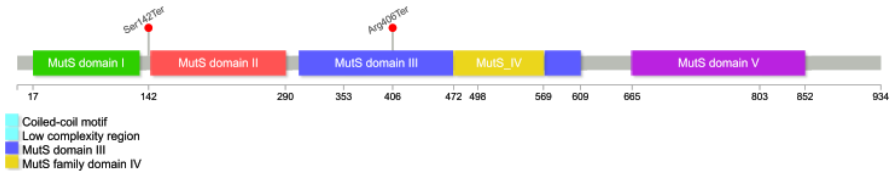

Figure 4-9: Schematic representation of the MSH2 Protein and positions of identified P/LP variants[1]

P8 and P9 both have c.1216C>T(p.Arg406Ter), but P9 has VUS mutation c.845G>A (p.Arg282Gln) on the *TP53* gene. P8 was a woman 44 years old with Lynch syndrome (Colon Cancer), and P9 was mother of P8, she has 69 years old with LS (UC, CC). P9 has two sisters, one suffering from pancreatic cancer and the other suffering from colon and breast cancer. Mother was suffering from uterine and pancreatic cancer and mother's sister was suffering from kidney cancer (Figure 4-10).

The c.1216C>T pathogenic mutation, located in coding exon 7 of the MSH2 gene, results from a C to T change at nucleotide position 1216. This varies the amino acid from an arginine to a stop codon within coding exon 7(figure4-9). This alter-

[1] - lollipop mutation diagram was designed by http://www.bioinformatics.com.cn/

ation causes loss of function by premature protein truncation or nonsense-mediated mRNA decay and also has been detected in multiple families meeting Amsterdam criteria(137).

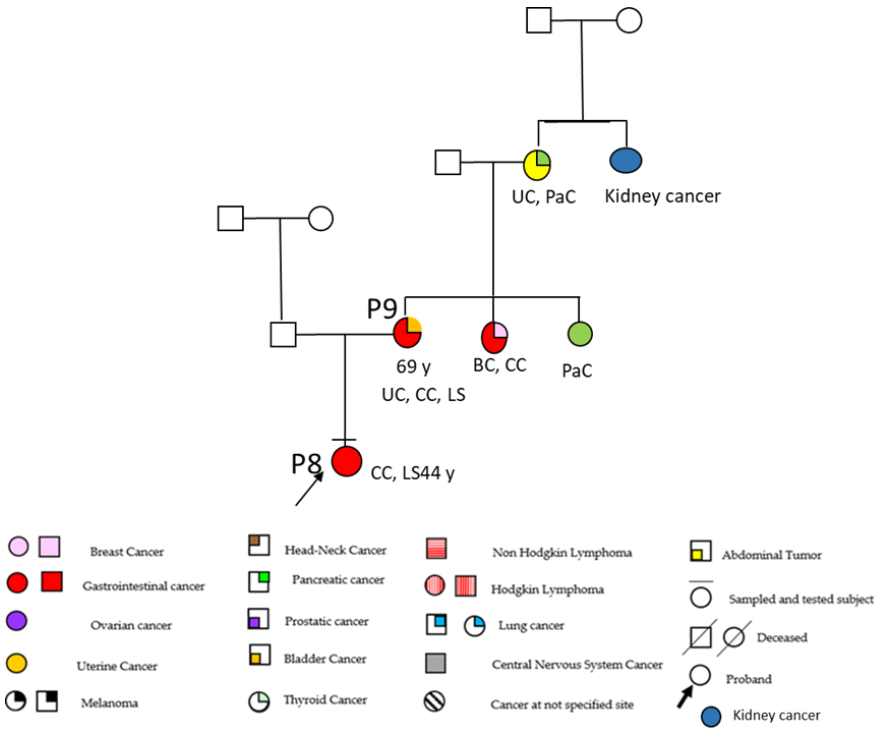

Figure 4-10: Pedigrees of family with P c.1216C>T(p.Arg406Ter) variant. Individuals with any cancer are shown as filled circles or square. Arrows show a proband with P variant in family. The tested subject is indicated with a horizontal line above the circle or square.

Proband P31 was a man with gastric cancer and family history for Gastric cancer who carried c.425C>G (P.Ser142Ter) on *MSH2*. This mutation has been reported in related to Lynch syndrome. The change creates a nonsense variant, which changes a Serine to a premature stop codon (TCA>TGA), and is expected to result in loss of normal protein function via either protein truncation or nonsense-mediated mRNA decay(figure4-9)(138).

4.2.6 MSH6

Probands P6, P7, P22 and P37 with lynch syndrome were carrier LP/P variants on *MSH6* gene.

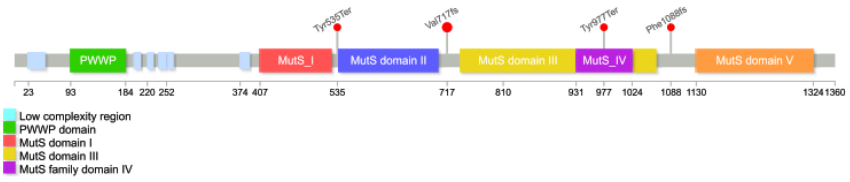

Figure 4-11: Schematic representation of the MSH6 Protein and positions of identified P/LP variants[1]

P6 was a woman 79 years old with Lynch syndromed with family history that carried c.3261_3262insC (p.Phe1088fs) P mutation. Her sister has CC and TC and son

of her sister suffers from Colon Cancer (figure 4-12).

The c.3261_3262insC variant is located in exon 5 *MSH6* gene that likely causes a frameshift and, result in a premature stop codon. Predicted it causes a truncated or absent MSH6 protein because of nonsense mediated decay, which are commonly known mechanisms for disease(figure4-10). The variant was seen in a large control population, ExAC, with an allele frequency of 211/120336 (1/570), which is approximately 12 times the estimated maximal expected allele frequency for a pathogenic MSH6 variant of 1/7031. The c.3261dupC variant in the *MSH6* gene is has been reported in multiple individuals with Lynch Syndrome, colon cancer , endometrial cancer(139).

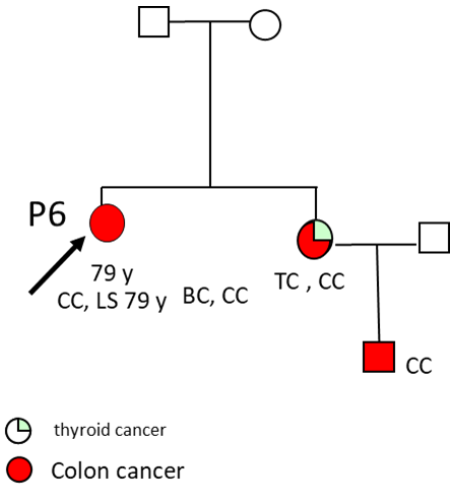

Figure 4-12: Pedigrees of P6. Individuals with any cancer are shown as filled circles or square. Arrows show a proband with P variant in family. The tested subject is indicated with a horizontal line above the circle or square.

Proband P7 was a man75 years-old with Lynch syndromed and Prostatic Cancer carrying c.2150_2153delTCAG (p.Val717fs) Pathogenic mutation. His family has different kind of the cancer: his father with Kidney cancer, his sister with UC, aunt had AML and one cousin had kidney cancer (figure 4-13).

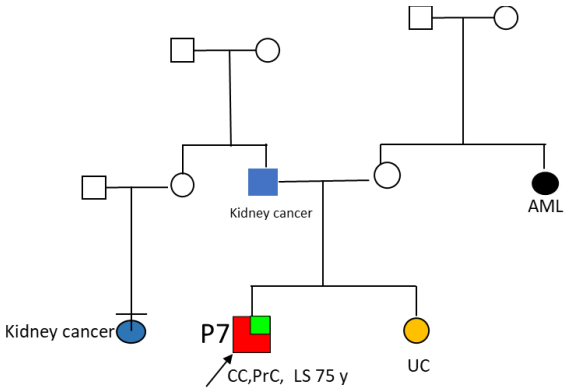

Figure 4-13. Pedigrees of P7..

The MSH6 c.2150_2153delTCAG (p.Val717Alafs) pathogenic mutation, located in

coding exon 4 of the MSH6 gene, results from a deletion of 4 nucleotides from position 2150 to 2153, creating a translational frameshift with premature termination codon (figure4-11). This alteration is expected to result in a truncated or absent MSH6 protein due to nonsense mediated mRNA decay. Therefore, this alteration is interpreted as a disease-causing mutation. This mutation has been previously reported in numerous patients diagnosed with CC, ovarian cancer and endometrial cancer(140).

c.1605C>A (p.Tyr535Ter) Pathogenic mutation on *MSH6* was observed on woman (P22) 82 years-old with UC and CC. She was diagnosed for lynch syndrome. Her mother and one of the her brother had CC, one of the brother had Prostatic cancer and other had lung cancer (figure4-14).

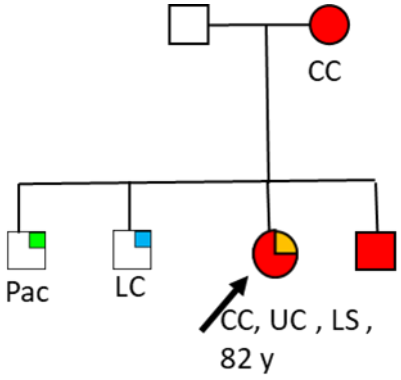

Figure 4-14: Pedigrees of P22 .

This sequence change results in a premature translational termination signal (p.Tyr535*) in the MSH6 gene. It probably causes an absent or disrupted MSH6 protein product(figure4-11) (90).

The patient P37 was woman 70 years old with Lynch syndrome, CC, UC and melanoma carrying c.2931C>A (p.Tyr977Ter) Pathogenic variant. She has one aunt with balder cancer, other aunt with breast cancer, uncle with kidney and balder cancer and his son with kidney cancer (figure 4-15).

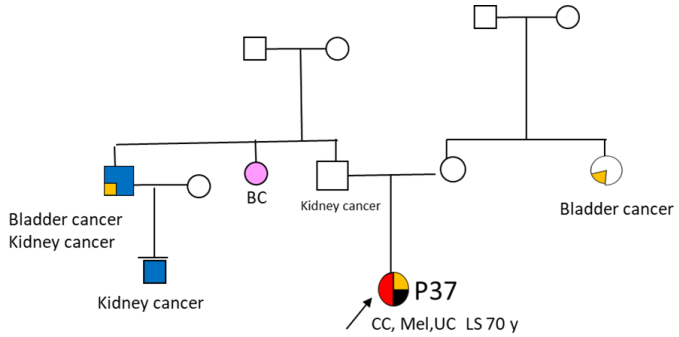

Figure 4-15: Pedigrees of P37.

The c.2931C>A pathogenic mutation, located in coding exon 4 of the MSH6 gene, results from a C to A substitution at nucleotide position 2931(figure4-10). This mutation creates a premature translational stop signal in the MSH6 gene. In such a way that the amino acid from a tyrosine to a stop codon within coding exon 4 (p.Tyr977Ter). This alteration was reported as a Swedish founder mutation. It is reported this alteration results in an absent or non-functional protein product truncation or nonsense-mediated mRNA decay. This variant may not affect RNA splicing. Loss-of-function variants in MSH6 are known to be pathogenic(141).

4.2.7 MUTYH

Four LP/P mutations were found in *MUTYH* gene in the 5 participants.

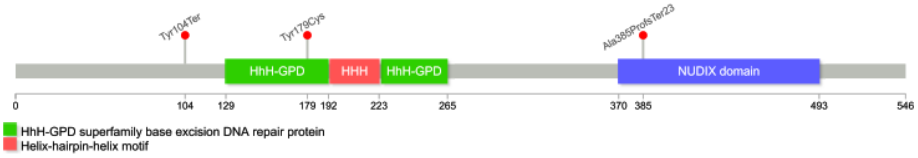

Figure 4-16: Schematic representation of the MUTYH Protein and positions of identified P/LP variants[1]

c.312C>A(p.Tyr104Ter) has been detected in a woman 53 years old with PaC (p19). The MUTYH c.312C>A variant causes a premature stop codon, predicted to result in a truncated or absent MUTYH protein because of the nonsense mediated decay (figure4-16). The variant has been reported in multiple patients in the articles in the homozygous and compound heterozygous state, and has been shown to be completely devoid of both glycosylase and DNA binding activities(142).

proband P21 carrying c.1147delC pathogenic variant (p.Ala385ProfsTer23) in MUTYH gene was a woman 60 years old with BC, TC. The c.1147delC located in coding exon 12 of the MUTYH gene, that results from a deletion of one nucleotide at nucleotide position 1147(figure4-16). This pathogenic sequence change causes an amino acid frameshift and creates a premature stop codon 23 amino acids downstream of the change (p.Ala385Profs*23). It is possible to cause a truncated or absent MUTYH protein because of aforementioned reasons. This alteration has been reported in a homozygous and compound heterozygous state in many suffers with polyposis and has been described as a European founder mutation (143).

[1] - lollipop mutation diagram was designed by http://www.bioinformatics.com.cn/

The functional researches demonstrate that the c.1147delC is completely defective in glycosylase and DNA binding activities and results in a protein free of glycosylase and DNA binding activity, as well as a 50-100% decrease in MUTYH protein expression levels compared to wild-type levels in a compound heterozygous(142).

c.933+3A>C Pathogenic mutation was observed in-patient 54-year-old (P29) with breast cancer and positive family history for breast cancer. She also was carrier for c.144+508G>T VUS mutation in RAD51D with unknown protein changing (144). *MUTYH* c.933+3A>C mutation affects a non-conserved intronic nucleotide that results in a fusion of exon 9 to exon 11. The variant has been reported in numerous colorectal cancer patients in the literature. It has been identified in 7/66554 of European chromosomes by the Exome Aggregation Consortium (ExAC, http://exac.broadinstitute.org; dbSNP rs587780751).

c.144+508G>T VUS mutation in RAD51D is a missense variant that replaces arginine with methionine at codon 54 of the RAD51D protein. The available evidence is not sufficient to determine the role of this variant in disease conclusively. Computational prediction shows that this mutation may not impact protein structure and function (internally defined REVEL score threshold <= 0.5). This variant has not been reported in suffers with hereditary cancer in the papers as well as has been identified in 2/251236 chromosomes in the general population by the Genome Aggregation Database[1]. The c.144+508G>T is classified as a Variant of Uncertain Significance (145).

The c.536A>G with protein changing p.Tyr179Cys was identified in 2 patients P25 and P31 with Polipioclon and gastric cancer respectively.

The p.Tyr179Cys variant causes extremely decreased DNA binding and adenine DNA glycosylase activity(figure4-16). The evidence for the origin of the common European mutations cited that two variants p.Tyr179Cys and p.Gly396Asp in *MUTYH* is pathogenic variants for MUTYH-associated polyposis[2] disease as well as is a known founder mutation in the European population and has been described in the gnomAD database (frequency of 0.25%) in the non-Finnish European population. That *MUTYH* c.536A>G has been identified in homozygous or

[1] - gnomAD
[2] - MAP

compound heterozygous form with other pathogenic *MUTYH* variants in affected individuals with colorectal cancer, familial adenomatous polyposis[1] or attenuated FAP. DNA sequence analysis of the *MUTYH* gene showed a sequence change, c.536A>G, in exon 7 that causes to change a highly conserved amino acid residue located in a domain of the MUTYH protein, p.Tyr179Cys(146).

The patient P25 has 69 years old suffering from Colon polyps. She carried c.6795A>G (p.Gln2265=) in *APC* gene as well. The c.6795A>G mutation has been classified as a Variant of Uncertain Significance. It affects codon 2265 of the *APC* mRNA. It does not alter the encoded amino acid sequence of the APC protein due to it is a silent change(figure4-16). The available evidence is not enough to determine the role of that in disease. This variant may create or strengthen a splice site based on the algorithms developed to predict the effect of sequence changes on RNA splicing (Clinvar).

Proband P31 with gastric cancer was a man 50 years old with family history for gastric cancer. He carried 2 pathogenic mutations c.425C>A (p.Ser142Ter) and c.536A>G (p.Tyr179Cys) in *MSH2* and *MUTYH* genes respectively(figure4-11 and 4-9).

4.2.8 PMS2

In the *PMS2* gene was observed a new pathogenic variant c.478C>T (p.Gln160Ter) in a man 66 years old (P33) with Lynch syndrome and lymphoma. Proband P33 carried c.1277A>T (p.Gln426Leu) VUS mutation in *MLH1* gene. Her mother and one aunt had BC, her father two brothers suffered from PaC, three of other brothers and one sister had Lung cancer and other sister had liver cancer (figure 4-17).

[1] - FAP

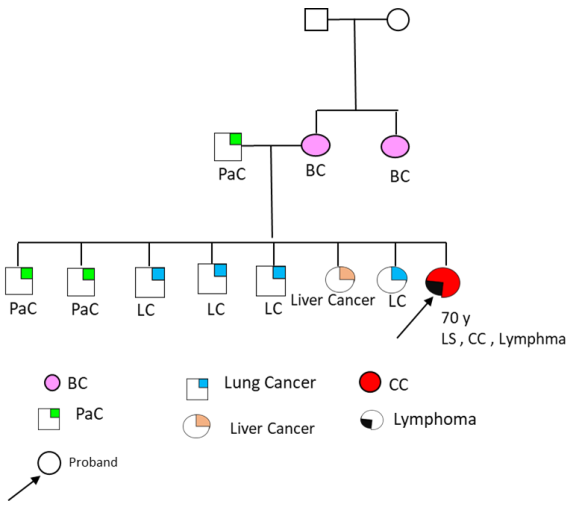

Figure 4-17: Pedigrees of the P33.

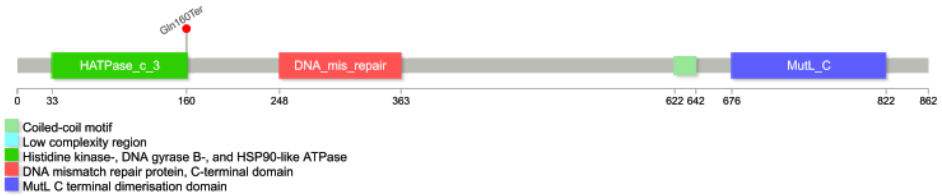

Figure 4-18: Schematic representation of the PMS2 Protein and positions of identified P/LP variant.[1]

The c.478C>T causes to create a premature translational stop signal (p.Gln160*) in the PMS2 gene. It may cause an absent or disrupted protein product. Loss-of-function variants in PMS2 are pathogenic(figure4-18). This variant has not been reported in the literature in individuals with PMS2-related disease(90).

c.1277A>T (p.Gln426Leu) located in coding exon 12 of the MLH1 gene, results from an A to T substitution at nucleotide position 1277. This sequence change results in replacing glutamine with leucine, an amino acid with dissimilar properties, at codon 426 of the MLH1 protein (p.Gln426Leu). The glutamine at codon 426 is weakly conserved, and there is a moderate physicochemical difference between glutamine and leucine, this alteration is predicted to be tolerated by in silico analysis. The *MLH1* c.1277A>T variant, is not reported in the medical literature but is reported in ClinVar (Variation ID: 231598). Because of the lack of clinical and functional data, the significance of this variant is uncertain at this time(147).

4.2.9 RECQL

In investigation in *RECQL* gene was identified a pathogenic mutation (c.362 363 delTA) in a man 72 years-old (P12) with Breast cancer at 59 years and pancreatic

[1] - lollipop mutation diagram was designed by http://www.bioinformatics.com.cn/

cancer at 70 with familiy history cancer. He had a sister deceased at 56 years for pulmonary cancer and BC at 38 years (Figure 4-20). This small deletion in RECQL was also excluded in the healthy 49 years old daughter of the proband.

Figure 4-19: Schematic representation of the RECQL Protein and positions of identified P/LP variant.[1]

In multiple independent studies were cited that damaging mutations in RECQL have been associated to increased breast cancer risk and to genomic instability as well as the germline mutations of RECQL were reported in affected individuals with hereditary Breast Cancer(figure4-19). The c.362 363 delTA is predicted to cause a premature stop codon with consequent production of a truncated protein and loss of the helicase activity. The Cys122Leufs*43 is located in the helicase Rec-A like domain A1 (amino acid resi-dues 63–281) of RECQL protein, containing the highly conserved signature helicase motifs of the SF-2 superfamily(61).

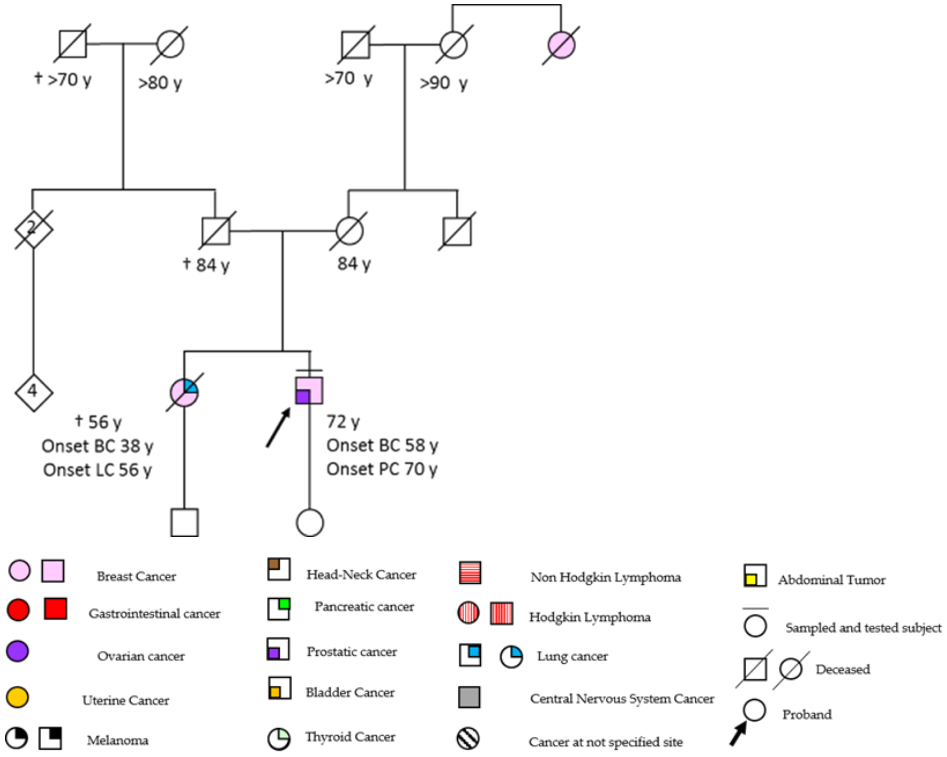

Figure 4-20. Pedigrees P12 with P/LP variants. Individuals with any cancer are shown. The tested subject is indicated with a horizontal line above the circle or square. BC, breast cancer; OC, ovarian cancer; GC, gastric cancer; PC, pancreatic cancer; CC, colon cancer; RC, colorectal cancer; PrC, prostatic cancer; LC, lung cancer;

4.2.10 TP53

A pathogenic mutation c.158G>A with protein changing p.Trp53Ter was found in *TP53* gene. Proband P27 carrying it was a 38 years-old female with breast cancer.

Figure 4-21: Schematic representation of the TP53 Protein and positions of identified P/LP variant[1]

The c.158G>A pathogenic mutation located in coding exon 3 of the TP53 gene, results from a G to A substitution at nucleotide position 158. This sequence change creates a premature translational stop signal (p.Trp53*) in the TP53 gene(figure4-21). The premature stop codons are typically deleterious in nature. Therefore, It causes an absent or disrupted protein product. However, Loss-of-function variants in TP53 are known to be pathogenic, this variant has not been reported in the literature in individuals with TP53-related disease(148).

4.2.11 PALB2

During study in *PALB2* gene, a likely pathogenic variant c.3350G>A (p. Arg1117Lys) was identified in a case 50 years-old-female (p24) with breast cancer.

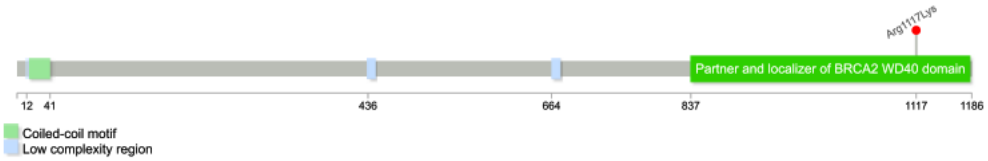

Figure 4-22: Schematic representation of the PALB2 Protein and positions of identified P/LP variant[2]

PALB2 c.3350G>A causes a conservative amino acid change located in the Partner and localiser of BRCA2, WD40 domain of the encoded protein sequence. c.3350G>A has been reported in the papers in sufferers with breast cancer. This sequence change (p. Arg1117Lys) replaces arginine, which is basic and polar, with lysine, which is basic and polar, at codon 1117 of the PALB2 protein as well as this mutation falls at the last nucleotide of exon 12, which is part of the consensus splice site for this exon(figure4-22)(149).

4.2.12 NBN

Investigation of the patient p17 (30 years old with breast cancer) was observed a

[1] - lollipop mutation diagram was designed by http://www.bioinformatics.com.cn/
[2] - lollipop mutation diagram was designed by http://www.bioinformatics.com.cn/

novel likely pathogenic nonsense variant c.1896G>A in *NBN* gene. She also has a VUS c.244T>C in *MRE11* gene.

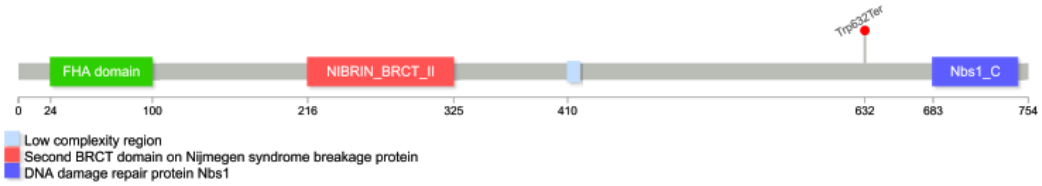

Figure 4-23: Schematic representation of the NBN Protein and positions of identified P/LP variant[1]

The variant c.1896G> A identified in the *NBN* gene has an early termination site of protein synthesis, with prediction of truncated protein (p. Trp632Ter). In Varsome, on the basis of the ACMG criteria it is classified as probably pathogenetic. It is not reported in the literature and is absent in the CLinVar , dbSNP and gnomAD databases(figure4-23).

MRE11 c.244T>C (p.Tyr82His) replaces tyrosine with histidine at codon 82 of the MRE11A protein. The tyrosine residue is extremely conserved and there is a moderate physicochemical difference between tyrosine and histidine. This is a rare missense variant with uncertain impact on protein function. Based on the now knowledge, there is no indication that this mutation causes disease as well as the evidence is not sufficient to prove that conclusively. Therefore, it has been classified as a Variant of Uncertain Significance (Clinvar).

4.2.13 CDKN2A

In this study, pathogenic variant c.193+1G>A was found in *CDKN2A* gene in affected individual (42 years-old-male) (p39) with multiple tumors (colon cancer, thyroid cancer, lung cancer and brain cancer).

This variant c.193+1G>A has been identified in several individuals with a history of malignant melanoma and shown that it has also been observed to segregate with disease in related individuals.

This variant consists of a G>A nucleotide substitution at the +1 position of intron 1B of the of the CDKN2A (p14ARF) gene. This variant destroys a canonical splice donor site and is expected to result in abnormal gene splicing, leading to either an abnormal message that causes disrupt RNA splicing and likely results in an absent or disrupted protein product. The CDKN2A gene encodes the p16 protein and, using an alternate reading frame, the p14-ARF protein as well. As exon 1B is only used by the p14-ARF protein, CDKN2A c.193+1G>A will not affect the p16 protein(150).

4.2.14 BARD1

c.1205C>A (p.Ser402Ter) pathogenic mutation was identified in *BARD1* in patient P34 who was a woman 52 years old with breast cancer and a family history of breast cancer.

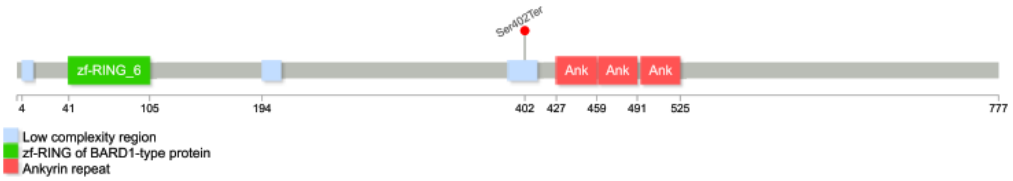

Figure 4-24: Schematic representation of the BARD1 Protein and positions of identified P/LP variant[1]

These variant changes 1 nucleotide (C to A substitution) in exon 4 of the *BARD1* gene at nucleotide position 1205. This changes the amino acid from a serine to a stop codon within coding exon 4, creating a premature translation stop signal. This alteration is expected to result in an absent or non-functional protein product or nonsense-mediated mRNA decay(figure4-24). Loss-of-function variants in *BARD1* are known to be pathogenic. This variant has not been reported in individuals affected with BARD1-related conditions and hereditary cancer in the literature. This variant is not present in population databases (ExAC no frequency)(151, 152).

4.2.15 CDH1

A Pathogenic variant, c.1003C>T, related to the *CDH1* gene was found in proband p41.

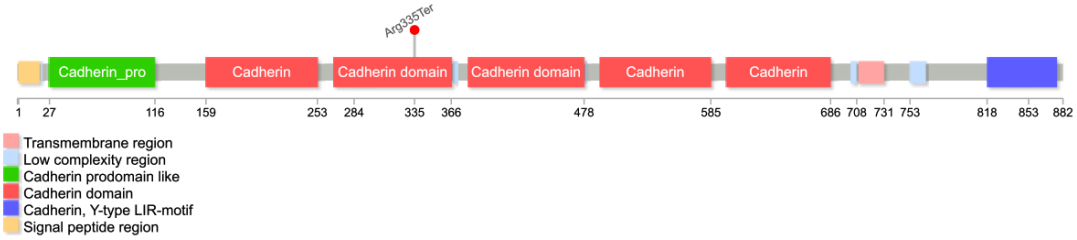

Figure 4-25: Schematic representation of the CDH1 Protein and positions of identified P/LP variant[2]

c.1003C>T (p.Arg335*) creates a nonsense variant, which changes an Arginine to a premature stop codon (CGA>TGA), and is predicted to lead to loss of normal protein function through either protein truncation (PVS1) or nonsense-mediated mRNA decay. This pathogenic mutation is denoted CDH1 c.1003C>T at the cDNA

[1] - lollipop mutation diagram was designed by http://www.bioinformatics.com.cn/
[2] - lollipop mutation diagram was designed by http://www.bioinformatics.com.cn/

level and p.Arg335Ter (R335X) at the protein level as well(figure4-25). This variant has been reported in many hereditary diffuse gastric cancer families as well as was identified co-segregate with the disease in multiple affected family members, with >7 meiosis observed across at least two families. This variant meets criteria to be classified as pathogenic based on the ACMG/AMP criteria applied as specified by the CDH1 Variant Curation Expert Panel: PVS1, PP1_Strong(153).

4.2.16 APC

In this study was observed P variant in *APC* c.6709C>T in a woman 55 years old with cc and LS (P26). She had one uncle and two sisters with CC also one of her sisters had TC.

The c.6709C>T pathogenic mutation located in coding exon 15 of the *APC* gene, results from a C to T substitution at nucleotide position 6709. This sequence change makes a premature translational stop signal (p.Arg2237*) in the APC gene. This variant the amino acid from an arginine to a stop codon within coding exon 15 and it is expected to disrupt the last 607 amino acid(s) of the APC protein. This mutation is expected to disrupt the EB1 and HDLG binding sites, that mediate interactions with the cytoskeleton. This disruption of the C-terminal portion of the protein is functionally important(figure4-26).

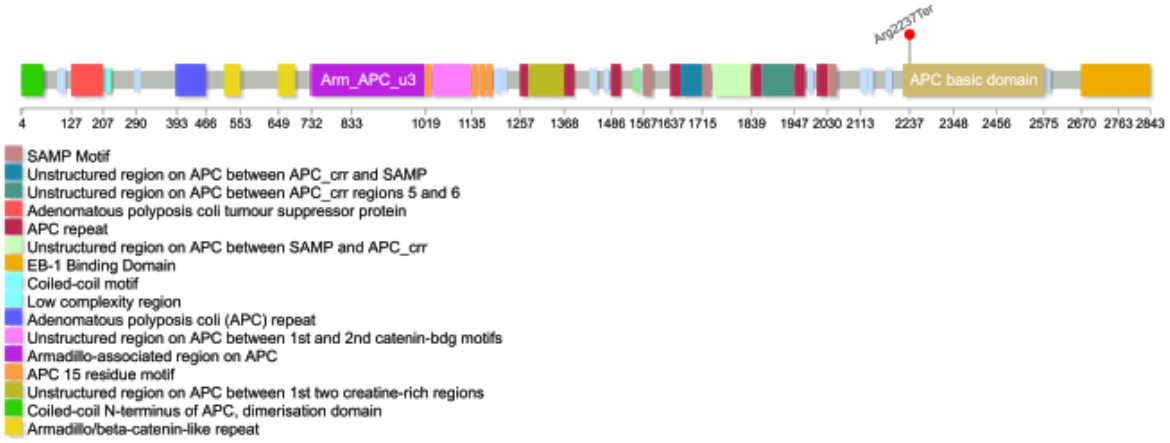

Figure 4-26: Schematic representation of the APC Protein and positions of identified P/LP variant[1]

The deletion of this region of the APC protein is causative of disease. This premature translational stop signal has been reported in patients with attenuated familial adenomatous polyposis(154, 155).

4.2.17 CHEK2

Four LP/P variants in *CHEK2* gene (c.1169A>C, c.470T>C, c.793-1G>A and c.1100delC) were found in this research.

proband P1 carrying c.1169A>C (p.Tyr390Ser) was woman 51 years old with Mullerian's sarcoma and family history of colon cancer and breast cancer (figure4-27)

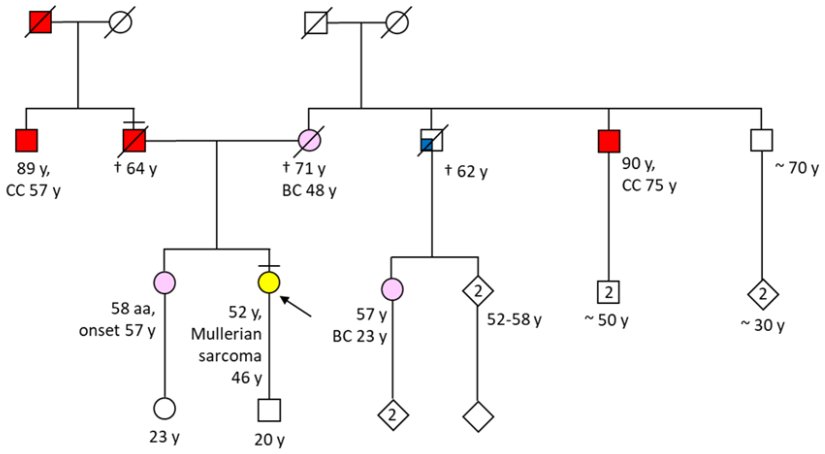

Figure 4-27.. Pedigrees P1 with P/LP variants. Her mother, her sister, and a maternal cousin indeed had BC at 47, 57, and 23 years, respectively. Two maternal uncles had also malignancies—one pulmonary and the other colorectal. However, colon cancer familiarity was stronger among paternal relatives, occurring in her father, uncle, and grandfather

Figure 4-28: Schematic representation of the CHEK2 Protein and positions of identified P/LP variant[1]

The c.1169A>C variant located in coding exon 10 of the CHEK2 gene, results from an A to C substitution at nucleotide position 1169. CHEK2 c.1169A>C (p.Tyr390Ser) causes the tyrosine at codon 390 to be replaced by serine located in the Protein kinase domain of the encoded CHEK2 protein sequence(figure4-28). This alteration has been reported in many patients undergoing multi-gene panel testing, including individuals diagnosed with breast and/or ovarian cancer. According to an in vitro kinase activity test combined with in silico models, this change had been related to no detectable kinase activity and was interpreted as likely deleterious(156).

The c.470T>C variant was discovered in two unrelated patients with unilateral hormone-responsive DC—invasive and moderately differentiated in one case (P2) (figure 4-29) and in situ and well-differentiated in the remaining (P4) (Figure 4-30).

Proband P2 also had other primitive malignancies, such as breast cancer at 42 years, Colon Cancer at 45 years, ovarian cancer at 46 years and Gastric Cancer at the age of 47. In addition, her father had multiple cancers: after a rectal carcinoma,

[1] - lollipop mutation diagram was designed by http://www.bioinformatics.com.cn/

a Prostate Cancer occurred more recently. Also, one of the cousins and grandfather had brain cancer.

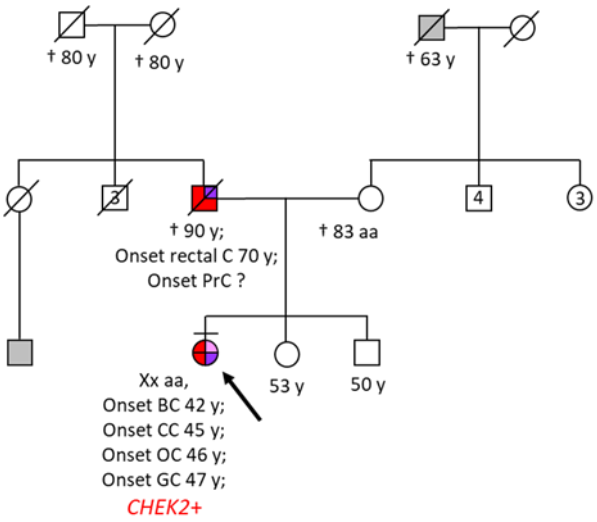

Figure 4-29. Pedigrees P2 with P/LP variants

Similarly, P4 had a familiarity with rectal carcinoma (maternal uncle) but referred to ovarian cancer in her younger sister and uterine adenocarcinoma in her mother as well (Figure 4-30). She also had the c.145G>A, p.(Ala49Thr) VUS in the PMS2 gene.

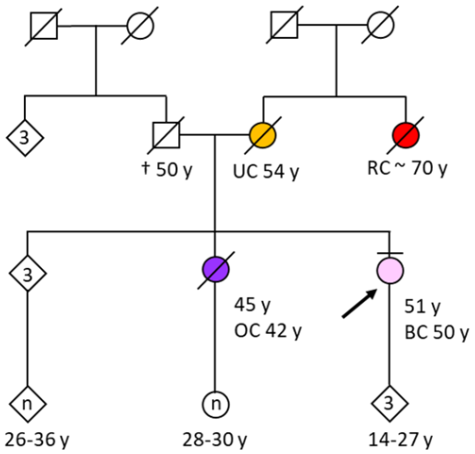

Figure 4-30. Pedigrees P4 with P/LP variants

The c.470T>C variant located in exon 4 results in an amino acid change, p.Ile157Thr. The p.Ile157Thr alteration affects a poorly conserved amino acid residue located in the Fork-head-associated[1] domain of the CHEK2 protein(figure4-28). The mutant allele does not affect the kinase activity of the protein but does affect the dimerization of the protein in a dominant negative manner, Therefore,

[1] - FHA

it is expected the p.Ile157Thr fails to bind to checkpoint proteins and inability to autophosphorylation. This c.470T>C (p.Ile157Thr) variant in the *CHEK2* gene is associated with cancer susceptibility, for example, significant association with increased risk of breast and colon cancer in large meta-analysis studies [odds ratios ~1.5], prostate cancer, CLL, colorectal cancer and other cancers and has been considered to be a susceptibility risk allele to cancer(157).

The c.793-1G>A splicing variant was observed in a 47-year-old woman (P5) with papillary thyroid cancer at the age of 27 and first-degree familiarity for BC. Her family had different malignancies, including osteosarcoma in her nephew, metachronous bilateral breast cancer in her deceased mother, Prostatic cancer and gastric cancer in the maternal uncle and grandfather, respectively. Prostatic cancer was observed in three paternal uncles, while a fourth paternal uncle deceased of lung cancer.

Segregation studies were done on the brother but not on his young son, who denied his consent, despite his previous diagnosis of osteosarcoma and the detection of the *CHEK2* LP variant in his father (Figure 4-31).

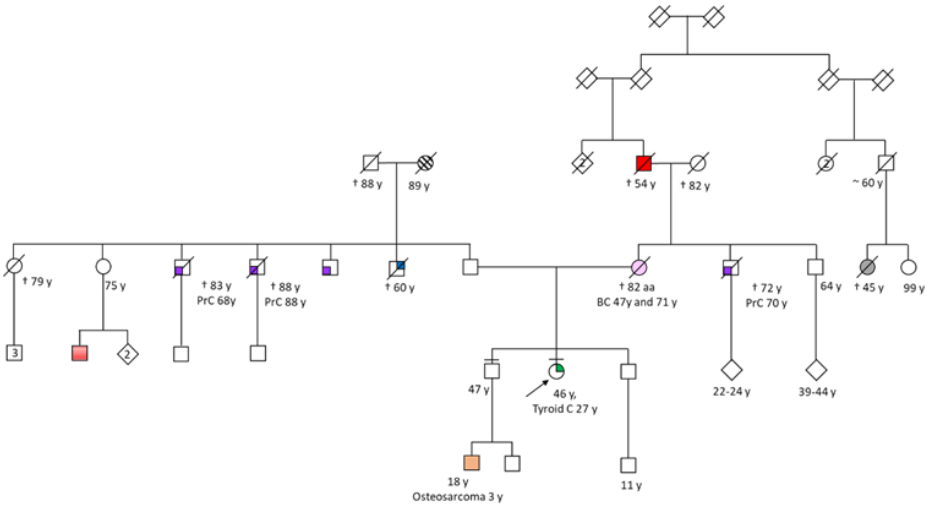

Figure 4-31. Pedigrees P5 with P/LP variants

The c.793-1G>A intronic variant results from a G to A substitution of one nucleotide upstream from coding exon 6 of the *CHEK2* gene. It results in an alternative splicing site downstream and a frameshift mutation, that leads to a premature stop codon and a consequent truncated protein p.Asp265Thrfs*10(figure4-28)(158). This variant has been reported in patients with breast cancer, prostatic cancer, and in affected individuals who underwent genetic testing for the high risk of hereditary cancer (159).

The c.1100del variant was identified in two women (P14 and P15) with bilateral metachronous hormone-responsive DC (56 and 66 years old respectively). Both sufferers had a post-menopausal onset and positive family history of breast cancer in first and/or second-degree maternal relatives. Proband P15 had diffuse mesenteric leiomyomatosis as well (Figure 4-32 and 4-33).

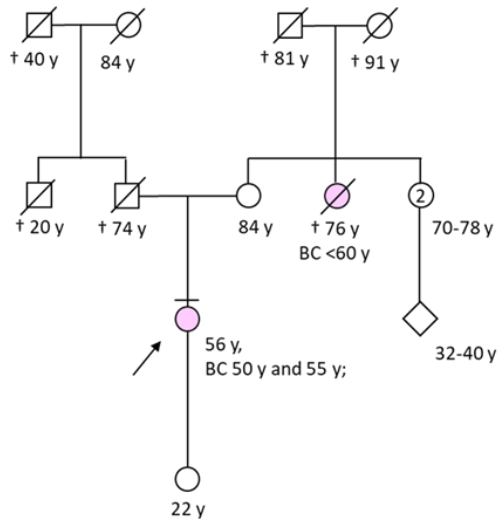

Figure 4-32. Pedigrees P14 with P/LP variants

The c.1100del variant results in a frameshift and premature stop codon, p.Thr367Metfs*15, responsible for the loss of the response to DNA damage and for the impairment of the CHEK2 kinase activity(figure4-28)(160). This mutation causes CHEK2 kinase protein truncation or nonsense-mediated decay in a gene for which loss-of-function is a known mechanism of disease. This variant has been reported in Li-Fraumeni Syndrome as well as in several cancer types as breast, ovarian, prostatic, colon, renal and thyroid malignancies (161).

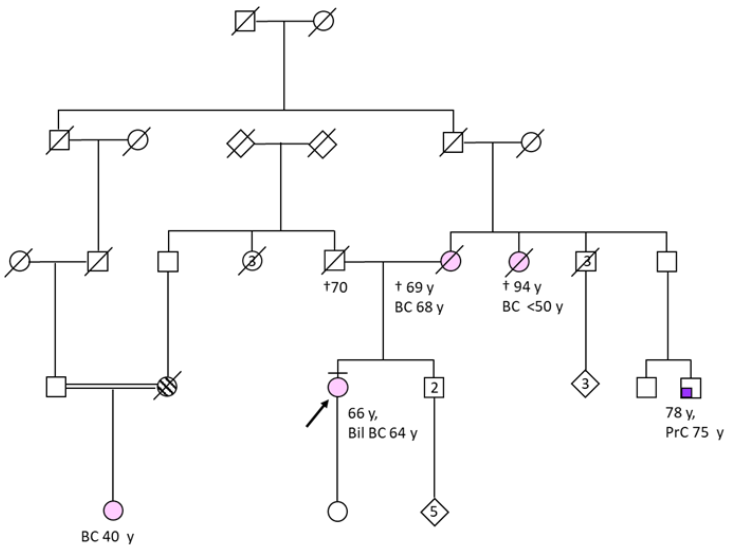

Figure 4-33. Pedigrees P15 with P/LP variants

4.3 MLPA result

38 patients were recommended for MLPA Panel. Two patients were positive; P220 for MHL1 and 230 were positive for check2.

Patient 220 showed a large deletion in exon 1-9 of the MHL1 gene. She was 51 years old with endometrial cancer. This mutation was detected in her two healthy twin sisters (figure 4-34).

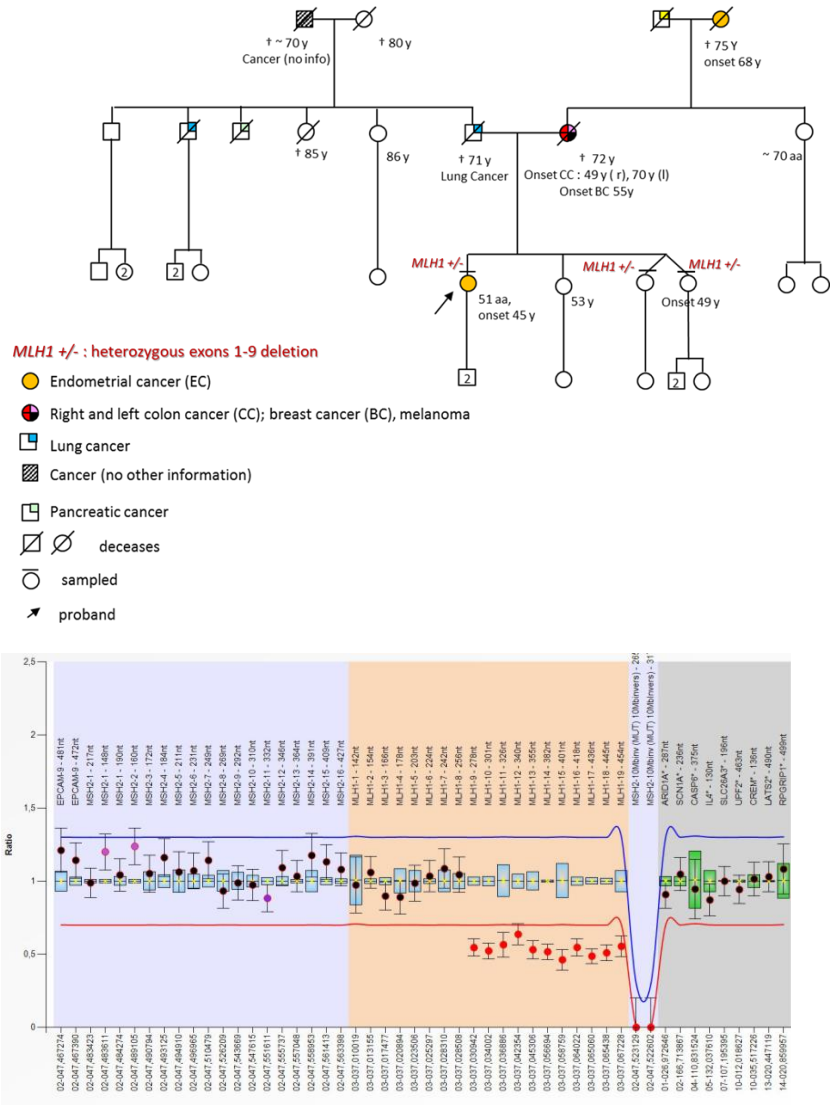

Figure 4-34 : Pedigrees P220 with Heterozygous pathogenic multiexonic deletion in MLH1 gene

P230 showed a large duplication in Chek 2 genes. She was a 49-year-old woman with breast cancer and a family history of breast and ovarian cancer(figure4-35).

The duplication including exons 6 to 13 identified by MLPA analysis on *CHEK2* was confirmed by long range PCR. For this purpose, the forward primer (LR-Forward) on intron 12 and the reverse primer (LR-Reverse) on intron 6 were designed.

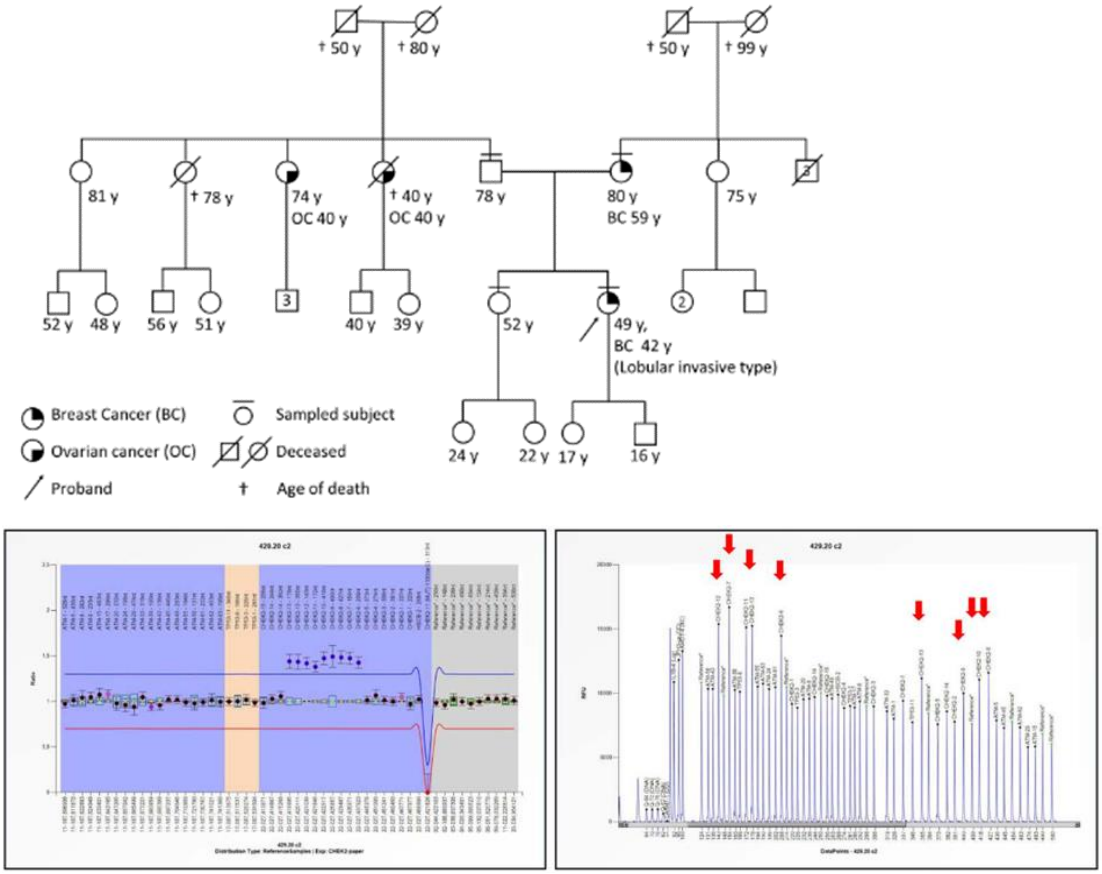

Figure 4-35: Pedigrees P220 with Heterozygous pathogenic multiexonic deletion in MLH1 gene

4.4 Genotype-Phenotype Correlations

4.4.1 Lynch Syndrome

Lynch syndrome is caused by autosomal dominant mutations to the major mismatch repair genes *MLH1, MSH2, MSH6* or *PMS2* as well as the *EPCAM* gene that inactivates *MSH2*. In addition, MUTYH-associated polyposis patients also show some phenotypic similarities to Lynch syndrome patients. MUTYH-associated polyposis patients show some phenotypic similarities to Lynch syndrome patients. The aim of this study was to investigate the prevalence of germline *MUTYH* mutations in a large series of LS patients. Lynch syndrome is associated with several cancers including endometrial, ovarian, colorectal cancer and stomach cancer.

41 of 416 patients were diagnosed with Lynch syndrome. 13 of them were male and 28 of them were female (Table4-4). 10 unique pathogenic mutations were identified on 11 probands (P6, P7, P8, P9, P20, P22, P26, P33, P37, P38 and P220). Two Patients P8 and

P9 had the same pathogenic variant (c.1216C>T, p. Arg406Ter). Three affected individuals P9, P33 and P38 carried VUS Mutations as well. 19 VUS mutations were observed in 13 Patients (P9, P10, P32, P33, P106, P108, P115, P135, P144, P160, 176, 183 and196). P10, P106 and P183 have two VUS variants, and P33 has 3 VUS mutations.

10 of the 11 probands with pathogenic variants had the mutation in related genes of lynch syndrome. In P6, P7, P22 and P37 patients were identified the LP/P variants c.3261_3262insC, c.2150_2153delTCAG, c.1605C>A and c.2931C>A in *MSH6* gene respectively. In P8 and P9 were observed c.1216C>T in *MSH2* gene.

The LP/P mutation c.1852_1854delAAG, c.210_213delAGAA in *MLH1* gene were detected on P20 and P38 respectively and Proband P220 had Heterozygous pathogenic multiexonic deletion in *MLH1* gene.

c.478C>T pathogenic mutation in PMS2 gene was seen in P33 patient. P26 was carrier of LP/P variants in APC gene (c.6709C>T). Patients P106 and 108 were carriers of VUS variants in the *MLH1* gene as well as P144 and P176 had VUS mutation in the *MSH6* gene (figure 4-34).

13 affected individuals (P37, P38, P26, P160, P6, P7, P8, P9, P,22, P33, P205, P206,P220) had family history of LS, but both of them (P205, P206) carried B/LB mutations. P206 suffered from BC.

Table 4-4: Characteristics of patients diagnosed with Lynch Syndrome

Sample ID	SEX	CLASSE	Diagnosis	AGE	Family History of Cancer (n# of 1st and 2nddegree affected relatives)	MLPA
P6	F	P	CC , (LS)	79	CC, TC	-
P7	M	P	CC , PrC, (LS)	75	Kidney, UC, AML	-
P8	F	P	CC , (LS)	44	CC, UC, BC, PaC, Kidney, LS	-
P9	F	P, VUS	CC , UC , (LS)	69	CC, BC, PaC, Kidney, LS	-
P20	M	P	CC , (LS)	58		-
P22	F	P	CC , UC , (LS)	82	CC, PrC, LC	-
P26	F	P	CC , (LS)	55	CC, TC	-
P33	M	P , 3 (VUS)	CC , Lymphoma, (LS)	66	BC, PaC, Liver C, LC	-
P37	F	P	CC, Mel , UC , (LS)	70	CC, Bladder, BC, Kidney	-
P38	F	P, VUS	CC , (LS)	53	CC, PaC, BrC	-
P106	M	(2)VUS	CC , (LS)	38	-	-
P108	F	VUS	CC , (LS)	64	-	-
P115	F	VUS	CC , (LS)	63	-	-
P135	M	VUS	CC , (LS)	61	-	lynch/ negative
P144	F	VUS	CC , (LS)	82	-	-
P160	F	VUS	BC , OC , (LS)	85	-	-
P176	F	VUS	CC , (LS)	65	-	-
P183	M	(2)VUS	CC , (LS)	50	-	-
P196	M	VUS	CC , (LS)	88	-	lynch/ negative

P10	M	(2) VUS	CC , (LS)	68	-	-
P32	F	VUS	CC , (LS)	66	-	-
P207	f	LB/B	CC , (LS)	119	-	-
P208	f	LB/B	CC , (LS)	52	-	-
P209	f	LB/B	CC , (LS)	71	-	-
P210	f	LB/B	CC , (LS)	57	-	lynch/ negative
P211	f	LB/B	CC , (LS)	74	-	
P212	f	LB/B	CC , (LS)	66	-	lynch/ negative
P213	M	LB/B	CC , (LS)	54	-	-
P214	f	LB/B	CC , (LS)	45	-	lynch/ negative
P215	M	LB/B	CC , (LS)	61	-	-
P216	M	LB/B	CC , (LS)	74	-	lynch/ negative
P217	f	LB/B	CC , (LS)	80	-	lynch/ negative
P218	f	LB/B	CC , (LS)	48	-	-
P205	f	LB/B	CC , BC (LS)	76	CC , (LS)	lynch/ negative
P219	f	LB/B	CC , (LS)	43	-	lynch/ negative
P220	f	LB/B	UC, (LS)	51	CC, BC, Mel, PaC, LC, EC	**MLH1/ positive**
P221	f	LB/B	CC , (LS)	58	-	lynch/negative
P222	f	LB/B	CC , (LS)	57	-	lynch/negative
P223	M	LB/B	CC , (LS)	80	-	lynch/negative
P224	M	LB/B	CC , (LS)	64	-	-
P225	M	LB/B	CC , (LS)	53	-	-
P206	M	LB/B	BC	85	BC , LS	Chek2/negative

BC, breast cancer; OC, ovarian cancer; PaC, pancreatic cancer; PrC, prostatic cancer; TC, Thyroid cancer; Mel, Melanoma; Gac, Gastric cancer; CC, Colon Cancer; UC, uterus cancer; LS, lynch syndrome; GIC, gastrointestinal cancer; AT, Ataxia-telangiectasia; C, cancer; pts, patients; F, female; M, male; Fam, familiarity;

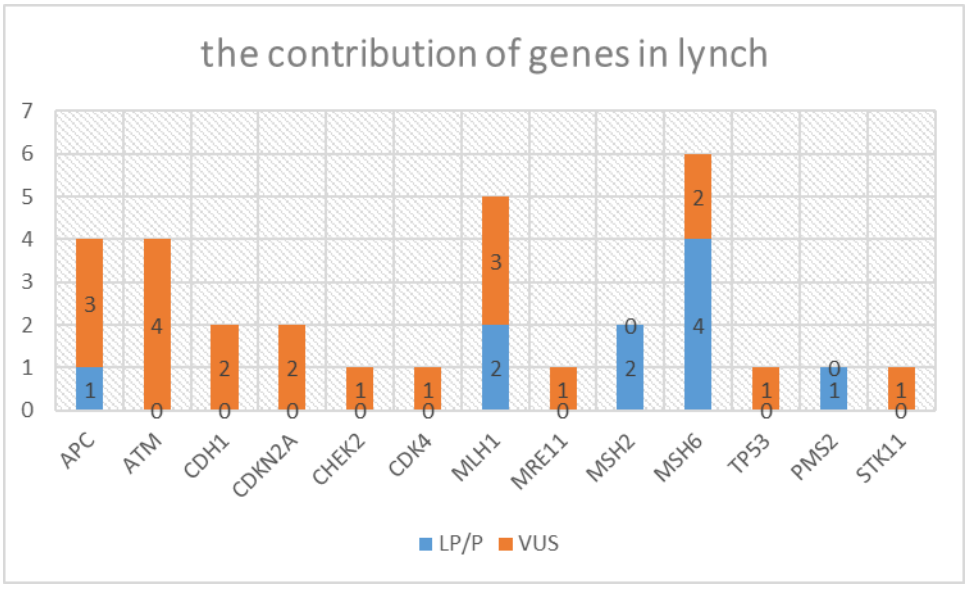

figure 4-36: the contribution of genes in lynch

Chapter 5

Discussion

5. Discussion

This study detected 37 unique P/LP variants in 17/28 DDR pathway genes in 41 of 416 probands by NGS and MLPA. The overall pathogenic variants rate in these 28 DDR genes was approximately 10% (41/416 unrelated BRCA1/2-negative cases). In addition, 161 probands carried VUS (~38.7%) (145 carriers VUS & LB/B, 16 carriers VUS & LP/P), whereas 230 (~55%) were negative (Figure 4-1). 30 of 41 proband carrying LP/P variants had a family history of cancer (~ 71.80%) which shows a strong relationship between these pathogenic mutations and hereditary cancers.

Different studies show that Mutation in 2 genes *CHEK2* and *ATM* accounted for approximately half of the mutations identified in genes other than *BRCA1/2* in Li-Fraumeni-like syndrome, BC, and other cancers, including prostatic, gastrointestinal and OC (162). In addition, the mutations in these genes confer a 2-fold to 3-fold increase in the risk of breast cancer because *CHEK2* and *ATM* are introduced as moderate-penetrance breast cancer susceptibility genes (163). The result of this study also confirmed this fact.

In current research, the most detected LP/P mutations were related to the *chek2* gene with four single variants in six patients and one large genomic arrangement in one proband, *ATM* gene with 5 mutations in five probands, *MUTYH* with 4 mutations in 5 patients, *MSH6* gene with four unique variants in four affected individuals, *RAD51* with four mutations in 4 probands, three single pathogenic variants in *MLH1* gene in 3 patients and a big deletion in this gene in one proband.

P/LP variants in *CHEK2* were detected in seven women and 5 of them suffered from mono-lateral and bilateral BC (P2, P4, P14, P15, P230). they referred to at least one or more different types of malignancies, including BC, OC, and gastrointestinal cancers, in their relatives. Moreover, the c.793-1G>A alternation was identified in a 47-year-old-female (p5) with a family history of osteosarcoma, PrC and young-onset-BC that suggested a Li-Fraumeni-like syndrome[1]. Of note, The proband had papillary thyroid cancer, which has been discussed in many studies of *CHEK2* patients (164)

The original studies of *CHEK2* mutations suggested a relationship between *CHEK2*

[1] - LFL

mutations and Li-Fraumeni syndrome, cancers of the breast, colon, ovary, endometrium, kidney, and stomach(164).

Unfortunately, patient P5 segregation studies did not be performed on the deceased parents and were not determined whether they had the variant or not. But was detected also in her asymptomatic brother. It may be inherited from the maternal branch, that they had suffered from breast, gastrointestinal, and brain cancers occurred, or from the paternal side, that three paternal uncles had PrC.

CHEK2 LP variant, c.1169A>C, p.(Tyr390Ser), were detected in a woman with a Mullerian sarcoma case (P1), referring BC and CC in relatives. The result of Monika Siołek et al study in 2015 suggested that in *CHEK2* mutations carriers the risk of sarcoma and non-Hodgkin lymphoma may also be increased (164).

c.1100delC mutation in the *Chek2* gene was observed in two women (P14, P15) with BBC. One of them suffers from diffuse mesenteric leiomyomatosis (P15). Both had a family history of breast cancer. In 2015, Monika Siołek et al reported that *chek2**1100delC was related to an increasing odds ratio of 5.7 for thyroid cancer (163). in 2021, Emma R Woodward and co-workers showed that this mutation results in ~2.5% of familial breast/ovarian cancer risk (165). in 2012, Maren Weischer et al suggested that *CHEK2* c.1100del heterozygotes have a two-fold risk of malignant melanoma compared to non-carriers (166). Nevertheless, the recurrence of BC, GC, and CC in many relatives of these two families (P14, P15) were in line with the typical phenotype of this gene.

The c.470T>C variant was found in two women with BC (P2, P4). Proband P2 also had other primitive malignancies, such as Colon Cancer, ovarian cancer and Gastric Cancer. In addition, her father had multiple cancers (rectal carcinoma, Prostate Cancer). Also, one of the cousins and grandfather had brain cancer. P4 had maternal relatives with rectal carcinoma, her sister had ovarian cancer and her mother had uterine adenocarcinoma. Unfortunately, parental segregation studies were not performed to prevent the investigation of the variation in her affected father. This mutation in the *CHEK2* gene is known that has an association with cancer susceptibility and causes a significantly increased risk of breast and colon cancer, prostate cancer, CLL, colorectal cancer and other cancers (166). CHEK2 damaging mutations have been detected in families with multiple primary cancers (164).

In a study by Davide Angeli et al in 2019 illustrated that *CHEK2* pathogenic mutation has been considered a proven genetic risk factor for male Breast Cancer (162). Several groups of researchers worldwide have confirmed this association (162), but the rate of the P/LP *CHEK2* variants in Italian male BC cases was poor (167), It was consistent with the results of this study. All detected pathogenic mutations of the *Chek2* gene were related to female patients. Only two men with VUS variants in this gene were found. Patient P78 had breast cancer and P183 had Lynch syndrome.

The an intragenic *CHEK2* duplication was detected in a 48-year-old woman (P230) ascertained for her personal and familial history of breast and ovarian cancer. The mutational analysis of the BRCA1/2 and reflex panel did not detect pathogenic variants with the NGS analysis. However, the kits used for CNV detection BRCA1 P002-D1 and BRCA2/CHEK2 P045D1 (MRC Holland, Amsterdam the Netherlands), featuring also probes for some regions where pathogenetic variants of the *CHEK2* gene occur, identified an apparent duplication of exon 9 of the *CHEK2* gene in the proband. The probes exploring exon 9 showed a relative peak ratio[1] value of about 1.5 when compared to the normal range value of reference probes (normal range: 0.7–1.3). This result was consistent with a duplication encompassing exon 9 of CHEK2, without further information on its extension. To confirm and better determine the extent of the duplication, a more comprehensive MLPA assay of the gene was performed, using a specific kit exploring all exons (1–15) of *CHEK2* (P190-D1). This allowed the identification of a large rearrangement in the *CHEK2* gene encompassing exons 6–13 on the proband. The molecular analysis of first-degree relatives confirmed the inheritance of the variant from the unaffected father and showed its presence in the asymptomatic sister. The duplication was absent in the mother, who was affected with post-menopausal breast cancer. A possibly pathogenic intragenic *CHEK2* tandem duplication encompassing exon 6–13 had already been identified by Tedaldi et al. in an unrelated Italian family(168). In the three-generation pedigree provided, there were three cases of breast cancer, a case of prostate cancer, a case of ovarian cancer and a case of leukemia. The variant was confirmed in two of the individuals who had developed breast cancer. Other individuals with cancer history were not available for analysis. The authors

[1] - RPR

demonstrated breakpoints in intron 5 (chr22:29111154, hg19) and Intron 13 (chr22:29088207) (168), however, despite the exons involved, the break points were different from the those found in the family we described, ruling out a possible founder effect in Italy.

Mono-allelic P/LP variants in this *ATM* gene are proven as moderate risk factors for malignancies, such as pancreatic, breast, prostatic, and other solid cancers. Germline ATM heterozygous carriers are approximately 0.75–1% of the population. Bi-allelic P/LP variants of this gene result in Ataxia–Telangiectasia[1], a neurodegenerative progressive disease complicated by immunodeficiency and cancer predisposition. In this research, five unique ATM pathogenic variants were detected in 5 probands with different cancers.

The P23 proband suffered from Breast Cancer and melanoma, and had a paternal family history of different types of malignancies previously described in ATM heterozygous patients; The P30 had ovarian cancer and the P40 had breast cancer. Several studies were shown that there is a notable prevalence of ATM mutations in breast and ovarian cancer families, because of this fact, ATM mutations could increase susceptibility to breast and ovarian cancer (169).

The P18 and P11 had gastric cancer and also, and P11 had a metastatic PaC. In his family, nine relatives deceased of gastric cancer. Although not frequently, pathogenic variants in ATM have been previously described in patients with GaC (170-172).

PaC remains one of the most lethal solid malignancies. The identification of damaging mutations in DDR system genes, including ATM. The combination of loss-of-function mutations in ATM is associated with pancreatic and prostate cancers and indicates the risk of breast and colorectal cancers (173). The PARP inhibitors that could have therapeutic potential in cancers with loss or mutation of ATM are opening up the possibility of new therapies.

Mutations in the *RAD51* gene, encoding a protein involved in HR, were detected in 5/416 female patients. Three of them suffered from OC, one of the three was a pathogenic mutation carrier of the *RAD51D* gene and the other carried LP/P mutations in the RAD51C gene. One woman had BC and the last woman had multiple tumours such as CC, CT, LC, and BrC. These genes, *RAD51C* and *RAD51D*, function

[1] - AT

in the same double-strand DNA break repair pathway as BRCA1 and BRCA2 and deleterious variants in them have been associated with a higher risk of epithelial ovarian carcinoma, especially with early onset (174). Although the mutations of *RAD51C* are a well-determined genetic risk factor for OC, their role of them in breast cancer is still debated (175). Peculiarly, in individuals with BC and detrimental mutations in RAD51, triple-negative cancers recurred many times (176, 177).

Mutations in the *PALB2* gene, encoding a BRCA2-interacting protein, are very common, containing 15% of the mutations in genes other than BRCA1/2. P24 with BC carried an LP/P variant of this gene. Also, 7 Probands P53, P63, P64, P111, P147, P186 and P192 had VUS mutations in the gene. All patients had BC except P186 was with Pac and Gastrophageal cancer, P147 with colon cancer, P 63 with OC and P111 had BC, OC, and PaC. Although initial reports showed that the risk of breast cancer with PALB2 of deleterious mutations was moderate, some recent researchers have declared that some variants confer a high risk. Previous studies have reported germline mutations in *PALB2* in about 3% of families with hereditary pancreatic cancer, but data are inconsistent regarding the risk of pancreatic cancer in Probands suffering from breast cancer with a germline mutation (51, 163).

Many studies suggested that *BARD1* is not only a breast cancer susceptibility gene but also a gene predisposing to triple-negative breast cancer[1]. Proband P34, carrying a pathogenic variant of *BARD1*, had breast cancer and relatives' history of breast cancer. In the study by Shimelis H et al in 2018 on 10,901 TNBC sufferers, it was appointed that *BARD1* was one of the most usual non-BRCA1/2 genes to mutate (178). Statistically, it was proven mutation in BARD1 has a significantly associated with a moderate to high risk of TNBC (0.5–0.7%). The same study indicated that the Pathogenic Variations in BARD1 were associated with a 21% risk for Caucasian patients and a 39% risk of TNBC for African American patients(178). In a different study in 2015 was done by Churpek J.E and et al on 289 African American patients, 144 of whom were patients of familial breast cancer, only one incidence of pathogenic variants in the BARD1 gene was found (178). The results of them were similar to the finding of this study.

[1] - TNBC

RECQL gene encodes a DNA helicase involved in homologous recombination repair and response to replication stress. RECQL pathogenic mutations cause to truncate or disrupted the RECQL protein or introduced missense mutations in its helicase domain(61). In this study, a man (P12) with BC and Pac as well as his relatives with BC and lung cancer was detected. This small deletion in RECQL was also identified in his healthy 49-year-old daughter. RECQL mutations may be a valuable biomarker for breast cancer. Applying novel DNA repair inhibitors targeting RECQL-related tumours may provide a new strategy for anti-cancer therapy. (61)

NBN gene encodes a protein called Nibrin, which participates in DNA repair. Mutations in the gene could result in Nijemen breakage syndrome, which may lead to an array of diseases, particularly increase susceptibility to cancer, including breast cancer. P17 was a 30-year-old- female with BC carrying LP mutation in the NBN gene (non-PVs in BRCA1/BRCA2). this mutation based on the ACMG criteria in Varsome is classified as probably pathogenetic. Unfortunately, there is not any report in the literature and is absent in the CLinVar, dbSNP and gnomAD databases. Also, P91, P129 and P174 were other probands carrying VUS mutations of the NBN gene. Several pieces of research demonstrated that polymorphic variants and defective mutations occurring in the NBN gene enhance the risk of breast cancer via the double-stranded break repair mechanism (179).

P39 with multiple tumours such as CC, CT, LC and BrC was the carrier for PV in CDKN2A. CDKN2A is well-known as a susceptibility gene for pancreatic cancer and melanoma, and also its germline variants have been associated with a broader range of neoplasms including head and neck squamous cell carcinomas, neural system tumours, breast carcinomas, as well as sarcomas.(180) T Debniak and coworkers 2005 reported that apparently, CDKN2A is a low penetrance breast cancer susceptibility gene in Poland. An association between malignant melanoma and breast cancer has been detected in several families with CDKN2A mutations in Poland, but the association should be confirmed in other populations (181).

Two other patients with breast cancer were carriers of a pathogenic mutation on genes 1 and 2, respectively. Previous studies have shown that TP53 is the most frequently mutated gene in cancer, such as breast cancer. P53 is a Protein coded

by the TP53 gene. Its activity is modulated by numerous posttranslational modifications, association with protein partners and regulators, and access to chromatin. Individuals with TP53 mutations should avoid radiation, because it may increase their risk of future malignancies. But when p53 is activated in response to chemotherapy, the constellation of targets is trans-activated and the degree of induction varies by tumour and cell type, therefore changing the fate of the cell (163, 182).

Recent studies indicated that although women carrying a *CDH1* gene mutation have an increased risk of developing lobular breast cancer, in different populations, they showed a significant relationship with other kinds of cancers such as diffuse gastric cancer or invasive ductal cancer. Probands with CDH1 mutations, who are at risk for gastric cancer, should consider prophylactic gastrectomy. Equally critical is the identification of family members without an inheritance (183).

5-10% of all colorectal cancer cases are hereditary. Hereditary colorectal cancer is divided into 2 distinct classes, hereditary non-polyposis colorectal cancers[1] and familial adenomatous polyposis[2].

Lynch syndrome (also known as hereditary nonpolyposis colorectal cancer, or HNPCC) is the most frequent hereditary colorectal cancer[3] syndrome and is characterized by germline mutations in the DNA mismatch repair[4] genes (MLH1, MSH2, MSH6, PMS2, and EPCAM) or deletion of the 3' end of EPCAM. In addition, MUTYH-associated polyposis patients also show some phenotypic similarities to Lynch syndrome patients. In FAP families, the mutation in the APC gene is related to numerous adenomatous polyps that frequently degrade into adenocarcinoma, for a lifetime risk of CRC approaching 100% and a median age at the cancer of 40 years. In addition to CRC, APC mutation carriers are at increased risk of other tumours, most notably benign desmoid tumours(184-186). In this study, 40 probands were diagnosed with lynch syndrome. 11 of them had PVs (P6, P7, P8, P9, P20, P22, P26, P33, P37, P38 and P220).

Patient P26 was the carrier for pathogenic mutation in the APC gene. both of her

[1] - HNPCC
[2] - FAP
[3] - CRC
[4] - MMR

sisters had CC and one of her uncles. In addition, mutations of the APC gene were reported in more than half of HNPCC, indicating that its mutation was a common molecular event and might play a critical role in the tumorigenesis of HNPCC(184).

numerous studies showed that *APC* mutations can be detected in both FAP and non-FAP tumours including HNPCC and sporadic colorectal ones (184). There was evidence that HNPCC lacked loss of heterozygosity[1] of the APC gene, being quite different from FAP (184). The result of the current study demonstrated one patient (P26) had PVs for APC and 16 patients were a carrier for VUS mutation in APC. P26, P183, and P10 suffered from Lynch syndrome. P134 and P25 had polyp colons. P54, P66, P72, P76, P93, P121, P127 had BC. In the end, P13, P136, P72 P143 and P204 suffered from OC, PaC, CC, GaC, and GIC respectively.

Although Mutations in APC have been detected in various cancers, including colorectal, pancreatic, and hepatocellular, they are reported to be much rarer in breast cancer(184).

Kashiwaba et al 1994 published the first report on *APC* mutations in breast cancer(187). They detected a mutation rate of 6 % (2 of 31) in primary breast carcinoma samples (188). Furuuchi et al 2000 screened almost the entire coding region using a yeast-based assay. They reported a mutation rate of 18 % (13 of 70) in primary breast cancers (189).

LP/P variants in the MMR genes *MLH1*, *MSH2* and *MSH6* were found in 11 patients (P20, P28, P38, P8, P9, P31, P6, P7, P22, P37, P220). Four probands P20, P28, P220 and P38 were carrier for PVs in MLH1. P28 suffered from BC and had a family history of breast cancer. But 3 patients P20, P220 and P28 had LS. Interestingly, albeit breast cancer is not included in the spectrum of Lynch Syndrome-related malignancies, an association between LS germline mutations and BC has been recently reported. Some researchers have suggested a higher risk of BC in affected individuals with LS and a higher frequency of MMR gene variants in BC cases (190). Studies able to clarify this association are needed also for therapeutic implications.

3 patients P8, P9 and P31 had PVs in MSH2. P8 and P9 suffered from LS (CC, UC), but P31 had GaC as well as a pathogenic mutation in the *MUTYH* gene. Carriers of

[1] - LOH

only the MSH2 mutation have a wider spectrum of malignancies such as malignancies of the small intestine, bladder, kidney, uterine cancer and CRC. All the tumours observed in known *MSH2*-only carriers occurred between the ages of 43 and 57 (185). In this study, the age range of patients carrying PVs[1] on MSH2 was about 56.5.

4 patients P6, P7, P22 and P37 had Lynch syndrome and PVs on the *MSH6* gene. Germline mutations in MSH6 account for 10%–20% of Lynch syndrome colorectal cancers caused by hereditary DNA mismatch repair gene mutations. The result of Laura Baglietto et al study in 2010 showed that MSH6 mutation carriers compared with incidence for the general population, had an eightfold increased incidence of colorectal cancer (HR = 7.6, 95% CI = 5.4 to 10.8), which was independent of sex and age. In addition, females carrying MSH6 mutation had a 26-fold increased incidence of endometrial cancer (HR = 25.5, 95% CI = 16.8 to 38.7) and a sixfold increased incidence of other cancers associated with Lynch syndrome (HR = 6.0, 95% CI = 3.4 to 10.7)(191).

in this research, P33 with Ls had PV mutation in *PMS2* as well as VUS in MLH1. The *PMS2* gene provides instructions for making a protein that plays an important role in repairing DNA. The PMS2 protein joins with another MLH1 protein to form a two-protein complex called a dimer. This complex coordinates the activities of other proteins that repair errors made during DNA replication. Variants in the *PMS2* gene have been reported in approximately 6 per cent of families with LS having an identified gene alteration(192).

In five cases P19, P21, P25, P29 and P31 had LP/P mutations in the *MUTYH* gene. P19 suffered from pancreatic cancer, P21 had BC and TC, P29, P31 and P31 also had Polipy colon, BC and GaC respectively. *MUTYH* is a DNA repair gene whose biallelic germline variants cause MUTYH-associated polyposis[2] syndrome. MAP is an autosomal recessive polyposis syndrome. Monoallelic MUTYH mutations have been reported in families with both breast and colorectal cancer. Several studies have reported extracolonic tumours in either monoallelic or biallelic carriers of MUTYH mutations such as gastroduodenal polyps and cancer, gastric polyp and

[1] - Pathogenic Variations
[2] - MAP

cancer, sebaceous adenomas and sebaceous carcinomas and brain tumours, especially meningioma (123).

There is some evidence of increased breast cancer risk in women with monoallelic variants. Overall, previous studies suggested that *MUTYH* pathogenic variants may have a role in male breast cancer[1]. In addition, they reported that MBC may be part of the tumour spectrum associated with MAP syndrome, with implications for the clinical management of patients and their relatives. A kin-cohort study of the relatives of MUTYH mutation carriers reported that monoallelic mutation carriers who had a family history of CRC were at increased risk of CRC with a hazard ratio of 2.9 (95%CI 1.2–7.0; p = 0.02). Aung Ko Win et al in 2012 reported that monoallelic mutation carriers with a family history of CRC had a twofold increased risk of CRC compared to the general population. A retrospective cohort analysis of obligate carriers of monoallelic mutations, being the parents of biallelic carriers, also estimated monoallelic mutation carriers had twice the CRC risk of the general population (SIR 2.12; 95%CI 1.30–3.28; p < 0.01)(123, 193).

VUS mutations in EPCAM have detected in 8 patients. 7 of them have BC P45, P57, P58, P59, P62, P98, and P133 and one of them suffered from gastric cancer (P179). The epithelial cell adhesion molecule gene[2] is overexpressed in some tumours, for example, in breast cancer, it has overexpressed in 30-40% of all cases and this increased expression correlates with poor prognosis (194). EPCAM mutation can influence MSH2 immunoexpression. Monoallelic deletions of the 3′ ends of EPCAM that silence the downstream gene, MSH2, cause a form of Lynch syndrome, which is a cancer predisposition syndrome associated with loss of DNA mismatch repair. Biallelic mutations in *EPCAM* lead to congenital tufting enteropathy[3], which is a rare chronic diarrheal disorder presenting in infancy. In addition, Epithelial cell adhesion molecule is often expressed in breast cancer, and its expression has been associated with poor prognosis. Recently, EPCAM has been applied as the target for antibody- and vaccine-based cancer immunotherapies. The differential association of EPCAM expression with prognosis in intrinsic subtypes has important implications for the development of EPCAM -targeted therapies in breast cancer. The use of EPCAM -specific monoclonal antibodies is a promising

[1] - MBC
[2] - EPCAM
[3] - CTE

treatment approach in these patients (195-197).

In this study, more than 30% of patients have carried VUS. This type of inconclusive result is a hard challenge to face. With the development of multi-gene panel analysis, the number of variants of unknown significance is increasing exponentially, and still, too often, their interpretation remains tangled and blurred. This raises numerous ethical and policy issues including communicating the significance of the results and possible clinical management options for patients. Woefully, in most cases, VUS lead to difficulty in risk assessment, sometimes overtreatment, and general anxiety in carriers. When any role of VUS varies and gets reclassified as potentially harmful, the practicing physicians would face the ethical and potential legal burden of contacting and explaining to patients. The currently available strategies to shed light on this issue are Bioinformatic analysis, functional studies, and periodical updates performed by international consortia (198). It is recommended that clinicians and laboratories should collaborate to guarantee periodical re-evaluations and updates on their variants to VUS carriers (198, 199).

The comparison of clinical and familial features between sufferers with and without P/LP variants brought out interesting suggestions. In this studied cohort, P/LP variants were more frequent in patients with BC (16 cases) and sufferers with gastrointestinal cancer (17 cases that 11 of whom suffered from LS). Previous reports declared that lynch syndrome accounts for 1–4% of all CRC cases (199), but current research reported it accounts for about 50% of patients with CRC. Lynch syndrome is one of the most common hereditary cancer syndromes and is characterized by the development of many cancers, such as colorectal cancer[1], endometrial cancer, ovarian cancer, stomach cancer and also, some studies have shown that other cancers, such as breast cancer, sarcoma, and adrenocortical carcinoma, occur in individuals with Lynch syndrome (199, 200).

Sigurdis Haraldsdottir et al in 2017 reported that founder mutations in MSH6 and *PMS2* prevail in Iceland, unlike most other populations(201). In the present study, Mutations in MSH6 were dominant in the Italian population. Recent advances caused further understanding of the molecular pathogenesis of this disease and help to improve diagnostic testing efficiency and surveillance strategies. Because

[1] - CRC

of that, The number of patients diagnosed with LS has recently increased.

The Increasing number of probands and the availability of segregation among relatives could give an opportunity to furtherly strengthen the association between the identified variants and the predisposition to cancer. Therefore, future studies investigating P/LP variants in DDR genes in numerous cancer cases, familial or sporadic, and in their relatives would be necessary to better discover the role of these genes in determining malignancies.

In conclusion, in this investigation, other than 13% of non-BRCA patients (298/416) are carriers of pathogenic variants in other genes (19/298), and most mutations were in genes associated with the BRCA1/2 DNA repair pathway (*CHEK2, RAD51C, ATM, MLH1, MSH2, RECQL, MUTYH, TP53, PALB2, CDKN2A*). The DDR genes panel may have an advantage for patients appropriately selected, particularly individuals with a personal or family history of more than 1 possible genetic syndrome. Therefore, Multiple-gene sequencing could notably increase their diagnostic power in them. Regarding LP/P mutations in genes involved in DNA damage response, other than BRCA, clinical guidance for such individuals is very important to ensure the development of new and more specific clinical management programs and pave the way to new therapeutic opportunities. The present of these variants explains the strong tumour recurrence in some families as well. Multiple-gene panel testing may also identify variations, like those in Lynch-related genes, that are Monitoring or prevention strategies for cancers which would not otherwise be justified by family history alone. Identifying an APC mutation in a patient with lynch syndrome and a family history of colon cancer may prompt monitoring of the role of the *APC* gene in lynch cancer. At least 30% of the mutations identified were in VUS class. Although cancer risk-management guidelines for these individuals are not yet available, identification of such mutations in combination with a family history may justify more intensive surveillance, prevention, or both as well as clinicians should anticipate and be prepared to address this with their patients. prospective research should clarify the clinical implications of cancer susceptibility genes and the role of the VUS mutations and will guide evidence-based management recommendations.

Chapter 6

Supplementary Materials

6. Supplementary Materials:

Table S 1 . Design of multigene panel.

Gene	Reference Sequence	Chr	Locus		Gene Uniformity
ATM	NM_000051.3	chr11	108117686	108236240	97.43%
PALB2	NM_024675.3	chr16	23614775	23649278	100%
MRE11A	NM_005591.3	chr11	94203632	94163157	100%
RAD50	NM_005732.3	chr5	131977865	131944422	100%
BARD1	NM_000465.3	chr2	215593395	215646238	100%
NBN	NM_002485.4	chr8	90990443	90982790	99.22%
BRIP1	NM_032043.2	chr17	59763192	59858371	96.28%
RAD51C	NM_058216.2	chr17	56809840	56787356	100%
RAD51D	NM_001142571.2	chr17	33434380	33446637	99.99%
STK11	NM_000455.4	chr19	1220575	1223176	94.69%
MSH2	NM_000251.2	chr2	47630326	47637516	99.99%
MLH1	NM_000249.3	chr3	37053497	37070428	100%
MSH6	NM_000179.2	chr2	48010368	48033795	99.49%
PMS2	NM_000535.6	chr7	6045518	6037059	100%
EPCAM	NM_002354.2	chr2	47596640	47606198	95.35%
MUTYH	NM_001128425.1	chr1	45798430	45800188	93.40%
RECQL1	NM_032941.2	chr12	21627770	21652509	91.39%
TP53	NM_000546.5	chr17	7576532	7578294	99.35%
PTEN	NM_000314.6	chr10	89711870	89725234	82.32%
CHEK2	NM_007194.3	chr22	29091693	29121117	98.88%
CDH1	NM_004360.4	chr16	68862072	68856133	98.34%
CDK4	NM_000075.3	chr12	58143232	58145130	100%
CDKN2A	NM_001195132.1	chr9	21970896	21968775	99.56%
SMAD4	NM_005359.5	chr18	48591788	48593562	100%
APC	NM_000038.5	chr5	112111321	112163708	100%

Custom Panel were designed with Ampliseq Designer V.7.0 (https://ampliseq.com/protected/startPage.action); features: 25 genes, 610 amplicons with size range 125–275 bp and size: 113.732 Kb; RefSeq, Reference Sequence; (https://www.ncbi.nlm.nih.gov/refseq/); chr, chromoso

Table S 2 . Clinical Characteristics of probands with VUS variants.

Sample ID	SEX	gene	protein	dbsnp	RESULT BRCA1/2 (NGS)	MLPA	Diagnosis	family History of Cancer n# of 1st and 2ndde Dgree affected relatives	AGE
P44	F	RAD50	p.Ile555Val	rs201120953	B		BC		58
P44	F	ATM	p.Tyr1961Cys	rs56399311	B , LB		BC		58
P45	F	EPCAM	p.His150Gln	rs864622724	B		BBC		59
P45	F	PTEN	p.?	rs370918174	B		BBC		59
P46	F	PMS2	p.Val717Met	rs201671325	B		BC		60
P47	F	CDH1	p.Ala592Thr	rs35187787	B		BBC		45
P48	F	MSH6	p.Thr327Ser	rs369568820	B , LB		BC		59
P49	F	MSH2	p.Asn596Ser	rs41295288	B		BC		57
P49	F	ATM	p.Ser49Cys	rs1800054	B		BC		57
P50	F	RAD50	p.?		B		BC , mel	BC, OTHER	49
P50	F	CHEK2	p.Ser379Cys	rs267606211	B		BC , mel	BC, OTHER	49
P51	F	MSH6	p.?|p.?	rs757825042	B		BC		64
P52	M	MRE11	p.Ala492Asp	rs61749249	B		BC		56
P53	F	PALB2	p.Gln1146Lys	rs879254033	B		BC		46
P54	F	APC	p.Asn1761Lys	rs933729249	B		BC		67
P55	F	ATM	p.?	rs3218711	B		BC		50
P56	F	BARD1	p.Cys71Ser	rs1060501308	B		BC		49
P57	F	EPCAM	p.Gln89His	rs146480420	B , VUS		BC		38
P58	F	EPCAM	p.Gln89His	rs146480420			BC		67
P59	M	EPCAM	p.?	rs114241106	B		BC		66
P60	F	MSH2	p.Ala272Val	rs34136999	B		BC		79
P3	F	RECQL	p.Pro126Leu		B		BC		44
P62	F	EPCAM	p.?	rs114241106	B		BC		42
P62	F	EPCAM	p.Ile277Met	rs115283528	B		BC		42
P63	F	PALB2	p.Gln1146Lys	rs879254033	B		OC		64
P64	F	MSH2	p.Pro616Arg	rs587779965	B		BBC	BC, OTHER	47
P64	F	PALB2	p.Tyr334Cys	rs200620434	B		BBC	BC, OTHER	47
P64	F	RAD51C	p.Arg312Trp	rs730881932	B		BBC	BC, OTHER	47
P65	F	MSH2	p.Ile145Met	rs63750124	B		BC		62
P66	F	APC	p.Glu1317Gln	rs1801166	B		BC		53
P66	F	BRIP1	p.Arg579Cys	rs28997571	B		BC		53
P67	F	BRIP1	p.(=)	rs4987050	B		OC	OC	53
P68	F	ATM	p.Leu1420Phe	rs1800058	B		BC		73
P69	F	PMS2	p.Ser418Phe	rs587782640	B , LB		BC		39
P69	F	MRE11	p.(=)	rs137868143	B , LB		BC		39
P4	F	PMS2	p.Ala49Thr	rs886039615	B		BC		51
P71	F	RAD50	p.Arg797Lys		B		BC		43
P72	F	APC	p.Arg1676Gly	rs370560998	B , LB		BBC , CC		83
P72	F	APC	p.Pro2467Thr	rs372305287	B , LB		BBC , CC		83
P73	F	ATM	p.Met1006Val	rs139893395	B		BBC		67
P73	F	ATM	p.Glu2689Gly	rs759779781	B		BBC		67
P73	F	CDH1	p.Gly879Ser	rs200911775	B		BBC		67
P74	F	MSH2	p.Gln61Pro	rs587779113	B		BC	PrC	58
P75	F	PTEN	p.?	rs370918174	B		BC		59
P76	F	APC	p.Ala2274Val	rs34919187	B		BC		41
P77	F	BARD1	p.Asn9_Arg13del	rs587781979	B		BC		43
P77	F	BARD1	p.Asn9_Arg13del	rs587781979	B		BC		43
P78	M	BARD1	p.Glu270Asp		B		BC	BC, OTHER	72
P78	M	CHEK2	p.Ser456Leu	rs876659827	B		BC	BC, OTHER	72
P79	M	MUTYH	p.?	rs890418965			BC		49
P79	M	CDH1	p.Arg43Trp				BC		49
P80	F	MLH1	p.Tyr566His	rs730881743	B		BC	BC, PrC	39
P9	F	TP53	p.Arg282Gln	rs730882008			CC , UC , (LS)	CC, BC, PaC, Kidney, (LS)	69
P82	F	CHEK2	p.?	rs587781279	B		BC		42
P83	F	STK11	p.Glu145Gly	rs369764220	B , LB		BC	BC, OTHER	59
P84	F	RAD51D	p.?	rs141690729	B		BC , PaC		74
P85	F	CHEK2	p.Glu239Lys	rs121908702	B		BC , mel		46
P86	F	ATM	p.Gly1522Ser	rs1064795495			BC		44
P87	F	CDH1	p.Ile225Thr	rs786203207	B		BC		36
P88	F	MSH6	p.(=)|p.?	rs61753796	B , LB		BC		52
P89	F	CDKN2A	p.Gly139Ser	rs587781733			BC		46
P90	F	PMS2	p.His479Asp	rs376344586	B		BC		52
P91	F	NBN	p.Arg89Gln	rs747315554	B		BC		54
P92	F	ATM	p.?	rs1333269885	B		BC		53
P93	F	APC	p.Ser2556Leu	rs761133356			BC		55
P93	F	ATM	p.His1568Arg	rs368830730			BC		55
P94	F	MSH6	p.Val734Glu	rs1060502883	B		BC		63
P95	F	PMS2	p.Gly497Asp	rs199739859	B , VUS		OC		54
P95	F	ATM	p.Cys219Arg	rs771685059	B , VUS		OC		54
P96	F	PTEN	p.?	rs201138705	B		BC		42
P96	F	BRIP1	p.Asn119Ser	rs889877039	B		BC		42
P97	F	CHEK2	p.?	rs121908700			BC		47
P98	M	MUTYH	p.Leu420Met	rs144079536			CC		56

P98	M	EPCAM	p.Phe105Tyr	rs924742537			CC		56
P98	M	CDH1	p.Thr823Ala	rs878854686			CC		56
P99	F	CDKN2A	p.His142Arg	rs759922342	B		BC		51
P100	F	ATM	p.Leu2261Val	rs757243222	B		BC		34
P13	F	APC	p.Val1789Leu		B		OC	OC, OTHER	53
P16	F	MSH6	p.Glu221Asp	rs41557217			OC		46
P103	F	MLH1	p.Leu294=	rs63751707	B		BC		59
P103	F	CDH1	p.Val493Leu	rs730881666	B		BC		59
P104	M	ATM	p.Ser978Pro	rs139552233	B , LB		BC		72
P105	F	MSH2	p.Tyr43Cys	rs17217723	B , VUS		BC		56
P106	M	MLH1	p.Phe753Ser	rs587778993			CC , (LS)		38
P106	M	ATM	p.?	rs56360226			CC , (LS)		38
P107	F	PTEN	p.Ala79Thr	rs202004587	B		BC	BC	61
P108	F	MLH1	p.Arg217Cys	rs4986984			CC , (LS)		64
P109	F	MLH1	p.Ser93Gly	rs41295282			BC		56
P110	F	MSH2	p.Ala272Val	rs34136999			OC		55
P111	F	ATM	p.Ile2511Phe	rs146069748	B		BC , OC , PaC		73
P111	F	PALB2	p.Leu278His	rs587778582	B		BC , OC , PaC		73
P112	F	CDH1	p.Ile225Thr	rs786203207	B		BC		54
P17	F	MRE11	p.Tyr82His	rs587781343			BC		30
P114	F	STK11	p.Pro326Thr	rs771632414			BBC		50
P115	F	CDK4	p.?	rs370609910			CC , (LS)		63
P116	F	SMAD4	p.Pro246Thr	rs876659967		Chek2/negative	BC		69
P117	F	BRIP1	p.Ala334Ser	rs535414791			BC , OC		69
P118	M	MLH1	p.Asn64Ser	rs63750952	B , LB		BC		54
P119	F	MSH6	p.Val1173Met	rs730881806	B		BC		47
P120	F	MLH1	p.Tyr646Cys	rs35045067	B , VUS		UC , NET PaC		68
P121	F	APC	p.His1054Pro	rs777538550	B		BC		74
P122	M	STK11	p.Arg301Gln	rs370222210	B		BC		61
P23	F	PMS2	p.His479Gln	rs63750685			BC		48
P124	F	BARD1	p.His483Leu	rs587781874	B		BC	BC, OTHER	50
P125	F	ATM	p.Ile2844Val	rs756230327	B , vus		BC		70
P126	F	ATM	p.Thr1926Ala	rs781448339	B		BC		46
P127	F	APC	p.Gln1230Arg	rs764706774	B		BC		54
P128	M	RAD50	p.Lys765Arg	rs587782573	B , VUS		PaC		34
P129	F	NBN	p.Asn142Ser	rs769414	B		BBC		64
P129	F	RAD51D	p.?	rs876659339	B		BBC		64
P130	F	MLH1	p.?	rs2308316	B , Vus		BC		52
P131	F	MLH1	p.?	rs200903126	B		BC		55
P132	F	BRIP1	p.Thr1142Arg	rs1279318199	B		BC		63
P132	F	CHEK2	p.Ser53Thr	rs371657037	B		BC		63
P133	F	EPCAM	p.His150Gln	rs864622724	B		BC		54
P133	F	MSH6	p.Pro1082Leu	rs191109849	B		BC		54
P134	F	APC	p.Leu1713Ser	rs587779797			Polipicolon		40
P135	M	CDH1	p.Arg598Gln	rs780759537		lynch/ negative	CC , (LS)		61
P136	F	MSH6	p.Pro1087Thr	rs63750998	B , VUS		PaC		47
P136	F	APC	p.Arg2566=	rs1060504883	B , VUS		PaC		47
P25	F	APC	p.Gln2265=	rs779065389			Polipicolon	Polipioclon	69
P138	F	ATM	p.Arg832Cys	rs2229022			GaC		54
P139	F	ATM	p.Ala799Val	rs199954262	B		BC		42
P140	F	ATM	p.?	rs199543313	B		BC		51
P141	F	PMS2	p.Gly207Glu	rs374704824	B		BC		51
P142	F	BARD1	p.Leu239Gln	rs200359745			BC , CC		56
P143	F	APC	p.Glu893Lys	rs199740875			GaC		61
P144	F	MSH6	p.?|p.?	rs3136363	B		CC , (LS)		82
P145	F	MLH1	p.Thr96Ala	rs770276731		lynch/ negative	CC		67
P146	F	MUTYH	p.Thr232Ser	rs587782351			BC		46
P147	F	PALB2	p.Asp871Gly	rs515726090			CC		67
P148	F	PMS2	p.Thr155Ile		B , VUS		BC , OC		52
P149	F	ATM	p.Cys532Tyr	rs35963548			BBC		55
P150	F	BARD1	p.Gly623Glu	rs587782252	B		BC		73
P151	F	MSH6	p.Cys765Ser	rs1114167712	B		BC		54
P152	F	BARD1	p.Pro454Ser	rs730881408	B		BC		48
P153	F	MSH6	p.Met703Thr	rs1064793189	B		BC		63
P154	M	MSH2	p.Tyr121Cys	rs587779971			Polipicolon		16
P155	F	MSH6	p.Thr605Ser	rs587781616		lynch/ negative	BC		60
P28	F	RET	p.?	rs1306444271			BC	BC	54
P157	F	CHEK2	p.Asp438Tyr	rs200050883	B		BC	BC	59
P29	F	RAD51D	p.?	rs876658172	B		BC	BC	54
P159	F	RET	p.?	rs1306444271	B		LMA		77
P160	F	MRE11	p.Asn556Ser	rs144896235			BC , OC , (LS)	UC, UC, BC	85
P161	F	RAD50	p.Arg1112Trp	rs773047090	B		BC	BC	73
P161	F	MEN1	p.?		B		BC	BC	73
P162	F	RET	p.?	rs1306444271	B		BC		70
P162	F	CHEK2	p.Ile364Thr	rs774179198	B		BC		70
P30	F	RET	p.?	rs1306444271			OC		81
P164	F	MRE11	p.Asn556Ser	rs144896235	B	lynch/ negative	BC	BC	43
P165	F	RET	p.?	rs1306444271		Chek2/negative	BC	BC	80
P166	M	RET	p.?	rs1306444271			BC		52

P167	F	ATM	p.Asn1081Ser	rs368111672			OC	OC	45
P51	F	MSH6	p.?|p.?	rs3136363	B		BC		64
P51	F	RET	p.?	rs1306444271	B		BC		64
P169	F	RET	p.?	rs117119161	B		BC	BC	46
P170	F	MRE11	p.Ala177Thr	rs142996063	B		BC		34
P33	M	MLH1	p.Gln426Leu	rs876659252			CC , Lymphoma, (LS)	BC, PaC, Liver C, LC	66
P33	M	ATM	p.Gln654Lys	rs528165789			CC , Lymphoma, (LS)	BC, PaC, Liver C, LC	66
P33	M	ATM	p.Leu2098Pro	rs587780631			CC , Lymphoma, (LS)	BC, PaC, Liver C, LC	66
P172	F	CDH1	p.Ala817Val	rs587782024	B , VUS		BC	BC + OC	76
P35	F	MSH2	p.Ala789Val	rs876660292			OC		41
P35	F	MSH2	p.Gln793Pro	rs876660291			OC		41
P174	M	NBN	p.?	rs756817252	B		BBC		69
P175	F	RAD50	p.Val117Phe	rs1237021808	B		BC	BC	61
P176	F	MSH6	p.Arg1334Trp|p.?	rs773763465			CC , (LS)		65
P177	F	STK11	p.Val368Met	rs1311925225	B		BC	BC	61
P178	F	RET	p.Thr1078Met	rs762952212			BC	BC + OC	56
P178	F	ATM	p.Arg23Gln	rs587779858			BC	BC + OC	56
P179	M	EPCAM	p.Arg173His|	rs771569331			GaC		82
P179	M	STK11	p.Gly394Ser	rs768780695			GaC		82
P180	F	CHEK2	p.Glu302Asp	rs587780190			BC , GaC		84
P181	F	MLH1	p.Ile657Thr	rs63750115	B , LB		BC	BC	56
P38	F	ATM	p.Arg924Trp	rs55723361			CC , (LS)	Brc, PaC, CC	53
P183	M	APC	p.Ser1371Gly				CC , (LS)		50
P183	M	CHEK2	p.Ser42Cys	rs1483975421			CC , (LS)		50
P184	F	MRE11	p.Ala177Val	rs773766504			BC		71
P185	F	APC	p.Arg1069Gly	rs375408871			BC	BC	60
P186	M	PALB2	p.Gly364Ser				GaPh C , PC		45
P187	F	BRIP1	p.Ile504=	rs876660478			GaC		46
P188	F	BRIP1	p.Trp816Cys	rs1064795352			BC		57
P189	F	BRIP1	p.Arg658Trp	rs786203170			BC		40
P190	F	STK11	p.?	rs2075606	B		BC	BC	64
P40	F	MUTYH	p.Ile223Val	rs200872702			BC	BC	48
P40	F	STK11	p.?	rs2075606			BC	BC	48
P192	M	PALB2	p.Lys294Glu	rs753676934		lynch/ negative	BC		81
P41	M	STK11	p.?	rs2075606	B		BC		74
P194	F	MRE11	p.Val38Leu	rs786202896			BC		55
P194	F	STK11	p.?	rs2075606			BC		55
P195	M	STK11	p.?	rs2075606			Cpr		72
P196	M	STK11	p.?	rs2075606		lynch/ negative	CC , (LS)		88
P197	F	STK11	p.?	rs2075606			BC		51
P198	F	STK11	p.?	rs2075606	B		BC		43
P199	F	STK11	p.?	rs2075606	B		BC		55
P200	F	BMPR1A	p.Met500Val	rs376651641	B		BC		79
P200	F	STK11	p.?	rs2075606	B		BC		79
P201	F	MSH2	p.Ile930Met	rs587779155	B		BC		60
P10	M	APC	p.Ser2531Phe				CC , (LS)		68
P203	M	APC	p.His1349Arg	rs748872251		lynch/ negative	GIC	GIC	41
P203	M	STK11	p.?	rs2075606		lynch/ negative	GIC	GIC	41

Abbreviations: BC, breast cancer; OC, ovarian cancer;PrC, prostatic cancer; LC, lung cancer; MEL, Melanoma;GaC, gastric cancer; CC, Colon Cancer; GaPh, Gastrophageal cancer; TC, tyroid cancer; PaC, pancreatic cancer; S, sarcoma; UC, uterine cancer; GIC Cancer, gastrointestinal tract cancer ; PC, peritoneal carcinosis (PC); BBC, bilateral Breat Cancer; AML, acute myeloid leukemia. B, benign; LB, likely benign, VUS, Variant of uncertain significance.

Table S 3. VUS variants detected in the sample.

gene	locus	transcript	coding	dbsnp	protein	type	function	GnomAD	count
APC	chr5:112176574	NM_000038.5	c.5283C>G	rs933729249	p.Asn1761Lys	SNV	missense	/	1
	chr5:112175240	NM_000038.5	c.3949G>C	rs1801166	p.Glu1317Gln	SNV	missense	0,004384	1
	chr5:112176317	NM_000038.5	c.5026A>G	rs370560998	p.Arg1676Gly	SNV	missense	0,0001279	1
	chr5:112178690	NM_000038.5	c.7399C>A	rs372305287	p.Pro2467Thr	SNV	missense	0,0001235	1
	chr5:112178112	NM_000038.5	c.6821C>T	rs34919187	p.Ala2274Val	SNV	missense	0,001131	1
	chr5:112178958	NM_000038.5	c.7667C>T	rs761133356	p.Ser2556Leu	SNV	missense	0,000012	1
	chr5:112176656	NM_000038.5	c.5365G>C		p.Val1789Leu	SNV	missense	/	1
	chr5:112174452	NM_000038.5	c.3161A>C	rs777538550	p.His1054Pro	SNV	missense	0,00002394	1
	chr5:112174980	NM_000038.5	c.3689A>G	rs764706774	p.Gln1230Arg	SNV	missense	0,000016	1
	chr5:112176429	NM_000038.5	c.5138T>C	rs587779797	p.Leu1713Ser	SNV	missense	0,000007996	1
	chr5:112178987	NM_000038.5	c.7696A>C	rs1060504883	p.Arg2566=	SNV	synonymous	/	1
	chr5:112178086	NM_000038.5	c.6795A>G	rs779065389	p.Gln2265=	SNV	synonymous	0,000007967	1

	chr5:112173968	NM_000038.5	c.2677G>A	rs199740875	p.Glu893Lys	SNV	missense	0,00002392	1
	chr5:112175402	NM_000038.6	c.4111A>G		p.Ser1371Gly	SNV	missense	/	1
	chr5:112174496	NM_000038.6	c.3205A>G	rs375408871	p.Arg1069Gly	SNV	missense	0,00001996	1
	chr5:112178883	NM_000038.6	c.7592C>T		p.Ser2531Phe	SNV	missense	/	1
	chr5:112175337	NM_000038.6	c.4046A>G	rs748872251	p.His1349Arg	SNV	missense	0,000003987	1
	chr11:108181006	NM_000051.3	c.5882A>G	rs56399311	p.Tyr1961Cys	SNV	missense	0,0000398	1
	chr11:108098576	NM_000051.3	c.146C>G	rs1800054	p.Ser49Cys	SNV	missense	0,007125	1
	chr11:108236264	NM_000051.3	c.9200C>G	rs3218711	p.?	SNV	unknown	0,005113	1
	chr11:108160350	NM_000051.3	c.4258C>T	rs1800058	p.Leu1420Phe	SNV	missense	0,01104	1
	chr11:108142072	NM_000051.3	c.3016A>G	rs139893395	p.Met1006Val	SNV	missense	0,00004774	1
	chr11:108205751	NM_000051.3	c.8066A>G	rs759779781	p.Glu2689Gly	SNV	missense	0,00000398	1
	chr11:108163473	NM_000051.3	c.4564G>A	rs1064795495	p.Gly1522Ser	SNV	missense	0,00000398	1
	chr11:108198368	NM_000051.3	c.6976-4A>G	rs1333269885	p.?	SNV	unknown	/	1
	chr11:108164131	NM_000051.3	c.4703A>G	rs368830730	p.His1568Arg	SNV	missense	0,0000398	1
	chr11:108114838	NM_000051.3	c.655T>C	rs771685059	p.Cys219Arg	SNV	missense	0,00000796	1
	chr11:108196245	NM_000051.3	c.6781C>G	rs757243222	p.Leu2261Val	SNV	missense	0,000004014	1
	chr11:108141988	NM_000051.3	c.2932T>C	rs139552233	p.Ser978Pro	SNV	missense	0,0009511	1
ATM	chr11:108235948	NM_000051.3	c.8987+3G>A	rs56360226	p.?	SNV	unknown	0,0001114	1
	chr11:108202186	NM_000051.3	c.7531A>T	rs146069748	p.Ile2511Phe	SNV	missense	0,000003984	1
	chr11:108216581	NM_000051.3	c.8530A>G	rs756230327	p.Ile2844Val	SNV	missense	0,00001194	1
	chr11:108180900	NM_000051.3	c.5776A>G	rs781448339	p.Thr1926Ala	SNV	missense	0,000003987	1
	chr11:108137925	NM_000051.3	c.2494C>T	rs2229022	p.Arg832Cys	SNV	missense	0,0002467	1
	chr11:108129732	NM_000051.3	c.2396C>T	rs199954262	p.Ala799Val	SNV	missense	0,00002388	1
	chr11:108143445	NM_000051.3	c.3154-4G>A	rs199543313	p.?	SNV	unknown	0,0002031	1
	chr11:108121787	NM_000051.3	c.1595G>A	rs35963548	p.Cys532Tyr	SNV	missense	0,0002558	1
	chr11:108143537	NM_000051.3	c.3242A>G	rs368111672	p.Asn1081Ser	SNV	missense	0,0001871	1
	chr11:108124602	NM_000051.4	c.1960C>A	rs528165789	p.Gln654Lys	SNV	missense	0,00002785	1
	chr11:108188194	NM_000051.4	c.6293T>C	rs587780631	p.Leu2098Pro	SNV	missense	0,000003978	1
	chr11:108098419	NM_000051.3	c.68G>A	rs587779858	p.Arg23Gln	SNV	missense	/	1
	chr11:108139268	NM_000051.4	c.2770C>T	rs55723361	p.Arg924Trp	SNV	missense	0,00004772	1
	chr2:215661789	NM_000465.3	c.211T>A	rs1060501308	p.Cys71Ser	SNV	missense	/	1
	chr2:215674253	NM_000465.3	c.26_40delACCGGCAGCCGAGGA|	rs587781979	p.Asn9_Arg13del|	INDEL	nonframeshiftDeletion|	/	1
BARD1	chr2:215674254	NM_000465.3	c.25_39delAACCGGCAGCCGAGG|	rs587781979	p.Asn9_Arg13del|	INDEL	nonframeshiftDeletion|	/	1
	chr2:215645788	NM_000465.3	c.810A>C		p.Glu270Asp	SNV	missense	/	1
	chr2:215632326	NM_000465.3	c.1448A>T	rs587781874	p.His483Leu	SNV	missense	0,00001592	1
	chr2:215645882	NM_000465.3	c.716T>A	rs200359745	p.Leu239Gln	SNV	missense	0,0001012	1
	chr2:215609826	NM_000465.3	c.1868G>A	rs587782252	p.Gly623Glu	SNV	missense	0,00001194	1
	chr2:215633991	NM_000465.3	c.1360C>T	rs730881408	p.Pro454Ser	SNV	missense	0,000003979	1
BMPR1A	chr10:88683375	NM_004329.3	c.1498A>G	rs376651641	p.Met500Val	SNV	missense	0,00003181	1
	chr17:59858260	NM_032043.2	c.1735C>T	rs28997571	p.Arg579Cys	SNV	missense	0,00008352	1
	chr17:59760948	NM_032043.2	c.3459T>C	rs4987050	p.(=)	SNV	synonymous	0,0007679	1
	chr17:59934442	NM_032043.2	c.356A>G	rs889877039	p.Asn119Ser	SNV	missense	0,000137	1
BRIP1	chr17:59878754	NM_032043.2	c.1000G>T	rs535414791	p.Ala334Ser	SNV	missense	0,00001194	1
	chr17:59760982	NM_032043.2	c.3425C>G	rs1279318199	p.Thr1142Arg	SNV	missense	0,000004007	1
	chr17:59861747	NM_032043.3	c.1512C>A	rs876660478	p.Ile504=	SNV	synonymous	/	1
	chr17:59793356	NM_032043.3	c.2448G>T	rs1064795352	p.Trp816Cys	SNV	missense	/	1
	chr17:59853887	NM_032043.3	c.1972C>T	rs786203170	p.Arg658Trp	SNV	missense	0,00003185	1
	chr16:68855966	NM_004360.4	c.1774G>A	rs35187787	p.Ala592Thr	SNV	missense	0,003253	1
	chr16:68867388	NM_004360.4	c.2635G>A	rs200911775	p.Gly879Ser	SNV	missense	0,0001273	1
	chr16:68772278	NM_004360.4	c.127C>T		p.Arg43Trp	SNV	missense	/	1
	chr16:68842738	NM_004360.4	c.674T>C	rs786203207	p.Ile225Thr	SNV	missense	/	1
CDH1	chr16:68867220	NM_004360.4	c.2467A>G	rs878854686	p.Thr823Ala	SNV	missense	/	1
	chr16:68849574	NM_004360.4	c.1477G>C	rs730881666	p.Val493Leu	SNV	missense	0,00003181	1
	chr16:68842738	NM_004360.4	c.674T>C	rs786203207	p.Ile225Thr	SNV	missense	/	1
	chr16:68855985	NM_004360.4	c.1793G>A	rs780759537	p.Arg598Gln	SNV	missense	0,00001591	1
	chr16:68867203	NM_004360.5	c.2450C>T	rs587782024	p.Ala817Val	SNV	missense	0,00002386	1
CDK4	chr12:58143104	NM_000075.3	c.684-4A>G|	rs370609910	p.?|	SNV	unknown|	0,000007959	1
CDKN2A	chr9:21970943	NM_001195132.1	c.415G>A	rs587781733	p.Gly139Ser	SNV	missense	/	1
	chr9:21970933	NM_001195132.1	c.425A>G	rs759922342	p.His142Arg	SNV	missense	0,00001215	1
	chr22:29091821	NM_007194.3	c.1136C>G	rs267606211	p.Ser379Cys	SNV	missense	/	1
	chr22:29091123	NM_007194.3	c.1367C>T	rs876659827	p.Ser456Leu	SNV	missense	/	1
	chr22:29121228	NM_007194.3	c.444+3A>G	rs587781279	p.?	SNV	unknown	/	1
	chr22:29107974	NM_007194.3	c.715G>A	rs121908702	p.Glu239Lys	SNV	missense	/	1
CHEK2	chr22:29121360	NM_007194.4	c.320-5T>A	rs121908700	p.?	SNV	unknown	0,0005702	1
	chr22:29130553	NM_007194.4	c.157T>A	rs371657037	p.Ser53Thr	SNV	missense	0,00004773	1
	chr22:29091178	NM_007194.4	c.1312G>T	rs200050883	p.Asp438Tyr	SNV	missense	0,0003864	1
	chr22:29092893	NM_007194.4	c.1091T>C	rs774179198	p.Ile364Thr	SNV	missense	0,00002388	1
	chr22:29099495	NM_007194.4	c.906A>C	rs587780190	p.Glu302Asp	SNV	missense	0,0000296	1

	chr22:29130585	NM_007194.4	c.125C>G	rs1483975421	p.Ser42Cys	SNV	missense	/	1
EPCAM	chr2:47602397	NM_002354.2	c.450C>G	rs864622724	p.His150Gln	SNV	missense	/	2
	chr2:47601029	NM_002354.2	c.267G>C	rs146480420	p.Gln89His	SNV	missense	0,00272	2
	chr2:47641430	NM_002354.2	c.77-11T>A	rs114241106	p.?	SNV	unknown	0,008101	2
	chr2:47601029	NM_002354.2	c.831A>G	rs115283528	p.Ile277Met	SNV	missense	0,002174	1
	chr2:47601076	NM_002354.2	c.314T>A	rs924742537	p.Phe105Tyr	SNV	missense	0,000003977	1
	chr2:47602397	NM_002354.2	c.450C>G	rs864622724	p.His150Gln	SNV	missense	/	1
	chr2:47604179	NM_002354.2	c.518G>A|	rs771569331	p.Arg173His|	SNV	missense|	0,00006366	1
MEN1	chr11:64572305	NM_000244.3	c.1366-17T>C		p.?	SNV	unknown	/	1
	chr3:37083787	NM_000249.3	c.1696T>C	rs730881743	p.Tyr566His	SNV	missense	/	1
	chr3:37059088	NM_000249.3	c.882C>G	rs63751707	p.Leu294=	SNV	synonymous	0,00006366	1
	chr3:37092131	NM_000249.3	c.2258T>C	rs587778993	p.Phe753Ser	SNV	missense	/	1
	chr3:37053562	NM_000249.3	c.649C>T	rs4986984	p.Arg217Cys	SNV	missense	0,0003703	1
	chr3:37042515	NM_000249.3	c.277A>G	rs41295282	p.Ser93Gly	SNV	missense	0,00001591	1
	chr3:37038184	NM_000249.3	c.191A>G	rs63750952	p.Asn64Ser	SNV	missense	0,00003579	1
	chr3:37090048	NM_000249.3	c.1937A>G	rs35045067	p.Tyr646Cys	SNV	missense	0,00004377	1
	chr3:37089991	NM_000249.3	c.1897-17C>G	rs2308316	p.?	SNV	unknown	0,0002427	1
	chr3:37092176	NM_000249.3	c.*32C>T	rs200903126	p.?	SNV	unknown	0,001578	1
	chr3:37042524	NM_000249.3	c.286A>G	rs770276731	p.Thr96Ala	SNV	missense	0,00003182	1
	chr3:37067366	NM_000249.4	c.1277A>T	rs876659252	p.Gln426Leu	SNV	missense	/	1
	chr3:37090081	NM_000249.4	c.1970T>C	rs63750115	p.Ile657Thr	SNV	missense	/	1
MRE11	chr11:94192599	NM_005591.3	c.1475C>A	rs61749249	p.Ala492Asp	SNV	missense	0,003449	1
	chr11:94204763	NM_005591.3	c.822T>C	rs137868143	p.(=)	SNV	synonymous	0,001403	1
	chr11:94219160	NM_005591.3	c.244T>C	rs587781343	p.Tyr82His	SNV	missense	0,00001991	1
	chr11:94180501	NM_005591.3	c.1667A>G	rs144896235	p.Asn556Ser	SNV	missense	0,0001234	2
	chr11:94211916	NM_005591.3	c.529G>A	rs142996063	p.Ala177Thr	SNV	missense	0,0002112	1
	chr11:94211915	NM_005591.4	c.530C>T	rs773766504	p.Ala177Val	SNV	missense	0,00003987	1
	chr11:94224040	NM_005591.4	c.112G>T	rs786202896	p.Val38Leu	SNV	missense	0,000003979	1
MSH2	chr2:47702191	NM_000251.2	c.1787A>G	rs41295288	p.Asn596Ser	SNV	missense	0,000318	1
	chr12:21643150	NM_000251.2	c.815C>T	rs34136999	p.Ala272Val	SNV	missense	0,0003031	2
	chr2:47702251	NM_000251.2	c.1847C>G	rs587779965	p.Pro616Arg	SNV	missense	/	1
	chr2:47637301	NM_000251.2	c.435T>G	rs63750124	p.Ile145Met	SNV	missense	0,0003222	1
	chr2:47630512	NM_000251.2	c.182A>C	rs587779113	p.Gln61Pro	SNV	missense	/	1
	chr2:47630458	NM_000251.2	c.128A>G	rs17217723	p.Tyr43Cys	SNV	missense	0,0000696	1
	chr2:47635690	NM_000251.2	c.362A>G	rs587779971	p.Tyr121Cys	SNV	missense	0,00003191	1
	chr2:47705566	NM_000251.3	c.2366C>T	rs876660292	p.Ala789Val	SNV	missense	/	1
	chr2:47705578	NM_000251.3	c.2378A>C	rs876660291	p.Gln793Pro	SNV	missense	/	1
	chr2:47710073	NM_000251.3	c.2790A>G	rs587779155	p.Ile930Met	SNV	missense	0,000004024	1
MSH6	chr2:48026102	NM_000179.2	c.980C>G	rs369568820	p.Thr327Ser	SNV	missense	0,00001989	1
	chr2:48033273	NM_000179.2	c.3647-70TTTTTG>T	rs757825042	p.?|p.?	INDEL	unknown|unknown	/	1
	chr2:48033725	NM_000179.2	c.3936T>C	rs61753796	p.(=)|p.?	SNV	synonymous|unknown	0,000012	1
	chr2:48027230	NM_000179.2	c.2108T>C	rs1064793189	p.Met703Thr	SNV	missense	/	1
	chr2:48027323	NM_000179.2	c.2201T>A	rs1060502883	p.Val734Glu	SNV	missense	/	1
	chr2:48025785	NM_000179.2	c.663A>C	rs41557217	p.Glu221Asp	SNV	missense	0,0007059	1
	chr2:48032127	NM_000179.2	c.3517G>A	rs730881806	p.Val1173Met	SNV	missense	0,000003979	1
	chr2:48030631	NM_000179.2	c.3245C>T	rs191109849	p.Pro1082Leu	SNV	missense	0,0003381	1
	chr2:48030645	NM_000179.2	c.3259C>A	rs63750998	p.Pro1087Thr	SNV	missense	0,00009545	1
	chr2:48033273	NM_000179.2	c.3647-69_3647-65delTTTTG	rs3136363	p.?|p.?	INDEL	unknown|unknown	0,002553	2
	chr2:48027416	NM_000179.2	c.2294G>C	rs1114167712	p.Cys765Ser	SNV	missense	/	1
	chr2:48026936	NM_000179.2	c.1814C>G	rs587781616	p.Thr605Ser	SNV	missense	0,00005257	1
MUTYH	chr2:48033789	NM_000179.2	c.4000C>T	rs773763465	p.Arg1334Trp|p.?	SNV	missense|unknown	0,0000368	1
	chr1:45798518	NM_001128425.1	c.505-12T>G	rs890418965	p.?	SNV	unknown	/	1
	chr1:45797157	NM_001128425.1	c.1258C>A	rs144079536	p.Leu420Met	SNV	missense	0,0005853	1
	chr1:45798157	NM_001128425.1	c.694A>T	rs587782351	p.Thr232Ser	SNV	missense	0,000003989	1
NBN	chr1:45798269	NM_001128425.2	c.667A>G	rs200872702	p.Ile223Val	SNV	missense	0,0002188	1
	chr8:90993657	NM_002485.4	c.266G>A	rs747315554	p.Arg89Gln	SNV	missense	0,00001193	1
	chr8:90993017	NM_002485.4	c.425A>G	rs769414	p.Asn142Ser	SNV	missense	0,00001215	1
PALB2	chr8:90992959	NM_002485.5	c.480+3A>G	rs756817252	p.?	SNV	unknown	0,000003982	1
	chr16:23614905	NM_024675.3	c.3436C>A	rs879254033	p.Gln1146Lys	SNV	missense	/	2
	chr16:23646987	NM_024675.4	c.880A>G	rs753676934	p.Lys294Glu	SNV	missense	0,000003978	1
	chr16:23646866	NM_024675.3	c.1001A>G	rs200620434	p.Tyr334Cys	SNV	missense	0,0001116	1
	chr16:23647033	NM_024675.3	c.833_834delTAinsAT	rs587778582	p.Leu278His	MNV	missense	/	1
PMS2	chr16:23637693	NM_024675.3	c.2612A>G	rs515726090	p.Asp871Gly	SNV	missense	0,00001591	1
	chr16:23646777	NM_024675.4	c.1090G>A		p.Gly364Ser	SNV	missense	/	1
	chr7:6022480	NM_000535.6	c.2149G>A	rs201671325	p.Val717Met	SNV	missense	0	1
	chr7:6027143	NM_000535.6	c.1253C>T	rs587782640	p.Ser418Phe	SNV	missense	/	1

	chr7:6045541	NM_000535.6	c.145G>A	rs886039615	p.Ala49Thr	SNV	missense	/	1
	chr7:6026961	NM_000535.6	c.1435C>G	rs376344586	p.His479Asp	SNV	missense	0,00001591	1
	chr7:6026906	NM_000535.6	c.1490G>A	rs199739859	p.Gly497Asp	SNV	missense	0,0001153	1
	chr7:6026959	NM_000535.6	c.1437C>G	rs63750685	p.His479Gln	SNV	missense	0,004243	1
PTEN	chr7:6038824	NM_000535.6	c.620G>A	rs374704824	p.Gly207Glu	SNV	missense	0,0003738	1
	chr7:6042157	NM_000535.6	c.464C>T		p.Thr155Ile	SNV	missense	/	1
	chr10:89690764	NM_000314.6	c.210-39A>G	rs370918174	p.?	SNV	unknown	0,002507	2
RAD50	chr10:89717588	NM_000314.6	c.635-22T>C	rs201138705	p.?	SNV	unknown	/	1
	chr10:89690828	NM_000314.6	c.235G>A	rs202004587	p.Ala79Thr	SNV	missense	0,0001079	1
	chr5:131927596	NM_005732.3	c.1663A>G	rs201120953	p.Ile555Val	SNV	missense	0,00007976	1
	chr5:131925536	NM_005732.3	c.1452+7T>G		p.?	SNV	unknown	/	1
	chr5:131939174	NM_005732.3	c.2390G>A		p.Arg797Lys	SNV	missense	/	1
	chr5:131939078	NM_005732.3	c.2294A>G	rs587782573	p.Lys765Arg	SNV	missense	0,000007967	1
RAD51C	chr5:131953931	NM_005732.3	c.3334A>T	rs773047090	p.Arg1112Trp	SNV	missense	0,00001196	1
RAD51D	chr5:131911604	NM_005732.4	c.349G>T	rs1237021808	p.Val117Phe	SNV	missense	0,00003184	1
	chr17:56801430	NM_058216.2	c.934C>T	rs730881932	p.Arg312Trp	SNV	missense	0,00001	1
	chr17:33434075	NR_037714.1|NM_133629.2	|c.145-575A>G	rs141690729	|p.?	SNV	|unknown	0,00003579	1
RECQL	chr17:33433504	NR_037714.1|NM_133629.2	|c.145-4T>G	rs876659339	|p.?	SNV	|unknown	/	1
RET	chr17:33445622	NR_037714.1|NM_133629.2	|c.144+508G>T	rs876658172	|p.?	SNV	|unknown	0,000007961	1
	chr2:47600591	NM_032941.2	c.377C>T		p.Pro126Leu	SNV	missense	/	1
	chr10:43624308	NM_020975.5	c.*591A>AT	rs1306444271	p.?	INDEL	unknown	0,00006429	7
SMAD4	chr10:43625361	NM_020975.5	c.*1644G>C	rs117119161	p.?	SNV	unknown	0,0002866	1
STK11	chr10:43623605	NM_020975.5	c.3233C>T	rs762952212	p.Thr1078Met	SNV	missense	0,00002386	1
	chr18:48584563	NM_005359.5	c.736C>A	rs876659967	p.Pro246Thr	SNV	missense	0,000007954	1
	chr19:1219382	NM_000455.4	c.434A>G	rs369764220	p.Glu145Gly	SNV	missense	0,0000134	1
	chr19:1223039	NM_000455.4	c.976C>A	rs771632414	p.Pro326Thr	SNV	missense	0,000004488	1
	chr19:1221987	NM_000455.4	c.902G>A	rs370222210	p.Arg301Gln	SNV	missense	0,00002223	1
	chr19:1223165	NM_000455.4	c.1102G>A	rs1311925225	p.Val368Met	SNV	missense	0,000008253	1
TP53	chr19:1226524	NM_000455.4	c.1180G>A	rs768780695	p.Gly394Ser	SNV	missense	0,0000169	1
	chr19:1220321	NM_000455.5	c.465-51T>C	rs2075606	p.?	SNV	unknown	0,2972	11
	chr17:7577093	NM_000546.5	c.845G>A	rs730882008	p.Arg282Gln	SNV	missense	0,000003977	1

VUS variants detected by 25 genes of cancer panel among 418 patients with a history familial/personal of cancer. Abbreviations: dbSNP, Single Nucleotide Polymorphism Database (https://www.ncbi.nlm.nih.gov/snp/); rs, reference SNP; HGVS: Human Genome Variation Society (http://www.HGVS.org/varnomen); GnomAD, Genome Aggregation Database (https://gnomad.broadinstitute.org/); ACMG, American College of Medical Genetics and Genomics; P, pathogenic; LP, likely pathogenic. Variants were annotated according to the current HGVS nomenclature; p.?, consequence on protein structure unknown.

6.1 Genotype-Phenotype correlations by snpXplorer

Genetic association studies are often applied to study the genetic basis of numerous human phenotypes. In this study, snpXplorer online software was used in order to increase the speed of review and integration of SNP annotations. snpXplorer combined association statistics from multiple studies and demonstrated regional information which includes SNP associations, recombination rates, structural variations, linkage disequilibrium patterns, eQTL[1], genes and gene expressions per tissue.

Regarding to a list of SNPs, snpXplorer performed variant-to-gene mapping and gene-set enrichment analysis aimed to identify molecular pathways that were

[1] - expression quantitative-trait-loci

overrepresented in the list of input SNPs. snpXplorer is freely available at https://snpxplorer.net.

The list of SNPs was inputted into snpXplorer software, Then they were run a functional annotation and enrichment analysis, and send the results by email.

The variant identifiers (chr:pos) pasted in the annotation section of snpXplorer, specifying chr:posas input type, Gene Ontology[1] and Reactome as gene-sets for the enrichment analysis, and Blood as GTEx tissue for eQTL (i.e. the default value).

The analysis showed that the N = 1551 variants were linked to a total of 28genes, with N = over 250 variants mapping to one gene, N = 70 variants mapping to two genes, N = 50 variants mapping to three genes, N = 20 variants mapping to four genes, N = 1 variant mapping to five genes, N = 4 variants mapping to 7, 8 genes (figure4-35 (b)). N = 107 variants were found to be coding variants and N = 246 variants were annotated based on their genomic position (figure 4-35 (a , d)). These plots not only inform the user about the effect of the SNPs of interest (such as a direct consequence on the protein sequence in case of coding SNPs, or a regulatory effect in case of eQTLs or intergenic SNPs), but also suggest the presence of more complex regions indicates, the number of genes associated with each SNP,

[1] - GO

(which normally increases for complex, gene-dense regions such as HLA-region or IGH-region) (figure4-35).

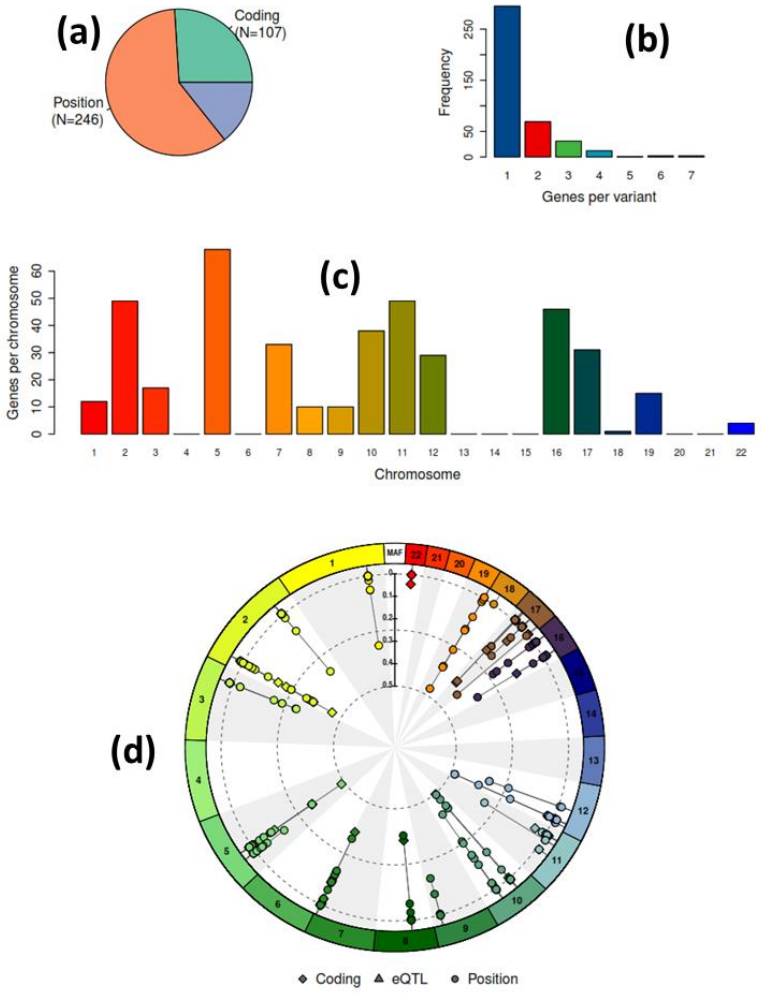

Figure S 1 : SNP- gene- mapping. A: The circular summary shows the frequency, the type and the chromosomal distribution of all input variants. b: barplot shows the number of genes associated with each variant. c: The central barplot shows the chromosomal distribution of all input variants. D: The circular plot shows the type of annotation of each genetic variant used as input (coding, eQTL or annotated by their positions) and also each variant's minor allele frequency and chromosomal distribution

With the resulting list of input SNPs and (likely) associated genes, the GWAS-Catalog and the datasets of structural variations for previously reported associations were compared. On one hand, we found a marked enrichment in the GWAS-Catalog-gene- overlap for Prostate Carcinoma, Type II Diabetes Mellitus, and Mean Corpuscular Hemoglobin Concentration, on other hand, it was observed a marked enrichment in the GWAS-Catalog-SNP-overlap for Breast Carcinoma, Melanoma, Body Mass Index, Lymphocyte Count and Uterine Fibroid (figure 4-36). The results of this analysis are practical in order to indicate other traits that were previously associated with the input SNPs. Such as, relationships between different traits were

discovered. in this study, it was suggested the involvement of Lymphocyte Count and Hemoglobin Concentration in cancer, is a known relationship.

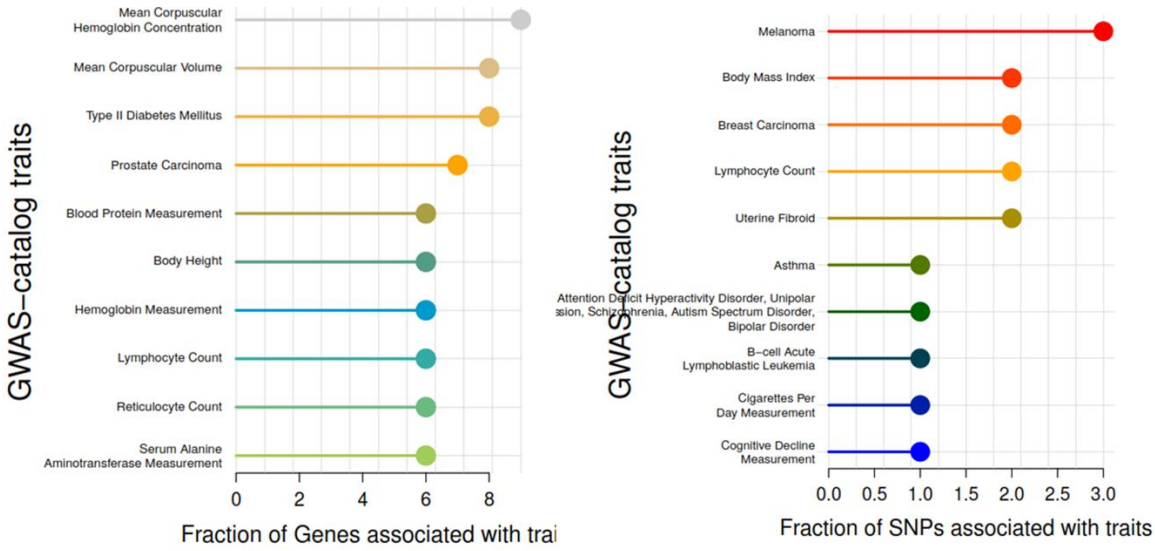

Figure S 2: Fraction of SNPs and Genes association with traits in the GWAS Catalog. The left graph shows the Fraction of genes (associated with input SNPs) previously associated with traits in the GWAS Catalog. Right graph shows the number of input SNPs was previously associated with traits in the GWAS catalogue.

Next, the software searched for all structural variations in a region of 10kb surrounding the input SNPs and detected that for 242 SNPs, a larger structural variation was present in the vicinity, including in three genes *PMS2*, *CHEK2* and *STK11* were found.

However, technological improvements now cause it possible to exactly measure SV alleles, these regions have been largely unexplored due to the complex nature of large SVs. This information on the SNP and SVs overlap is useful to investigate the functional effect of SVs, and could be applied to prioritize certain genomic regions. Finally, It was done a (sampling-based) gene-set enrichment analysis to detect molecular pathways enriched within the set of genes associated with the input variants. The gene-set enrichment analysis is performed using the Gost function from the R package gprofiler2. several gene-set sources, such as Gene Ontology(GO:BP), KEGG, Reactome, and Wiki-pathways were selected and Blood as tissue for the eQTL analysis. The full table of the gene-set enrichment analysis comprising all tested terms and their relative sampling-based p-values was performed. To simplify the interpretation of the gene-set enrichment results, the snpXplorer clustered the significantly enriched terms from GO based on a semantic similarity

measure by using REVIGO[1] (Figure4-37) and a term-based clustering approach (Figure 4-38)(Table S5).

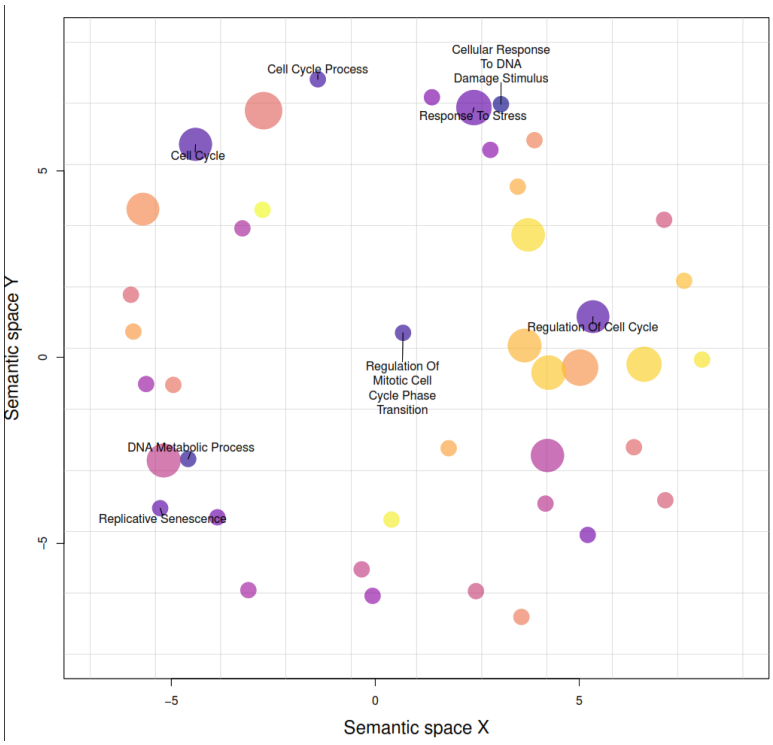

Figure S 3: clustring-GO-term. REVIGO plot demonstrated the remaining GO terms after removing redundancy based on a semantic similarity measure. The colour of each dot shows the significance (the darker, the further significant), while the size of the dot codes for the number of similar terms removed from REVIGO

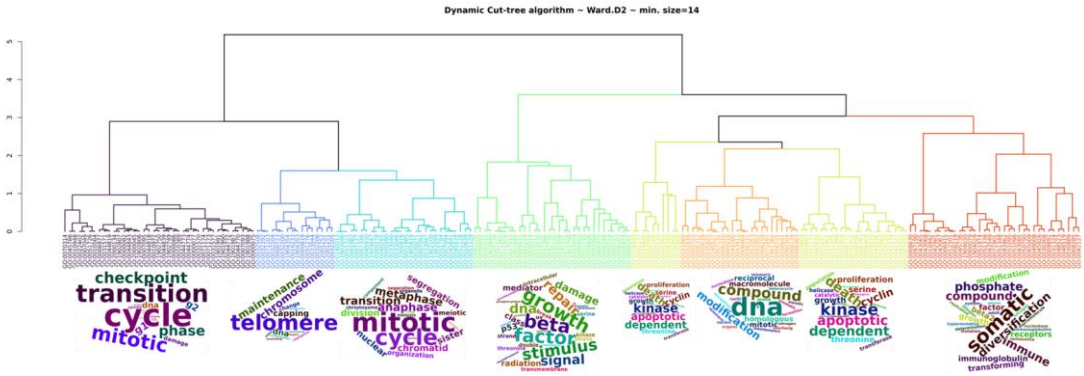

Figure S 4: Results of the term-based clustering approach. it was applied Lin as a semantic similarity measure to calculate the similarity between all GO terms. Then, It was used Ward-d2 as a clustering algorithm, and a dynamic cut tree algorithm to highlight clusters. For each cluster, it was applied word clouds of the most frequent words to describe each cluster.

Both methods are helpful because they give an overview of the most relevant biological processes related to the input SNPs. Word clouds help the interpretation of the set of GO terms of each cluster, and they were generated regarding the previous

[1] - reduce + visualize gene ontology

studies of the input SNPs (figure 4-38). In This study, the clustering approach identified seven essential clusters of GO terms (Figure4-38, 4-39).

The snpXplorer by REVIGO tools reduced redundant terms based on a semantic similarity measure and demonstrated enrichment outcomes in an embedded space through eigenvalue decomposition of the pairwise distance matrix. Also, the Lin tool was used as a semantic distance measure for REVIGO and a term-based clustering approach.

Firstly, a semantic similarity matrix between all enriched terms was calculated, and subsequently hierarchical clustering to the obtained distance matrix was applied. Finally, the semantic similarity heatmap was generated (Figure4-33). All tables describing REVIGO analysis and a term-based clustering-approach were brought in the supplementary chapter.

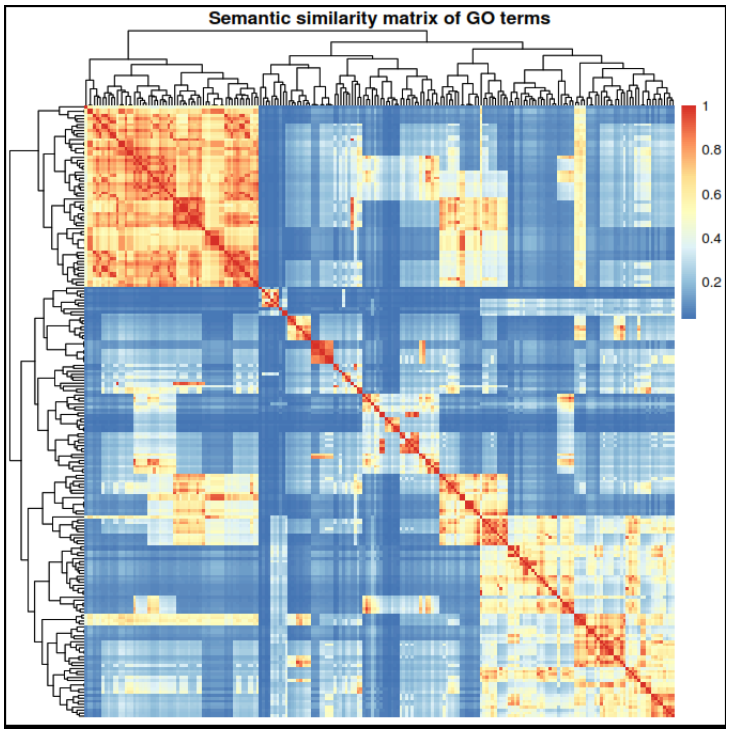

Figure S 5: Semantic similarity matrix (Pheatmap). The pheatmap shows the semantic similarity matrix values between all significantly enriched Gene Ontology biological processes terms (y-axis) and using snpXplorer functional annotation section (x-axis). The terms of study (x-axis) are ordered based on their assigned cluster as a result of the term-based clustering approach. Substantial similarity patterns are noticeable in the pheatmap that red cluster are cluster 1 and have more No suggestions. The more similar GO terms are, they are red, and the less similar, they are blue.

6.2 Siteography

- Enigma database https://brcaexchange.org/variants
- Clinvar database https://www.ncbi.nlm.nih.gov/clinvar/
- Snp database https://www.ncbi.nlm.nih.gov/snp
- Genome Aggregation Database (gnomAD) https://gnomad.broadinstitute.org
- Revigo http://revigo.irb.hr/
- snpxplorer http://snpxplorer.net/
- reactome https://reactome.org/
- Human Splicing Finder (HSF) software prediction http://umd.be/Redirect.html
- Human Genome Variation Society (HGVS) https://varnomen.hgvs.org/
- Insight database (https://www.insight-group.org/variants/databases/).
- IGV (Integrative Genomics Viewer) softwarehttp://software.broadinstitute.org/software/igv/
- Ion Ampliseq designer software https://www.ampliseq.com/login/login.action
- LSDBs (Locus-Specific Mutation Databases) https://grenada.lumc.nl/LSDB_list/lsdbs
- Mutation Taster software prediction http://www.mutationtaster.org
- (NCCN) Guidelineshttps://www.nccn.org/professionals/physician_gls/default.aspx
- Proveansoftware prediction http://provean.jcvi.org/seq_submit.php
- Reference Sequence chromosome; (https://www.ncbi.nlm.nih.gov/refseq/
- SIFT software prediction https://sift.bii.a-star.edu.sg
- Varsome database https://varsome.com/

Chapter 7

Reference

7. References

1. Loeb KR, Loeb LA. Significance of multiple mutations in cancer. Carcinogenesis. 2000;21(3):379-85.
2. Curtin NJ. DNA repair dysregulation from cancer driver to therapeutic target. Nature Reviews Cancer. 2012;12(12):801-17.
3. Howard MJ, Wilson SH. DNA scanning by base excision repair enzymes and implications for pathway coordination. DNA repair. 2018;71:101-7.
4. Takata Ki, Wood RD. Bypass specialists operate together. The EMBO Journal. 2009;28(4):313-4.
5. Rajewsky M, Engelbergs J, Thomale J, Schweer T. Relevance of DNA repair to carcinogenesis and cancer therapy. Genes and Environment in Cancer. 1998:127-46.
6. Mogos RAM, Popovici R, Tanase AE, Calistru T, Popovici P, Grigore M, et al. New approaches in ovarian cancer based on genetics and carcinogenesis hypotheses. Experimental and Therapeutic Medicine. 2022;23(6):1-4.
7. Bartek J. DNA damage response, genetic instability and cancer: from mechanistic insights to personalized treatment. Molecular oncology. 2011;5(4):303.
8. Marcello M-S, Monica B, De Maria Marchiano R. DNA Damage Repair Pathways in Cancer Stem Cells. 2012.
9. Denissenko MF, Pao A, Tang M-s, Pfeifer GP. Preferential formation of benzo [a] pyrene adducts at lung cancer mutational hotspots in P53. Science. 1996;274(5286):430-2.
10. Spivak G. Nucleotide excision repair in humans. DNA repair. 2015;36:13-8.
11. Li Z, Pearlman AH, Hsieh P. DNA mismatch repair and the DNA damage response. DNA repair. 2016;38:94-101.
12. de Boer J, van der Horst G, Hoeijmakers J. Mutant mice as a model of human diseases of aging. Defective transcription and premature aging. Tijdschrift Voor Gerontologie en Geriatrie. 1999;30(4):168-70.
13. Ye L, Morishima A, Kohara K, Miki T. Helicases and aging. Cellular and Molecular Life Sciences CMLS. 2000;57(5):716-30.
14. Chakraverty RK, Hickson ID. Defending genome integrity during DNA replication: a proposed role for RecQ family helicases. Bioessays. 1999;21(4):286-94.
15. Shen J-C, Loeb LA. Werner syndrome exonuclease catalyzes structure-dependent degradation of DNA. Nucleic acids research. 2000;28(17):3260-8.
16. Boiteux S, Jinks-Robertson S. DNA repair mechanisms and the bypass of DNA damage in Saccharomyces cerevisiae. Genetics. 2013;193(4):1025-64.
17. Boiteux S, Guillet M. Abasic sites in DNA: repair and biological consequences in Saccharomyces cerevisiae. DNA repair. 2004;3(1):1-12.
18. Curia MC, Catalano T, Aceto GM. MUTYH: Not just polyposis. World Journal of Clinical Oncology. 2020;11(7):428.
19. Vaughn C, Sancar A. Mechanisms and maps of nucleotide excision repair. 2020.
20. de Laat WL, Jaspers NG, Hoeijmakers JH. Molecular mechanism of nucleotide excision repair. Genes & development. 1999;13(7):768-85.

21. Fuss JO, Cooper PK. DNA repair: dynamic defenders against cancer and aging. PLoS biology. 2006;4(6):e203.
22. Smith ML, Ford JM, Hollander MC, Bortnick RA, Amundson SA, Seo YR, et al. p53-mediated DNA repair responses to UV radiation: studies of mouse cells lacking p53, p21, and/or gadd45 genes. Molecular and cellular biology. 2000;20(10):3705-14.
23. Kubo Y, Hoffmann B, Goltz K, Schnakenberg U, Jahr H, Merkel R, et al. Different frequency of cyclic tensile strain relates to anabolic/catabolic conditions consistent with immunohistochemical staining intensity in tenocytes. International journal of molecular sciences. 2020;21(3):1082.
24. Yasuda T, Kagawa W, Ogi T, Kato TA, Suzuki T, Dohmae N, et al. Novel function of HATs and HDACs in homologous recombination through acetylation of human RAD52 at double-strand break sites. PLoS genetics. 2018;14(3):e1007277.
25. Perkhofer L, Schmitt A, Romero Carrasco MC, Ihle M, Hampp S, Ruess DA, et al. ATM Deficiency Generating Genomic Instability Sensitizes Pancreatic Ductal Adenocarcinoma Cells to Therapy-Induced DNA DamageATM-Deficient PDAC and Targeted Therapies. Cancer research. 2017;77(20):5576-90.
26. Smith HL, Southgate H, Tweddle DA, Curtin NJ. DNA damage checkpoint kinases in cancer. Expert reviews in molecular medicine. 2020;22.
27. Cragun D, Weidner A, Tezak A, Clouse K, Pal T. Cancer risk management among female BRCA1/2, PALB2, CHEK2, and ATM carriers. Breast cancer research and treatment. 2020;182(2):421-8.
28. Pandey M, Raghavan SC. DNA double-strand break repair in mammals. Journal of Radiation and Cancer Research. 2017;8(2):93.
29. Mirzoeva OK, Petrini JH. DNA damage-dependent nuclear dynamics of the Mre11 complex. Molecular and cellular biology. 2001;21(1):281-8.
30. Chen C-C, Feng W, Lim PX, Kass EM, Jasin M. Homology-directed repair and the role of BRCA1, BRCA2, and related proteins in genome integrity and cancer. Annual review of cancer biology. 2018;2:313-36.
31. Pietrasik S, Zajac G, Morawiec J, Soszynski M, Fila M, Blasiak J. Interplay between BRCA1 and GADD45A and its potential for nucleotide excision repair in breast cancer pathogenesis. International journal of molecular sciences. 2020;21(3):870.
32. Bridge G, Rashid S, Martin SA. DNA mismatch repair and oxidative DNA damage: implications for cancer biology and treatment. Cancers. 2014;6(3):1597-614.
33. Cherkasova V. DNA mismatch repair in mammals. Genome Stability: Elsevier; 2021. p. 323-47.
34. Jensen NM, Dalsgaard T, Jakobsen M, Nielsen RR, Sørensen CB, Bolund L, et al. An update on targeted gene repair in mammalian cells: methods and mechanisms. Journal of biomedical science. 2011;18(1):1-14.
35. Heinen CD. Mismatch repair defects and Lynch syndrome: The role of the basic scientist in the battle against cancer. DNA repair. 2016;38:127-34.
36. Tiwari V, Wilson III DM. DNA damage and associated DNA repair defects in disease and premature aging. The American Journal of Human Genetics. 2019;105(2):237-57.
37. Peña-Diaz J, Rasmussen LJ. Approaches to diagnose DNA mismatch repair gene defects in cancer. DNA repair. 2016;38:147-54.

38. Shiovitz S, Korde LA. Genetics of breast cancer: a topic in evolution. Annals of Oncology. 2015;26(7):1291-9.
39. Germani A, Petrucci S, De Marchis L, Libi F, Savio C, Amanti C, et al. Beyond BRCA1 and BRCA2: deleterious variants in DNA repair pathway genes in Italian families with breast/ovarian and pancreatic cancers. Journal of clinical medicine. 2020;9(9):3003.
40. Kobayashi H, Ohno S, Sasaki Y, Matsuura M. Hereditary breast and ovarian cancer susceptibility genes. Oncology reports. 2013;30(3):1019-29.
41. Fackenthal JD, Olopade OI. Breast cancer risk associated with BRCA1 and BRCA2 in diverse populations. Nature Reviews Cancer. 2007;7(12):937-48.
42. Tung NM, Garber JE. BRCA1/2 testing: therapeutic implications for breast cancer management. British Journal of Cancer. 2018;119(2):141-52.
43. Li J, Zhao B, Huang T, Qin Z, Wang SM. Human BRCA pathogenic variants were originated during recent human history. Life science alliance. 2022;5(5).
44. Minde DP, Anvarian Z, Rüdiger SG, Maurice MM. Messing up disorder: how do missense mutations in the tumor suppressor protein APC lead to cancer? Molecular cancer. 2011;10(1):1-9.
45. Scheuermann JC, de Ayala Alonso AG, Oktaba K, Ly-Hartig N, McGinty RK, Fraterman S, et al. Histone H2A deubiquitinase activity of the Polycomb repressive complex PR-DUB. Nature. 2010;465(7295):243-7.
46. Banin S. Moyal L, Shieh S, Taya Y, Anderson CW, Chessa L, Smorodinsky NI, Prives C, Reiss Y, Shiloh Y, and Ziv Y. Enhanced phosphorylation of p53 by ATM in response to DNA damage Science. 1998;281:1674-7.
47. Canman CE, Lim D-S, Cimprich KA, Taya Y, Tamai K, Sakaguchi K, et al. Activation of the ATM kinase by ionizing radiation and phosphorylation of p53. Science. 1998;281(5383):1677-9.
48. Carranza D, Vega AK, Torres-Rusillo S, Montero E, Martinez LJ, Santamaría M, et al. Molecular and functional characterization of a cohort of Spanish patients with ataxia-telangiectasia. Neuromolecular medicine. 2017;19(1):161-74.
49. Stadler ZK, Thom P, Robson ME, Weitzel JN, Kauff ND, Hurley KE, et al. Genome-wide association studies of cancer. Journal of clinical oncology. 2010;28(27):4255.
50. Renault A-L, Mebirouk N, Fuhrmann L, Bataillon G, Cavaciuti E, Le Gal D, et al. Morphology and genomic hallmarks of breast tumours developed by ATM deleterious variant carriers. Breast Cancer Research. 2018;20(1):1-18.
51. Tung N, Lin NU, Kidd J, Allen BA, Singh N, Wenstrup RJ, et al. Frequency of germline mutations in 25 cancer susceptibility genes in a sequential series of patients with breast cancer. Journal of Clinical Oncology. 2016;34(13):1460.
52. Hsu H-M, Wang H-C, Chen S-T, Hsu G-C, Shen C-Y, Yu J-C. Breast cancer risk is associated with the genes encoding the DNA double-strand break repair Mre11/Rad50/Nbs1 complex. Cancer Epidemiology Biomarkers & Prevention. 2007;16(10):2024-32.
53. Woo J-S, Chung MS, Paik SS. Clinicopathological Significance of SMAD4 Expression in Breast Cancer. Journal of Breast Disease. 2019;7(2):52-8.
54. Cybulski C, Wokołorczyk D, Jakubowska A, Huzarski T, Byrski T, Gronwald J, et al. Risk of breast cancer in women with a CHEK2 mutation with and without a family history of breast cancer. Journal of Clinical Oncology. 2011;29(28):3747-52.
55. Guilford P, Humar B, Blair V. Hereditary diffuse gastric cancer: translation of CDH1 germline mutations into clinical practice. Gastric cancer. 2010;13(1):1-10.

56. Mueller J, Gazzoli I, Bandipalliam P, Garber JE, Syngal S, Kolodner RD. Comprehensive molecular analysis of mismatch repair gene defects in suspected Lynch syndrome (hereditary nonpolyposis colorectal cancer) cases. Cancer research. 2009;69(17):7053-61.
57. Claes K, Dahan K, Tejpar S, De Paepe A, Bonduelle M, Abramowicz M, et al. The genetics of familial adenomatous polyposis (FAP) and MutYH-associated polyposis (MAP). Acta gastro-enterologica Belgica. 2011;74(3):421-6.
58. Nielsen M, Morreau H, Vasen HF, Hes FJ. MUTYH-associated polyposis (MAP). Critical reviews in oncology/hematology. 2011;79(1):1-16.
59. Wu S, Zhou J, Zhang K, Chen H, Luo M, Lu Y, et al. Molecular mechanisms of PALB2 function and its role in breast cancer management. Front Oncol 10: 301. 2020.
60. Ambasta RK, Dave I, Kumar P. PTEN-Somatic mutations causing cancer: Cancer Genome Sequence Analysis. Journal of Pharmacy Research. 2011;4(10):3568-74.
61. Banerjee T, Brosh Jr RM. RECQL: a new breast cancer susceptibility gene. Cell Cycle. 2015;14(22):3540-3.
62. Wang Y-S, Chen J, Cui F, Wang H, Wang S, Hang W, et al. LKB1 is a DNA damage response protein that regulates cellular sensitivity to PARP inhibitors. Oncotarget. 2016;7(45):73389.
63. Silwal-Pandit L, Langerød A, Børresen-Dale A-L. TP53 mutations in breast and ovarian cancer. Cold Spring Harbor perspectives in medicine. 2017;7(1):a026252.
64. Evans DGR, Kallionpää RA, Clementi M, Trevisson E, Mautner V-F, Howell SJ, et al. Breast cancer in neurofibromatosis 1: survival and risk of contralateral breast cancer in a five country cohort study. Genetics in Medicine. 2020;22(2):398-406.
65. Janatova M, Soukupova J, Stribrna J, Kleiblova P, Vocka M, Boudova P, et al. Mutation analysis of the RAD51C and RAD51D genes in high-risk ovarian cancer patients and families from the Czech Republic. PLoS One. 2015;10(6):e0127711.
66. Wu Y, Zhang Y, Pi H, Sheng Y. Current therapeutic progress of CDK4/6 inhibitors in breast cancer. Cancer Management and Research. 2020;12:3477.
67. Aftab A, Shahzad S, Hussain HMJ, Khan R, Irum S, Tabassum S. CDKN2A/P16INK4A variants association with breast cancer and their in-silico analysis. Breast Cancer. 2019;26(1):11-28.
68. Chen WY, Garber JE, Higham S, Schneider KA, Davis KB, Deffenbaugh AM, et al. BRCA1/2 genetic testing in the community setting. Journal of Clinical Oncology. 2002;20(22):4485-92.
69. Park J, Huang D, Chang YJ, Lim MC, Myung S-K. Oral contraceptives and risk of breast cancer and ovarian cancer in women with a BRCA1 or BRCA2 mutation: a meta-analysis of observational studies. Carcinogenesis. 2022;43(3):231-42.
70. Roila F, Fumi G, Ruggeri B, Antonuzzo A, Ripamonti C, Fatigoni S, et al. Prevalence, characteristics, and treatment of fatigue in oncological cancer patients in Italy: a cross-sectional study of the Italian Network for Supportive Care in Cancer (NICSO). Supportive Care in Cancer. 2019;27(3):1041-7.
71. Incorvaia L, Fanale D, Badalamenti G, Bono M, Calò V, Cancelliere D, et al. Hereditary breast and ovarian cancer in families from Southern Italy (Sicily)—prevalence and geographic distribution of pathogenic variants in BRCA1/2 genes. Cancers. 2020;12(5):1158.

72. Melchor L, Benítez J. The complex genetic landscape of familial breast cancer. Human genetics. 2013;132(8):845-63.
73. Lichtenstein P, Holm NV, Verkasalo PK, Iliadou A, Kaprio J, Koskenvuo M, et al. Environmental and heritable factors in the causation of cancer—analyses of cohorts of twins from Sweden, Denmark, and Finland. New England journal of medicine. 2000;343(2):78-85.
74. Powell SM, Petersen GM, Krush AJ, Booker S, Jen J, Giardiello FM, et al. Molecular diagnosis of familial adenomatous polyposis. New England Journal of Medicine. 1993;329(27):1982-7.
75. Jansen AM, Crobach S, Geurts-Giele WR, van den Akker BE, Garcia MV, Ruano D, et al. Distinct patterns of somatic mosaicism in the APC gene in neoplasms from patients with unexplained adenomatous polyposis. Gastroenterology. 2017;152(3):546-9. e3.
76. Pérez-Cabornero L, Sanz MI, Sampedro EV, Aras EL, Becares AA, Pino CM, et al. Frequency of Rearrangements in Lynch Syndrome Cases Associated with MSH2: Characterization of a New Deletion Involving both EPCAM and the 5′ Part of MSH2Deletions in MSH2-Associated Lynch Syndrome Cases. Cancer Prevention Research. 2011;4(10):1556-62.
77. South CD, Hampel H, Comeras I, Westman JA, Frankel WL, de la Chapelle A. The frequency of Muir-Torre syndrome among Lynch syndrome families. Journal of the national cancer institute. 2008;100(4):277-81.
78. Boland PM, Yurgelun MB, Boland CR. Recent progress in Lynch syndrome and other familial colorectal cancer syndromes. CA: a cancer journal for clinicians. 2018;68(3):217-31.
79. Vasen H, Mecklin J-P, Khan PM, Lynch H. The international collaborative group on hereditary non-polyposis colorectal cancer (ICG-HNPCC). Diseases of the Colon & Rectum. 1991;34(5):424-5.
80. Aaltonen LA, Peltomäki P, Leach FS, Sistonen P, Pylkkänen L, Mecklin J-P, et al. Clues to the pathogenesis of familial colorectal cancer. Science. 1993;260(5109):812-6.
81. Vasen HF, Watson P, Mecklin JP, Lynch HT. New clinical criteria for hereditary nonpolyposis colorectal cancer (HNPCC, Lynch syndrome) proposed by the International Collaborative group on HNPCC. Gastroenterology. 1999;116(6):1453-6.
82. Umar A, Boland CR, Terdiman JP, Syngal S, Chapelle Adl, Rüschoff J, et al. Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. Journal of the National Cancer Institute. 2004;96(4):261-8.
83. Cohen SA, Pritchard CC, Jarvik GP. Lynch syndrome: from screening to diagnosis to treatment in the era of modern molecular oncology. Annual review of genomics and human genetics. 2019;20:293-307.
84. Ponti G, Castellsagué E, Ruini C, Percesepe A, Tomasi A. Mismatch repair genes founder mutations and cancer susceptibility in Lynch syndrome. Clinical genetics. 2015;87(6):507-16.
85. Raskin L, Schwenter F, Freytsis M, Tischkowitz M, Wong N, Chong G, et al. Characterization of two Ashkenazi Jewish founder mutations in MSH6 gene causing Lynch syndrome. Clinical genetics. 2011;79(6):512-22.
86. Gebert J, Gelincik O, Oezcan-Wahlbrink M, Marshall JD, Hernandez-Sanchez A, Urban K, et al. Recurrent frameshift neoantigen vaccine elicits protective

immunity with reduced tumor burden and improved overall survival in a Lynch syndrome mouse model. Gastroenterology. 2021;161(4):1288-302. e13.
87. Chang L, Chang M, Chang HM, Chang F. Microsatellite instability: a predictive biomarker for cancer immunotherapy. Applied Immunohistochemistry & Molecular Morphology. 2018;26(2):e15-e21.
88. Duraturo F, Liccardo R, De Rosa M, Izzo P. Genetics, diagnosis and treatment of Lynch syndrome: Old lessons and current challenges. Oncology letters. 2019;17(3):3048-54.
89. Rasmussen LJ, Heinen CD, Royer-Pokora B, Drost M, Tavtigian S, Hofstra RM, et al. Pathological assessment of mismatch repair gene variants in Lynch syndrome: past, present, and future. Human mutation. 2012;33(12):1617-25.
90. Thompson BA, Spurdle AB, Plazzer J-P, Greenblatt MS, Akagi K, Al-Mulla F, et al. Application of a 5-tiered scheme for standardized classification of 2,360 unique mismatch repair gene variants in the InSiGHT locus-specific database. Nature genetics. 2014;46(2):107-15.
91. Manahan ER, Kuerer HM, Sebastian M, Hughes KS, Boughey JC, Euhus DM, et al. Consensus guidelines on genetictesting for hereditary breast cancer from the American Society of Breast Surgeons. Annals of surgical oncology. 2019;26(10):3025-31.
92. Kohn AB, Moroz TP, Barnes JP, Netherton M, Moroz LL. Single-cell semiconductor sequencing. Biological Aging: Springer; 2013. p. 247-84.
93. Roy S, Coldren C, Karunamurthy A, Kip NS, Klee EW, Lincoln SE, et al. Standards and guidelines for validating next-generation sequencing bioinformatics pipelines: a joint recommendation of the Association for Molecular Pathology and the College of American Pathologists. The Journal of Molecular Diagnostics. 2018;20(1):4-27.
94. Austin-Tse CA, Jobanputra V, Perry DL, Bick D, Taft RJ, Venner E, et al. Best practices for the interpretation and reporting of clinical whole genome sequencing. NPJ genomic medicine. 2022;7(1):1-13.
95. Barnell EK, Ronning P, Campbell KM, Krysiak K, Ainscough BJ, Ramirez C, et al. Standard operating procedure for somatic variant refinement of tumor sequencing data. bioRxiv. 2018:266262.
96. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genetics in medicine. 2015;17(5):405-23.
97. Thusberg J, Olatubosun A, Vihinen M. Performance of mutation pathogenicity prediction methods on missense variants. Human mutation. 2011;32(4):358-68.
98. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. Nature methods. 2010;7(4):248-9.
99. Orlando L, Schiavone P, Fedele P, Calvani N, Nacci A, Rizzo P, et al. Molecularly targeted endocrine therapies for breast cancer. Cancer treatment reviews. 2010;36:S67-S71.
100. Pernas S, Tolaney SM. HER2-positive breast cancer: new therapeutic frontiers and overcoming resistance. Therapeutic advances in medical oncology. 2019;11:1758835919833519.

101. Mabuchi S, Sugiyama T, Kimura T. Clear cell carcinoma of the ovary: molecular insights and future therapeutic perspectives. Journal of gynecologic oncology. 2016;27(3).
102. Farmer H, McCabe N, Lord CJ, Tutt AN, Johnson DA, Richardson TB, et al. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. Nature. 2005;434(7035):917-21.
103. Engbrecht M, Mangerich A. The nucleolus and PARP1 in cancer biology. Cancers. 2020;12(7):1813.
104. Murai J, Huang S-yN, Das BB, Renaud A, Zhang Y, Doroshow JH, et al. Trapping of PARP1 and PARP2 by clinical PARP inhibitors. Cancer research. 2012;72(21):5588-99.
105. Min A, Im S-A. PARP inhibitors as therapeutics: beyond modulation of PARylation. Cancers. 2020;12(2):394.
106. Lord CJ, Ashworth A. BRCAness revisited. Nature Reviews Cancer. 2016;16(2):110-20.
107. Network CGAR. Integrated genomic analyses of ovarian carcinoma. Nature. 2011;474(7353):609.
108. Mateo J, Carreira S, Sandhu S, Miranda S, Mossop H, Perez-Lopez R, et al. DNA-repair defects and olaparib in metastatic prostate cancer. New England Journal of Medicine. 2015;373(18):1697-708.
109. Waddell N, Pajic M, Patch A-M, Chang DK, Kassahn KS, Bailey P, et al. Whole genomes redefine the mutational landscape of pancreatic cancer. Nature. 2015;518(7540):495-501.
110. Pommier Y, O'Connor MJ, De Bono J. Laying a trap to kill cancer cells: PARP inhibitors and their mechanisms of action. Science translational medicine. 2016;8(362):362ps17-ps17.
111. Fong PC, Yap TA, Boss DS, Carden CP, Mergui-Roelvink M, Gourley C, et al. Poly (ADP)-ribose polymerase inhibition: frequent durable responses in BRCA carrier ovarian cancer correlating with platinum-free interval. Journal of clinical oncology. 2010;28(15):2512-9.
112. Pujade-Lauraine E, Ledermann JA, Selle F, Gebski V, Penson RT, Oza AM, et al. Olaparib tablets as maintenance therapy in patients with platinum-sensitive, relapsed ovarian cancer and a BRCA1/2 mutation (SOLO2/ENGOT-Ov21): a double-blind, randomised, placebo-controlled, phase 3 trial. The lancet oncology. 2017;18(9):1274-84.
113. Shih HA, Couch FJ, Nathanson KL, Blackwood MA, Rebbeck TR, Armstrong KA, et al. BRCA1 and BRCA2 mutation frequency in women evaluated in a breast cancer risk evaluation clinic. Journal of Clinical Oncology. 2002;20(4):994-9.
114. Shiloh Y. ATM and related protein kinases: safeguarding genome integrity. Nature reviews cancer. 2003;3(3):155-68.
115. Cantor SB, Guillemette S. Hereditary breast cancer and the BRCA1-associated FANCJ/BACH1/BRIP1. Future oncology. 2011;7(2):253-61.
116. Tung N, Silver DP. Chek2 DNA damage response pathway and inherited breast cancer risk. Journal of clinical oncology. 2011;29(28):3813-5.
117. Park J-Y, Singh TR, Nassar N, Zhang F, Freund M, Hanenberg H, et al. Breast cancer-associated missense mutants of the PALB2 WD40 domain, which directly binds RAD51C, RAD51 and BRCA2, disrupt DNA repair. Oncogene. 2014;33(40):4803-12.

118. Santana dos Santos E, Lallemand F, Petitalot A, Caputo SM, Rouleau E. HRness in breast and ovarian cancers. International Journal of Molecular Sciences. 2020;21(11):3850.
119. Yu Q, Sicinska E, Geng Y, Ahnström M, Zagozdzon A, Kong Y, et al. Requirement for CDK4 kinase function in breast cancer. Cancer cell. 2006;9(1):23-32.
120. Singla RK, Wang X, Gundamaraju R, Joon S, Tsagkaris C, Behzad S, et al. Natural products derived from medicinal plants and microbes might act as a game-changer in breast cancer: a comprehensive review of preclinical and clinical studies. Critical Reviews in Food Science and Nutrition. 2022:1-45.
121. Sarrió D, Moreno-Bueno G, Hardisson D, Sánchez-Estévez C, Guo M, Herman JG, et al. Epigenetic and genetic alterations of APC and CDH1 genes in lobular breast cancer: relationships with abnormal E-cadherin and catenin expression and microsatellite instability. International journal of cancer. 2003;106(2):208-15.
122. Vierkoetter KR, Ayabe AR, VanDrunen M, Ahn HJ, Shimizu DM, Terada KY. Lynch Syndrome in patients with clear cell and endometrioid cancers of the ovary. Gynecologic oncology. 2014;135(1):81-4.
123. Rizzolo P, Silvestri V, Bucalo A, Zelli V, Valentini V, Catucci I, et al. Contribution of MUTYH variants to male breast cancer risk: results from a multicenter study in Italy. Frontiers in Oncology. 2018;8:583.
124. Matsumoto K, Nishimura M, Onoe T, Sakai H, Urakawa Y, Onda T, et al. PARP inhibitors for BRCA wild type ovarian cancer; gene alterations, homologous recombination deficiency and combination therapy. Japanese journal of clinical oncology. 2019.
125. Becker-Catania SG, Chen G, Hwang MJ, Wang Z, Sun X, Sanal O, et al. Ataxia-telangiectasia: phenotype/genotype studies of ATM protein expression, mutations, and radiosensitivity. Molecular genetics and metabolism. 2000;70(2):122-33.
126. Stankovic T, Kidd A, Sutcliffe A, McGuire G, Robinson P, Weber P, et al. ATM mutations and phenotypes in ataxia-telangiectasia families in the British Isles: expression of mutant ATM and the risk of leukemia, lymphoma, and breast cancer. The American Journal of Human Genetics. 1998;62(2):334-45.
127. Gilad S, Chessa L, Khosravi R, Russell P, Galanty Y, Piane M, et al. Genotype-phenotype relationships in ataxia-telangiectasia and variants. The American Journal of Human Genetics. 1998;62(3):551-61.
128. Cavalieri S, Funaro A, Porcedda P, Turinetto V, Migone N, Gatti RA, et al. ATM mutations in Italian families with ataxia telangiectasia include two distinct large genomic deletions. Human Mutation. 2006;27(10):1061-.
129. Chessa L, Piane M, Magliozzi M, Torrente I, Savio C, Lulli P, et al. Founder effects for ATM gene mutations in Italian Ataxia Telangiectasia families. Annals of human genetics. 2009;73(5):532-9.
130. Huang Y, Yang L, Wang J, Yang F, Xiao Y, Xia R, et al. Twelve novel Atm mutations identified in Chinese ataxia telangiectasia patients. Neuromolecular Medicine. 2013;15(3):536-40.
131. Meindl A, Hellebrand H, Wiek C, Erven V, Wappenschmidt B, Niederacher D, et al. Germline mutations in breast and ovarian cancer pedigrees establish RAD51C as a human cancer susceptibility gene. Nature genetics. 2010;42(5):410-4.
132. Kraus C, Hoyer J, Vasileiou G, Wunderle M, Lux MP, Fasching PA, et al. Gene panel sequencing in familial breast/ovarian cancer patients identifies multiple

novel mutations also in genes others than BRCA1/2. International journal of cancer. 2017;140(1):95-102.
133. Pelttari LM, Heikkinen T, Thompson D, Kallioniemi A, Schleutker J, Holli K, et al. RAD51C is a susceptibility gene for ovarian cancer. Human molecular genetics. 2011;20(16):3278-88.
134. Schild D, Lio Y-c, Collins DW, Tsomondo T, Chen DJ. Evidence for simultaneous protein interactions between human Rad51 paralogs. Journal of Biological Chemistry. 2000;275(22):16443-9.
135. Thompson BA, Goldgar DE, Paterson C, Clendenning M, Walters R, Arnold S, et al. A multifactorial likelihood model for MMR gene variant classification incorporating probabilities based on sequence bioinformatics and tumor characteristics: a report from the Colon Cancer Family Registry. Human mutation. 2013;34(1):200-9.
136. Sjursen W, McPhillips M, Scott RJ, Talseth-Palmer BA. Lynch syndrome mutation spectrum in New South Wales, Australia, including 55 novel mutations. Molecular genetics & genomic medicine. 2016;4(2):223-31.
137. Latham A, Srinivasan P, Kemel Y, Shia J, Bandlamudi C, Mandelker D, et al. Microsatellite instability is associated with the presence of Lynch syndrome pan-cancer. Journal of clinical oncology. 2019;37(4):286.
138. Rossi BM, Palmero EI, López-Kostner F, Sarroca C, Vaccaro CA, Spirandelli F, et al. A survey of the clinicopathological and molecular characteristics of patients with suspected Lynch syndrome in Latin America. Bmc Cancer. 2017;17(1):1-26.
139. Berginc G, Bračko M, Ravnik-Glavač M, Glavač D. Screening for germline mutations of MLH1, MSH2, MSH6 and PMS2 genes in Slovenian colorectal cancer patients: implications for a population specific detection strategy of Lynch syndrome. Familial cancer. 2009;8(4):421-9.
140. Walsh T, Casadei S, Lee MK, Pennil CC, Nord AS, Thornton AM, et al. Mutations in 12 genes for inherited ovarian, fallopian tube, and peritoneal carcinoma identified by massively parallel sequencing. Proceedings of the National Academy of Sciences. 2011;108(44):18032-7.
141. DeRycke MS, Gunawardena S, Balcom JR, Pickart AM, Waltman LA, French AJ, et al. Targeted sequencing of 36 known or putative colorectal cancer susceptibility genes. Molecular genetics & genomic medicine. 2017;5(5):553-69.
142. Ali M, Kim H, Cleary S, Cupples C, Gallinger S, Bristow R. Characterization of mutant MUTYH proteins associated with familial colorectal cancer. Gastroenterology. 2008;135(2):499-507. e1.
143. Eliason K, Hendrickson B, Judkins T, Norton M, Leclair B, Lyon E, et al. The potential for increased clinical sensitivity in genetic testing for polyposis colorectal cancer through the analysis of MYH mutations in North American patients. Journal of medical genetics. 2005;42(1):95-6.
144. Avezzu A, Agostini M, Pucciarelli S, Lise M, Urso ED, Mammi I, et al. The role of MYH gene in genetic predisposition to colorectal cancer: another piece of the puzzle. Cancer letters. 2008;268(2):308-13.
145. Baldock RA, Pressimone CA, Baird JM, Khodakov A, Luong TT, Grundy MK, et al. RAD51D splice variants and cancer-associated mutations reveal XRCC2 interaction to be critical for homologous recombination. DNA repair. 2019;76:99-107.
146. Molatore S, Russo MT, D'Agostino VG, Barone F, Matsumoto Y, Albertini AM, et al. MUTYH mutations associated with familial adenomatous polyposis:

functional characterization by a mammalian cell-based assay. Human mutation. 2010;31(2):159-66.
147. Nykamp K, Anderson M, Powers M, Garcia J, Herrera B, Ho Y-Y, et al. Sherloc: a comprehensive refinement of the ACMG–AMP variant classification criteria. Genetics in Medicine. 2017;19(10):1105-17.
148. Ruijs MW, Verhoef S, Rookus MA, Pruntel R, van der Hout AH, Hogervorst FB, et al. TP53 germline mutation testing in 180 families suspected of Li–Fraumeni syndrome: mutation detection rate and relative frequency of cancers in different familial phenotypes. Journal of medical genetics. 2010;47(6):421-8.
149. Girard E, Eon-Marchais S, Olaso R, Renault Al, Damiola F, Dondon Mg, et al. Familial breast cancer and DNA repair genes: Insights into known and novel susceptibility genes from the GENESIS study, and implications for multigene panel testing. International journal of cancer. 2019;144(8):1962-74.
150. Harland M, Taylor CF, Chambers PA, Kukalizch K, Randerson-Moor JA, Gruis NA, et al. A mutation hotspot at the p14ARF splice site. Oncogene. 2005;24(28):4604-8.
151. Ratajska M, Antoszewska E, Piskorz A, Brozek I, Borg Å, Kusmierek H, et al. Cancer predisposing BARD1 mutations in breast–ovarian cancer families. Breast cancer research and treatment. 2012;131(1):89-97.
152. De Brakeleer S, De Grève J, Loris R, Janin N, Lissens W, Sermijn E, et al. Cancer predisposing missense and protein truncating BARD1 mutations in non-BRCA1 or BRCA2 breast cancer families. Human mutation. 2010;31(3):E1175-E85.
153. Slavin T, Neuhausen SL, Rybak C, Solomon I, Nehoray B, Blazer K, et al. Genetic gastric cancer susceptibility in the international clinical cancer genomics community research network. Cancer genetics. 2017;216:111-9.
154. Kerr SE, Thomas CB, Thibodeau SN, Ferber MJ, Halling KC. APC germline mutations in individuals being evaluated for familial adenomatous polyposis: a review of the Mayo Clinic experience with 1591 consecutive tests. The Journal of Molecular Diagnostics. 2013;15(1):31-43.
155. Wen Y, Eng CH, Schmoranzer J, Cabrera-Poch N, Morris EJ, Chen M, et al. EB1 and APC bind to mDia to stabilize microtubules downstream of Rho and promote cell migration. Nature cell biology. 2004;6(9):820-30.
156. Maxwell KN, Wubbenhorst B, D'Andrea K, Garman B, Long JM, Powers J, et al. Prevalence of mutations in a panel of breast cancer susceptibility genes in BRCA1/2-negative patients with early-onset breast cancer. Genetics in medicine. 2015;17(8):630-8.
157. Falck J, Mailand N, Syljuåsen RG, Bartek J, Lukas J. The ATM–Chk2–Cdc25A checkpoint pathway guards against radioresistant DNA synthesis. Nature. 2001;410(6830):842-7.
158. Agiannitopoulos K, Papadopoulou E, Tsaousis GN, Pepe G, Kampouri S, Kocdor MA, et al. Characterization of the c. 793-1G> A splicing variant in CHEK2 gene as pathogenic: a case report. BMC Medical Genetics. 2019;20(1):1-4.
159. Leedom TP, LaDuca H, McFarland R, Li S, Dolinsky JS, Chao EC. Breast cancer risk is similar for CHEK2 founder and non-founder mutation carriers. Cancer Genetics. 2016;209(9):403-7.
160. Roeb W, Higgins J, King M-C. Response to DNA damage of CHEK2 missense mutations in familial breast cancer. Human molecular genetics. 2012;21(12):2738-44.

161. Hale V, Weischer M, Park JY. CHEK2∗ 1100delC mutation and risk of prostate cancer. Prostate Cancer. 2014;2014.
162. Angeli D, Salvi S, Tedaldi G. Genetic predisposition to breast and ovarian cancers: how many and which genes to test? International journal of molecular sciences. 2020;21(3):1128.
163. Tung N, Battelli C, Allen B, Kaldate R, Bhatnagar S, Bowles K, et al. Frequency of mutations in individuals with breast cancer referred for BRCA 1 and BRCA 2 testing using next-generation sequencing with a 25-gene panel. Cancer. 2015;121(1):25-33.
164. Siołek M, Cybulski C, Gąsior-Perczak D, Kowalik A, Kozak-Klonowska B, Kowalska A, et al. CHEK 2 mutations and the risk of papillary thyroid cancer. International Journal of Cancer. 2015;137(3):548-52.
165. Woodward ER, van Veen EM, Forde C, Harkness EF, Byers HJ, Ellingford JM, et al. Clinical utility of testing for PALB2 and CHEK2 c. 1100delC in breast and ovarian cancer. Genetics in Medicine. 2021;23(10):1969-76.
166. Weischer M, Heerfordt IM, Bojesen SE, Eigentler T, Garbe C, Röcken M, et al. CHEK2* 1100delC and risk of malignant melanoma: Danish and German studies and meta-analysis. Journal of investigative dermatology. 2012;132(2):299-303.
167. Rizzolo P, Zelli V, Silvestri V, Valentini V, Zanna I, Bianchi S, et al. Insight into genetic susceptibility to male breast cancer by multigene panel testing: Results from a multicenter study in Italy. International Journal of Cancer. 2019;145(2):390-400.
168. Tedaldi G, Danesi R, Zampiga V, Tebaldi M, Bedei L, Zoli W, et al. First evidence of a large CHEK2 duplication involved in cancer predisposition in an Italian family with hereditary breast cancer. BMC cancer. 2014;14(1):1-5.
169. Thorstenson YR, Roxas A, Kroiss R, Jenkins MA, Yu KM, Bachrich T, et al. Contributions of ATM mutations to familial breast and ovarian cancer. Cancer research. 2003;63(12):3325-33.
170. Hansford S, Kaurah P, Li-Chang H, Woo M, Senz J, Pinheiro H, et al. Hereditary diffuse gastric cancer syndrome: CDH1 mutations and beyond. JAMA oncology. 2015;1(1):23-32.
171. Huang D-S, Tao H-Q, He X-J, Long M, Yu S, Xia Y-J, et al. Prevalence of deleterious ATM germline mutations in gastric cancer patients. Oncotarget. 2015;6(38):40953.
172. Zhou J, Zhao Z, Zhang Y, Bao C, Cui L, Cai S, et al. Pathogenic germline mutations in chinese patients with gastric cancer identified by next-generation sequencing. Oncology. 2020;98(8):583-8.
173. Helgason H, Rafnar T, Olafsdottir HS, Jonasson JG, Sigurdsson A, Stacey SN, et al. Loss-of-function variants in ATM confer risk of gastric cancer. Nature genetics. 2015;47(8):906-10.
174. Arvai KJ, Roberts ME, Torene RI, Susswein LR, Marshall ML, Zhang Z, et al. Age-adjusted association of homologous recombination genes with ovarian cancer using clinical exomes as controls. Hereditary cancer in clinical practice. 2019;17(1):1-7.
175. Akbari MR, Tonin P, Foulkes WD, Ghadirian P, Tischkowitz M, Narod SA. RAD51C germline mutations in breast and ovarian cancer patients. Breast Cancer Research. 2010;12(4):1-2.
176. Hoyer J, Vasileiou G, Uebe S, Wunderle M, Kraus C, Fasching PA, et al. Addition of triple negativity of breast cancer as an indicator for germline mutations

in predisposing genes increases sensitivity of clinical selection criteria. BMC cancer. 2018;18(1):1-11.
177. Couch FJ, Hart SN, Sharma P, Toland AE, Wang X, Miron P, et al. Inherited mutations in 17 breast cancer susceptibility genes among a large triple-negative breast cancer cohort unselected for family history of breast cancer. Journal of clinical oncology. 2015;33(4):304.
178. Śniadecki M, Brzeziński M, Darecka K, Klasa-Mazurkiewicz D, Poniewierza P, Krzeszowiec M, et al. BARD1 and breast cancer: the possibility of creating screening tests and new preventive and therapeutic pathways for predisposed women. Genes. 2020;11(11):1251.
179. Uzunoglu H, Korak T, Ergul E, Uren N, Sazci A, Utkan NZ, et al. Association of the nibrin gene (NBN) variants with breast cancer. Biomedical Reports. 2016;4(3):369-73.
180. Chan SH, Chiang J, Ngeow J. CDKN2A germline alterations and the relevance of genotype-phenotype associations in cancer predisposition. Hereditary cancer in clinical practice. 2021;19(1):1-9.
181. Dębniak T, Gorski B, Huzarski T, Byrski T, Cybulski C, Mackiewicz A, et al. A common variant of CDKN2A (p16) predisposes to breast cancer. Journal of medical genetics. 2005;42(10):763-5.
182. Ungerleider NA, Rao SG, Shahbandi A, Yee D, Niu T, Frey WD, et al. Breast cancer survival predicted by TP53 mutation status differs markedly depending on treatment. Breast Cancer Research. 2018;20(1):1-8.
183. Adib E, El Zarif T, Nassar AH, Akl EW, Abou Alaiwi S, Mouhieddine TH, et al. CDH1 germline variants are enriched in patients with colorectal cancer, gastric cancer, and breast cancer. British journal of cancer. 2022;126(5):797-803.
184. Huang J, Zheng S, Jin S-H, Zhang S-Z. Somatic mutations of APC gene in carcinomas from hereditary non-polyposis colorectal cancer patients. World Journal of Gastroenterology. 2004;10(6):834.
185. Li Y, Fan L, Zheng J, Nie X, Sun Y, Feng Q, et al. Lynch syndrome pre-screening and comprehensive characterization in a multi-center large cohort of Chinese patients with colorectal cancer. Cancer biology & medicine. 2022:j. issn. 2095-3941.2021. 0585.
186. Gray PN, Tsai P, Chen D, Wu S, Hoo J, Mu W, et al. TumorNext-Lynch-MMR: a comprehensive next generation sequencing assay for the detection of germline and somatic mutations in genes associated with mismatch repair deficiency and Lynch syndrome. Oncotarget. 2018;9(29):20304.
187. Kashiwaba M, Tamura G, Ishida M. Aberrations of theAPC gene in primary breast carcinoma. Journal of cancer research and clinical oncology. 1994;120(12):727-31.
188. Sørlie T, Bukholm I, Børresen-Dale A. Truncating somatic mutation in exon 15 of the APC gene is a rare event in human breast carcinomas. Mutations in brief no. 179. Online. Human mutation. 1998;12(3):215-.
189. Furuuchi K, Tada M, Yamada H, Kataoka A, Furuuchi N, Hamada J-i, et al. Somatic mutations of the APC gene in primary breast cancers. The American journal of pathology. 2000;156(6):1997-2005.
190. Nguyen-Dumont T, Steen JA, Winship I, Park DJ, Pope BJ, Hammet F, et al. Mismatch repair gene pathogenic germline variants in a population-based cohort of breast cancer. Familial Cancer. 2020;19(3):197-202.

191. Baglietto L, Lindor NM, Dowty JG, White DM, Wagner A, Gomez Garcia EB, et al. Risks of Lynch syndrome cancers for MSH6 mutation carriers. Journal of the national cancer institute. 2010;102(3):193-201.
192. Borràs E, Pineda M, Cadiñanos J, Del Valle J, Brieger A, Hinrichsen I, et al. Refining the role of PMS2 in Lynch syndrome: germline mutational analysis improved by comprehensive assessment of variants. Journal of medical genetics. 2013;50(8):552-63.
193. Win AK, Cleary SP, Dowty JG, Baron JA, Young JP, Buchanan DD, et al. Cancer risks for monoallelic MUTYH mutation carriers with a family history of colorectal cancer. International journal of cancer. 2011;129(9):2256-62.
194. Gostner JM, Fong D, Wrulich OA, Lehne F, Zitt M, Hermann M, et al. Effects of EpCAM overexpression on human breast cancer cell lines. BMC cancer. 2011;11(1):1-14.
195. Osta WA, Chen Y, Mikhitarian K, Mitas M, Salem M, Hannun YA, et al. EpCAM is overexpressed in breast cancer and is a potential target for breast cancer gene therapy. Cancer research. 2004;64(16):5818-24.
196. Forones NM, Lima FT, Martin RP, Martins L, Teixeira PVL, Pesquero JB, et al. Mismatch Repair Genes and EPCAM germline mutations in patients with gastric or colorectal cancer with suspected of Lynch syndrome. American Society of Clinical Oncology; 2018.
197. Pathak SJ, Mueller JL, Okamoto K, Das B, Hertecant J, Greenhalgh L, et al. EPCAM mutation update: Variants associated with congenital tufting enteropathy and Lynch syndrome. Human mutation. 2019;40(2):142-61.
198. Mullai N. The importance of variants of unknown significance (VUS) in BRCA mutation. Wolters Kluwer Health; 2021.
199. Li X, Liu G, Wu W. Recent advances in Lynch syndrome. Experimental Hematology & Oncology. 2021;10(1):1-8.
200. Idos G, Valle L. Lynch syndrome. 2021.
201. Haraldsdottir S, Rafnar T, Frankel WL, Einarsdottir S, Sigurdsson A, Hampel H, et al. Comprehensive population-wide analysis of Lynch syndrome in Iceland reveals founder mutations in MSH6 and PMS2. Nature communications. 2017;8(1):1-11.