UNIVERSITA' DEGLI STUDI DI PADOVA

DIPARTIMENTO DI SCIENZE ONCOLOGICHE E CHIRURGICHE
SEZIONE DI ONCOLOGIA

SCUOLA DI DOTTORATO DI RICERCA IN ONCOLOGIA
E ONCOLOGIA CHIRURGICA
XXII ciclo

STRATEGIES FOR THE IDENTIFICATION OF ALLELES INVOLVED IN HEREDITARY BREAST CANCER

DIRETTORE: CH.MA PROF. PAOLA ZANOVELLO
SUPERVISORE: DR. MARCO MONTAGNA
CORRELATORE: CH.MA PROF. EMMA D'ANDREA

DOTTORANDO: DR. CINZIA CASELLA

31 DICEMBRE 2009

Alla mia mamma

Index

Riassunto

Il tumore della mammella e' la neoplasia più frequente nelle donne dell'Occidente. Ad oggi sono stati identificati alcuni fattori di rischio per lo sviluppo della neoplasia mammaria, ma il maggiormente significativo è la presenza di storia familiare, che è associata alla presenza nei membri della famiglia di alterazioni germinali predisponenti. Fino ad oggi sono stati identificati alcuni geni responsabili dell'aumento di rischio per lo sviluppo del tumore della mammella, tra cui i più rilevanti sono i geni ad alta penetranza BRCA1 e BRCA2. Tuttavia più del 70% dell'eccesso di rischio che si riscontra in casi familiari di tumore della mammella rispetto la popolazione generale non trova ad oggi un riscontro in alterazioni genetiche predisponenti.

Durante il mio dottorato mi sono occupata dell'identificazione di alterazioni genetiche coinvolte nella predisposizione e progressione del tumore eredo-familiare della mammella in cui sono state escluse mutazioni chiaramente patogenetiche dei geni BRCA1/BRCA2.

In un primo progetto ho partecipato alla caratterizzazione del ruolo patogenetico della variante BRCA1 p.Val1688del, identificata come variante di incerto significato clinico (*Unclassified Variant*, UV) e riscontrata frequentemente in pazienti provenienti dalla regione Veneto. In assenza di un chiaro effetto deleterio sulla funzionalità della proteina, le UV non possono essere utilizzate per l'identificazione e la sorveglianza degli individui ad alto rischio per lo sviluppo della neoplasia mammaria. Allo scopo di chiarire il ruolo patogenetico della delezione p.Val1688del, abbiamo utilizzato un approccio multi fattoriale, descritto da Golgar *et al*. (2004), attraverso il quale evidenze di diversa natura vengono integrate per derivare un rapporto di probabilità in favore della patogenicità o neutralità della variante. A questo scopo, i membri di 12 famiglie portatrici della variante sono stati analizzati per ottenere dati quali la co-segregazione della variante con il fenotipo tumorale, la co-presenza di mutazioni patogenetiche nel gene BRCA1, l'istopatologia del tumore, la perdita di eterozigosi (LOH) e la conservazione filogenetica del residuo aminoacidico alterato. Assunta l'indipendenza delle evidenze raccolte, i rapporti di probabilità associati a ciascuna evidenza sperimentale e clinica sono stati combinati,

ottenendo un valore di 349000:1 in favore del ruolo patogenetico della variante, che supera di molto il *cut off* di 1000:1 stabilito per poter classificare la variante come alterazione predisponente.
Il secondo progetto di cui mi sono occupata riguarda la definizione del profilo genetico della linea cellulare di carcinoma mammario HCC1500, derivata da una paziente con probabile tumore eredo-familiare della mammella. Nonostante l'età precoce di insorgenza e la storia familiare di tumore della mammella e del colon, il coinvolgimento dei geni BRCA1, BRCA2 e TP53 è stato escluso mediante *screening* mutazionale. La linea HCC1500 è stata quindi analizzata mediante due approcci complementari: i) l'identificazione di mutazioni nonsenso mediante l'utilizzo di un'approccio di analisi recentemente descritto chiamato GINI (*Gene Identification by NMD Inhibition*; Noesie and Dietz, 2001), e ii) l'identificazione di alterazioni numeriche mediante l'utilizzo di *array* per la genotipizzazione di un elevato numero di SNP, una recente tecnologia che permette un'elevata risoluzione di analisi. Mediante questo approccio è stato possibile identificare singoli geni mutati con specifiche alterazioni puntiformi o a causa di più estesi riarrangiamenti genomici. In particolare, nel gene PNLIPRP3 abbiamo identificato una sostituzione intronica che, modificando lo splicing del pre-mRNA, potrebbe dare origine ad un trascritto con mutazione troncante. Inoltre, l'elevata risoluzione dell'analisi per lo studio delle alterazioni cromosomiche numeriche, ci ha permesso di identificare singoli o pochi geni coinvolti in delezioni in omozigosi, amplificazioni di piccole regioni e riarrangiamenti intragenici derivanti da eventi di amplificazione o delezione. Sulla base di criteri quali la funzione biologica e mutazioni ritenute rilevanti per la selettività con cui alterano specifiche regioni codificanti, abbiamo selezionato alcuni geni che, prioritariamente ad altri, andrebbero confermati e analizzati in maggior dettaglio con ulteriori studi..

Un terzo progetto che ho seguito direttamente, ma che comunque non tratterò nella mia tesi, riguarda l'identificazione di fattori genetici coinvolti nella modificazione del rischio per i soggetti portatori di mutazione BRCA1 e BRCA2. Tale progetto si inserisce nell'ambito di un consorzio internazionale, denominato CIMBA (*Consortium of Investigators of Modifiers of BRCA1/2*; (Chenevix-Trench *et al*., 2007). SNP identificati prevalentemente attraverso studi di

associazione *genome-wide* sono stati genotipizzati in 213 *carriers* di mutazione BRCA1/2 della nostra casistica mediante discriminazione allelica con sonde TaqMan. L'analisi statistica è stata effettuata sui dati genotipici provenienti da più di 30 gruppi di lavoro, permettendo la raccolta di dati derivanti da migliaia di individui. In questo modo è possibile ottenere la potenza statistica necessaria per l'identificazione dell'associazione tra polimorfismi con debole effetto e un aumentato rischio di sviluppo del tumore della mammella in specifiche sottoclassi di soggetti già portatori di mutazione nei geni ad alta penetranza BRCA1/2. Effettuati su più di 9400 e 5600 *carriers* di mutazione rispettivamente di BRCA1 e BRCA2, tali studi hanno permesso di identificare un'associazione significativa tra gli SNP rs3817198 (LPS1) e rs13387042 (2q35) ed un aumentato rischio in soggetti portatori di mutazioni BRCA2, e, nel secondo caso BRCA1 o BRCA2. L'identificazione e studio di fattori genetici modificatori permetterà di acquisire una più approfondita conoscenza della predisposizione e patogenesi del tumore della ereditario mammella, ma anche di stimare con migliore precisione il rischio di sviluppo della neoplasia così come di fornire utili informazioni per l'identificazione di nuovi approcci terapeutici.

Summary

Breast cancer is the most common malignancy affecting women in the Western world. At present, several risk factors have been identified and, among them, the most significant is a positive family history for the disease due to the presence of inherited predisposing germline alterations. Despite the discovery of many susceptibility genes, among which the highly penetrant BRCA1 and BRCA2 are the most relevant, more than 70% of the genetic predisposition to breast cancer remains unexplained.

During my three-years Ph.D., I used different analytical approaches to investigate DNA alterations and loci that could be relevantly implicated in the aetiology and progression of hereditary breast cancers.

As first project on non informative BRCA1/2 families, I was involved in the assessment of the pathogenetic role of the BRCA1 p.Val1688del allele, which recurs in the Northeast Italy. Unclassified variants (UVs) provide a considerable challenge in the clinical management of high-risk families. Indeed, the absence of conclusive data on UVs pathogenicity, due to the lack of an evident deleterious effect on protein function, prevents their use in the identification of individuals predisposed to breast cancer development. To assess the pathogenetic role of BRCA1 p.Val1688del, we used the integrated multifactorial likelihood model described by Goldgar *et al*. (2004). Several independent evidences were derived from epidemiologic and linkage studies as well as histopathology, LOH and evolutionary conservation analyses, in 12 families carrying the p.Val1688del allele. All the evidences were than properly evaluated to obtain, from each, the probability of the observed data given that the variant is a deleterious mutation, against the probability of the observed data under the hypothesis that the variant is neutral. The likelihood ratios were integrated to obtain a combined odd of variant causality, resulting in a clear assessment of the BRCA1 p.Val1688del pathogenicity. Indeed, the final integrated odd of 349,000:1 in favor of disease causality largely exceeds the established cut off 1,000:1 needed to state the deleterious effect of the variant.

The second project, described herein, focused on the identification of molecular alterations involved in non-BRCA1/BRCA2 breast cancers. To this purpose, I performed a comprehensive genomic profile of the breast cancer cell line HCC1500, which was established from a patient who probably harboured a predisposing germline mutation affecting genes other than BRCA1 and BRCA2. Indeed, despite the early age of disease development and a positive family history for breast and colon cancer, we did not detect any deleterious BRCA1, BRCA2 or TP53 alteration. The HCC1500 cell line was analysed by two complementary approaches: i) a novel combined strategy named GINI (Gene Identification by NMD Inhibition; Noesie and Dietz, 2001) for the identification of nonsense mutations at the transcriptome level, and ii) a copy number alteration (CNA) analysis that was achieved by array-based high-density SNP genotyping. This approach led to the identification of genes specifically mutated by either point mutations or major genomic rearrangements. In particular, we identified an intronic substitution in the PNLIPRP3 gene, which is predicted to create a PTC by altering pre-mRNA splicing. Moreover, high-resolution CNA analysis allowed the detection of small homozygously deleted regions as well as focal amplifications, involving single or few genes. A gene selection based on the specific biological function and/or type of mutation and/or previously reported involvement in cancer pathways, allowed the definition of a relatively small number of candidate breast cancer genes that will deserve further and more specific investigation.

Finally, among the minor research projects that I will not discuss in my thesis, I was also directly involved in studies carried out by our research group as part of an international consortium, named CIMBA (Consortium of Investigators of Modifiers of BRCA1/2), whose aim is the identification of genetic factors implicated in the modification of BRCA1 and BRCA2 penetrance (Chenevix-Trench *et al*., 2007). Candidate single nucleotide polymorphisms (SNPs), most of which derived from genome-wide association studies, were genotyped in 213 BRCA1/2 mutation carriers from our cohort by the taqman allele discrimination technique, and subsequently statistically analyzed together with the data obtained from more than 30 study groups world-wide. The genotyping of SNP alleles in thousands of BRCA1/BRCA2 mutation carriers provide the necessary

statistical power for the identification of significant association between common SNPs with typical low effects and specific subclasses of individuals at higher risk. These studies were carried out on more than 9400 BRCA1 and 5600 BRCA2 mutation carriers and allowed the identification of a significant association between rs3817198 in LPS1 and rs13387042 at 2q35 (already found to be associated with increased breast cancer risks in the general population) and increased breast cancer risk in BRCA2 and both BRCA1 and BRCA2 mutation carriers, respectively. The identification of genetic modifiers of breast cancer risk not only will favor a better understanding of breast cancer predisposition and pathogenesis but will also allow more precise cancer risk estimates and will represent a useful tool for the design of new therapeutic approaches.

Introduction

1. Breast Cancer: Environmental and Genetic Susceptibility Factors

With an average lifetime risk of 8-10%, breast cancer is the most common malignancy in women in the Western world, accounting for 22% of all female cancers and being, after lung cancer, the second cause of cancer death in women (Ripperger *et al.*, 2009; Oldenburg *et al.*, 2007).

To date, many risk factors involved in the aetiology and development of breast cancer have been identified. Among these, breast cancer incidence rates correlate with gender (approximately, only one out of every 150 breast cancers occurs in males), and with ethnic origin; a difference of 5-10 fold in the incidence and mortality rates has been observed between low- (Far East, Africa and South America) and high-risk (North America and Northern Europe) areas. Moreover, breast cancer incidence show age specific patterns: very low before age 25, it increases up to 100-fold by age 45. After menopause, there is a great divergence in the breast cancer risk among the four different continents (Dumitrescu and Cotarla, 2005). In many epidemiological studies, the extent and duration of exposure to sex hormones has also been consistently identified as a risk factor. Lifetime exposure to endogenous sex hormones have been studied in relation to breast cancer risk: early age at menarche, late age at menopause, nulliparity and late age at first full-term pregnancy increase breast cancer risk, whereas other factors including early age at first full-term pregnancy, higher parity and prolonged breast-feeding are protective against breast carcinogenesis (Dumitrescu and Cotarla, 2005).

The most important and well established risk factor for breast cancer is a positive family history. Epidemiological observations underlining the clustering of breast cancer cases within families dates back to Roman times. Breast cancer is twice as common in women with an affected first-degree relative; the risk of tumor development increases with the number of affected relatives, and it is greater for women with relatives affected at younger age, with bilateral disease or a history of benign breast disease (Oldenburg *et al.*, 2007). The predominant component of the familial aggregation in breast cancer is due to genetic predisposition, as it has been demonstrated by twin studies. Indeed, the risk of a monozygotic twin is substantially higher than that of a dizygotic twin of

an affected individual (Turnbull and Rahman, 2008). It is estimated that 5-10% of all breast cancers arise due to a single high-penetrance, inherited, predisposing allele, whereas the other 90-95% of cases are considered sporadic. Compared with sporadic cancers, hereditary breast cancers arise usually at an earlier age and are often multifocal or bilateral. To date, only 20-25% of the overall excess familial risk is explained by mutations in known breast cancer susceptibility genes (Oldenburg *et al.*, 2007), and more than 70% of the genetic predisposition to familial breast cancer remains unexplained.

Since 1990s, linkage analysis, mutational screening based upon candidate gene approach and, more recently, genome-wide association studies, have allowed the identification of three classes of breast cancer susceptibility alleles: high-penetrance alleles, rare moderate-penetrance alleles, and common low-penetrance alleles. These three groups of breast cancer predisposing alleles confer different levels of risk and have different prevalence in the population (Stratton and Rahman, 2008).

2. Breast Cancer Susceptibility Genes

2.1 High-Risk Breast Cancer Susceptibility Genes

2.1.1 BRCA1 and BRCA2 Penetrance

The hereditary breast and ovarian cancer syndrome is the most frequent autosomal dominant disorder associated with a high breast cancer risk. This syndrome arises as a consequence of germline mutations of BRCA1 and BRCA2 genes (BReast Cancer), conferring a greater than tenfold relative risk of breast cancer (Stratton and Rahman, 2008). Some features within a family pedigree suggest the presence of BRCA1 or BRCA2 alterations: i) two or more individuals in the family with breast and/or ovarian cancer, ii) early onset of breast cancer cases (i.e. before age 40), iii) multifocal and bilateral breast cancer or breast and ovarian cancer in the same individual, iv) male breast cancer cases, and v) Ashkenazi Jewish ancestry (Prucka *et al.*, 2008). Estimates of the penetrance of BRCA1 and BRCA2 mutations (i.e. lifetime risk of mutation carrier of developing the disease) vary according to the nature of cases studied. Studies based on high-risk families, which could be enriched for

other familial risk factors, suggested that the risk of breast cancer by age 70 might have been as high as 87% for BRCA1 and 84% for BRCA2 mutations carriers (Turnbull and Rahman, 2008). A meta-analysis of 22 population-based studies of breast cancer cases, unselected for family history, estimated lower risks: 65% for BRCA1, and 45% for BRCA2 mutation carriers (Antoniou *et al.*, 2003). BRCA1 and BRCA2 are also high-penetrance ovarian cancer genes. Within multiple-cases families, the cumulative risk of ovarian cancer at age 70 in BRCA1 and BRCA2 mutation carriers is 63% and 27%, respectively, whereas population-based studies suggest an average cumulative risk by age 70 years of 39% and 11% for BRCA1 and BRCA2 mutation carriers, respectively (Antoniou *et al.*, 2003).

Though to a much lower extent, BRCA1 and BRCA2 mutation carriers are at increased risk of other cancer types, such as colon and prostate cancer (BRCA1 carriers) and male breast, prostate, laryngeal and pancreatic cancer (BRCA2 carriers) (Oldenburg *et al.*, 2007).

2.1.2 Structure and Functions of BRCA1 and BRCA2

BRCA1 and BRCA2 genes were identified in 1994 and 1995, respectively, through linkage analysis and positional cloning (Narod and Foulkes, 2004). Both BRCA1 and BRCA2 are large genes: BRCA1 has 23 exons, spans approximately 100 kb of genomic DNA and encodes a 1,863 amino acid protein, while BRCA2 has 27 exons, spans about 70 kb, and encodes a protein of 3,418 amino acids. BRCA1 and BRCA2 proteins play a central role in pathways responsible for the integrity of the genome and the maintenance of chromosome stability (Narod and Foulkes, 2004). For this reason, they are included among the caretaker genes that act as sensor of DNA damage.

BRCA1 protein contains three well known domains involved in protein functions (Figure 1A). The amino-terminal RING finger domain (Cys_3-His-Cys_4) mediates important protein interactions. The central region of the protein contains two nuclear localization sequences (NLS), responsible of the mainly nuclear localization of BRCA1.

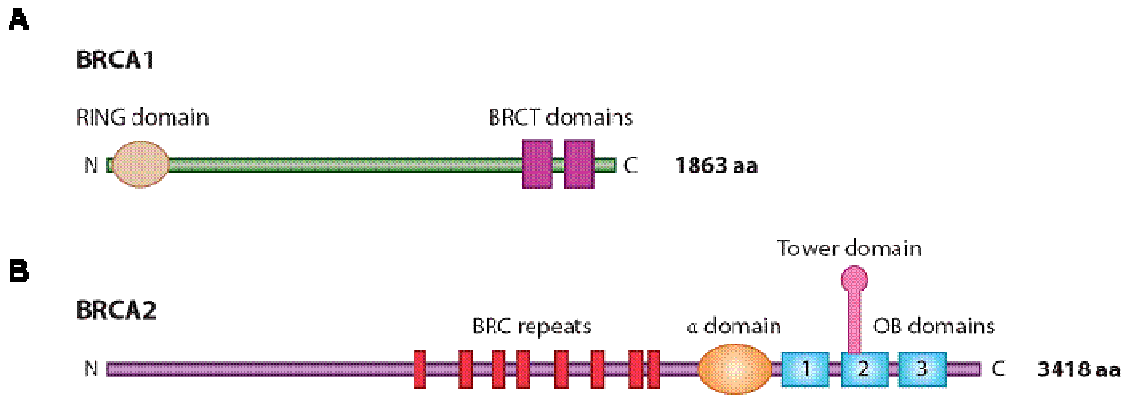

Figure 1. BRCA1 and BRCA2 proteins. Schematic representation of functional domains involved in protein-protein or protein-DNA interactions (modified from Venkitaraman, 2009).

The carboxyl-terminal BRCT domain is a tandem repeat of about 95 amino acids found in a variety of proteins involved in cell cycle checkpoints and DNA repair. Indeed, BRCA1 is involved in the repair of double strand DNA breaks by homologous recombination (HR), co-localizing and associating with the RAD51 recombinase in nuclear foci during S and G2 phases of the cell cycle. BRCA1 is also involved in the nonhomologous end-joining (NHEJ) mechanism of double-strand break repair, and in the nucleotide-excision repair (NER) process, which deals with the repair of single-strand breaks (Narod and Foulkes, 2004; Gudmundsdottir and Ashworth, 2006). When DNA is damaged, BRCA1 is rapidly phosphorylated by the key DNA damage response kinases ATM (Ataxia Telangiectasia Mutated), ATR (Ataxia Telangiectasia and RAD3-Related), and CHK2 (cell cycle checkpoint kinase). Some of the phosphorylated sites allow BRCA1 to modulate the function of specific downstream target proteins, which are involved in the activation of DNA repair response and cell cycle checkpoints (Gudmundsdottir and Ashworth, 2006). Indeed, in addition to its direct role in DNA repair process, BRCA1 is also involved in cell cycle arrest (G1-S, intra-S-phase, and G2-M checkpoints), providing the proper timing for DNA repair processes.

Furthermore, BRCA1 has also been implicated in other biological processes including: i) ubiquitination, through the ubiquitin ligase activity of BRCA1:BARD1 (BRCA1-associated Ring Domain 1) heterodimer, ii) transcriptional regulation,

via its direct interaction with the RNA polymerase II holoenzyme complex and several transcription factors (ER-α, p53, STAT1, c-Myc, CtIP, ZBRK1), and iii) chromatin remodeling, through BRCA1 association with histone deacetylases HDAC1 and HDAC2, as well as components of the SWI/SNF-related chromatin-remodeling complex (Mullan *et al.*, 2006).

The central region of the BRCA2 protein contains eight BRC repeat motifs consisting of a repetitive sequence of approximately 70 amino acids that are essential for the DNA repair protein function. The NLSs sequences and the single-strand DNA-binding domains (DBD) are located In the C-terminal region. This protein region is composed of a conserved globular domain made-up primarily by α-helices, and three OB folds (oligonucleotide/oligosaccharide-binding). Moreover, a largely α-helical domain (termed tower domain) emerges from the second OB-fold and mediates BRCA2 interaction with double-stranded DNA by means of its three-helix bundle motif (Figure 1B).

BRCA2 plays an important role in the control of the cellular recombination processes. Via its BRC motifs, BRCA2 interacts with and regulates RAD51 recombinase during HR, targeting RAD51 to single-stranded DNA generated at a double-stranded break site. Moreover, BRCA2 stabilizes RAD51 nucleoprotein filaments through the interaction with its C-terminal domain named TR2. Like BRCA1, BRCA2 function is regulated by phosphorylation events that occur in the G2 phase of cell cycle, which impair TR2 interaction with RAD51. Thus, BRCA2 phosphorylation is thought to coordinate the DNA repair process with the entry of the cell in the mitotic phase (Thorslund and West, 2007).

Additional BRCA2 functions include the control of the normal homologous recombination processeses (crossing-over) between parental chromosomes in germline cells in combination with RAD51 and the meiosis-specific recombinase DMC1 (Gudmundsdottir and Ashworth, 2006).

2.1.3 BRCA1 and BRCA2 Mutations and Variants of Unknown Significance

The present knowledge on the molecular basis of hereditary breast cancer has allowed the application of genetic tests for the identification of germline mutations that could explain the inherited predisposition to tumor development. Genetic testing of cancer susceptibility genes is now widely applied in clinical practice to predict risk of developing cancer.

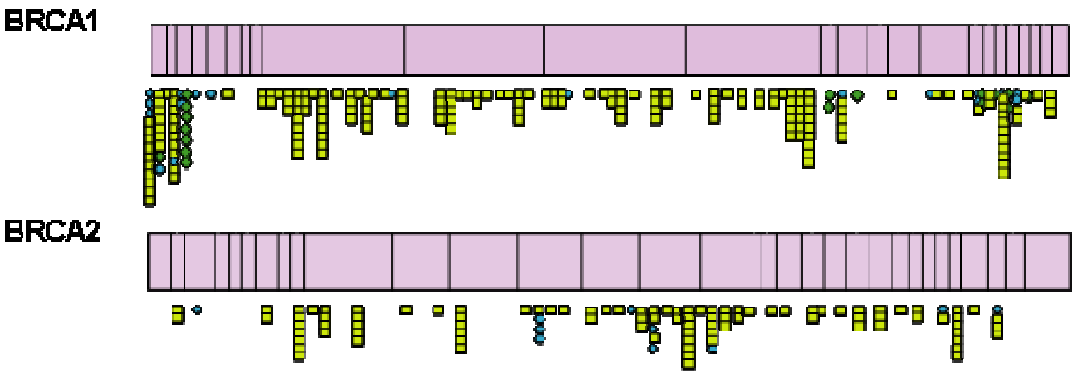

Figure 2. BRCA1 and BRCA2 point mutations spectrum. Missense mutations (●), nonsense mutations (■), and mutations affecting splice-sites (◆) are schematically reported. Mutations spans the entire coding sequence of both genes without mutation hot-spots (modified from http://research.nhgri.nih. gov/bic/).

The mutational screening of the entire coding sequence of BRCA1 and BRCA2 genes (Figure 2) is performed on affected individuals of high-risk families, and lead to three possible outcomes. Genetic testing may detect pathogenic changes, which means sequence alterations that clearly impair protein function (typically truncating mutations). A positive result allows for the extension of the test to other family members and the enrollment of mutation carriers in clinical surveillance and prevention strategies. A non-informative result is given when no variation in DNA sequence is detected. This could be ascribed to insufficient sensitivity of the current technology employed or to the presence of phenocopies, (i.e. breast cancer cases that do not share the genetic predisposition of the other affected family members), as well as to genes other than BRCA1 and BRCA2. In the absence of an informative result, the risk assessment must be based on the clinical history of the patient and her/his family. A third possibility is the detection of a sequence variant with uncertain

clinical significance (UVs, unclassified variants). In other words, it is not possible to readily predict the effect of the variant on gene function and consequently its influence on cancer risk. Genetic variants of uncertain significance are typically missense mutations and small in-frame deletions (1-3 amino acids). The result of genetic test should be considered not informative and medical management should be based on patient's personal and familial cancer history (Berliner and Fay, 2007). Such variants represent a considerable challenge for their clinical implications. In the public database BIC (Brest Cancer Information Core, http://research.nhgri.nih.gov/bic) more than 8,000 affected individuals are reported as carriers of BRCA1/BRCA2 UVs. These data highlight the importance of clarifying the role of UVs to improve genetic counselling and clinical management for a relevant number of families. To this purpose, different epidemiological approaches could be applied. However, the majority of BRCA1 and BRCA2 variants are individually rare and many have been reported only in single families, a factor that clearly impairs the power efficiency of epidemiologic studies aimed to assess UV's pathogenicity. Goldgar *et al.* (2004) described a multifactorial model that provides the probability that a variant is a deleterious mutation rather than a neutral nucleotide change, integrating the contribution to odds of causation of several independent factors. This kind of approach represent a powerful tool to establish the tumor predisposing nature of BRCA1 and BRCA2 variants, thus improving genetic counselling and medical management of high-risk families.

2.1.4 Breast Cancer-Associated Cancer Predisposition Syndromes

In addition to the hereditary breast/ovarian cancer syndrome, there are some other rare cancer predisposing syndromes associated with an increased breast cancer risk (Figure 3).

The Li-Fraumeni syndrome is a rare disease caused by germline mutations of TP53 gene, and breast cancer is one of the several neoplasias that frequently occur at a young age. Though the estimated penetrance of TP53 mutations is 28-56% by age of 45 years, constitutive TP53 mutations are uncommon in non-Li-Fraumeni breast cancer families and the attributable risk of TP53 mutations to familial breast cancer is very low (Oldenburg *et al.*, 2007).

An increased breast cancer risk is present in individuals with the rare Peutz-Jeghers syndrome carrying heterozygous germline mutations in STK11 tumor suppressor gene, which encodes a serine/threonine kinase. The probability of developing cancer by age 65 is estimated to be about 50%, and in particular breast cancer risk ranges between 29% and 54% (Oldenburg *et al.*, 2007).

Another gene associated with the increased breast cancer risk is PTEN, a lipid phosphatase that functions as a tumor suppressor through the negative regulation of cell-survival signaling pathways. Germline mutations of PTEN give rise to the Cowden syndrome, a multiple hamartoma syndrome that includes increased risk of benign and malignant tumors of the breast, thyroid, and endometrium. Women carrying a PTEN mutation have a 25-50% lifetime breast cancer risk. However, no mutations in the PTEN gene have been detected in breast cancer families without features of Cowden syndrome, and PTEN somatic mutations are very rare.

E-cadherin gene CDH1 plays a central role in the maintenance of cell differentiation and the normal architecture of epithelial tissues. Germline CDH1 mutations are associated with hereditary diffuse gastric cancer syndrome, and within families this mutation is associated with an increased risk of breast cancer. Indeed, the lifetime risk of developing breast cancer in CDH1 mutation carriers was estimated at 20-40%. In particular, somatic CDH1 mutations have been found in women with lobular breast cancer.

In addition, also germline mutations in neurofibromatosis type I gene (NF1) and in NBN gene, responsible of Nijmegen breakage syndrome, are described to be associated with a moderately increase of breast cancer risk (Ripperger *et al.*, 2009). Altogether, these conditions account for tiny fraction, about 1-2%, of the hereditary breast cancers with the majority of the familial risk remaining unexplained. Genome-wide linkage analyses in non-BRCA1/BRCA2 families have failed to identify additional high-penetrance susceptibility genes, suggesting that, if they exist, each of them should accounts for a very small fraction of familial risk

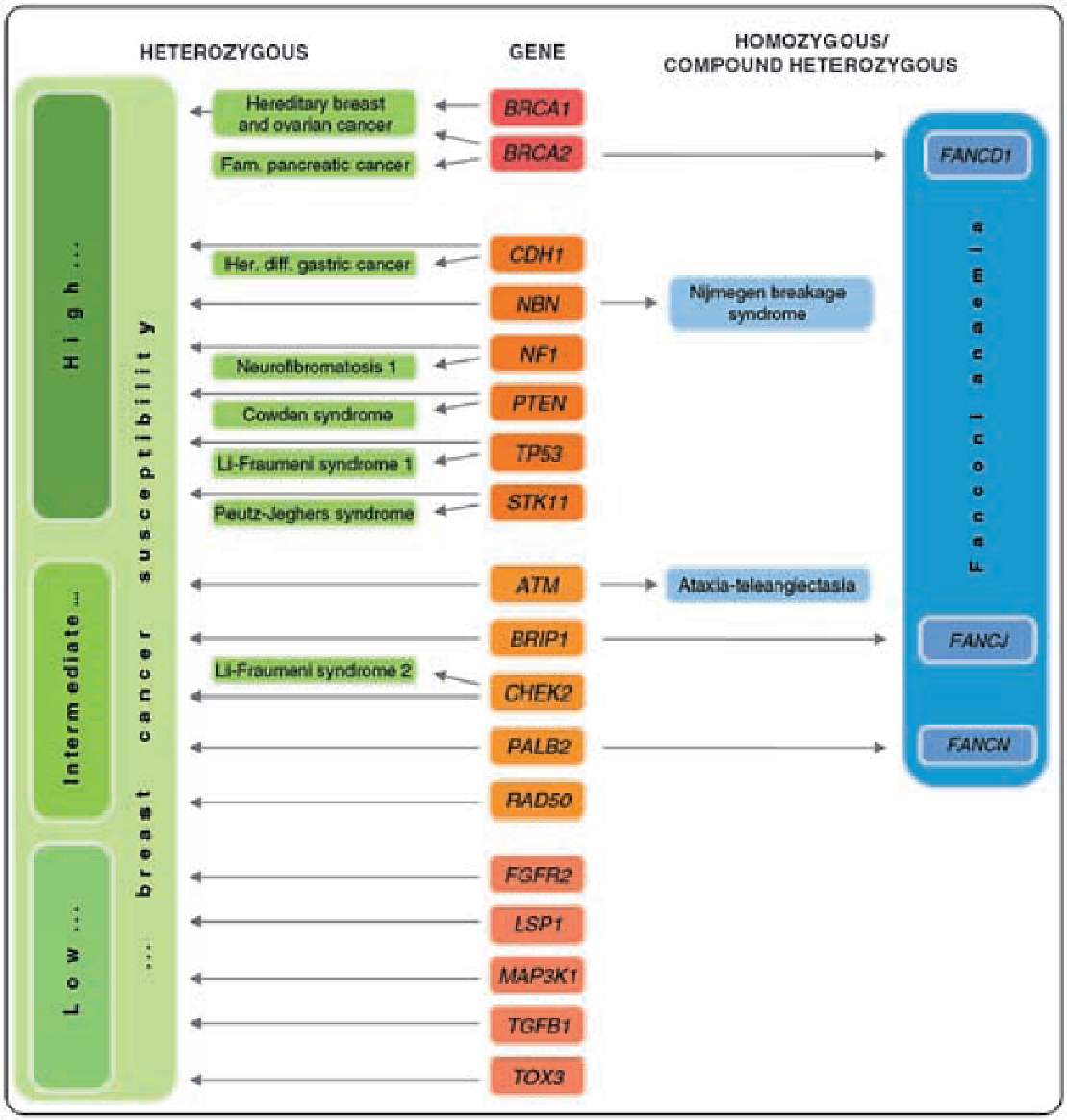

Figure 3. Brest cancer susceptibility genes. To date three class of predisposing genes are known: high-penetrance genes, rare intermediate-penetrance genes, and common low-penetrance alleles. Bi-allelic germline mutation in some of these genes associated with an increased breast cancer risk give rise to distinct severe disorders (from Ripperger *et al*., 2009).

2.2 Rare Intermediated-Penetrance Breast Cancer Predisposition Genes

To date, despite significant efforts made to identify other highly penetrant breast cancer predisposing genes, a BRCA3 gene has not been identified yet, likely due to the genetic heterogeneity of families non-BRCA1/BRCA2, and/or a low mutation frequency of the putative genes involved (Oldenburg *et al.*, 2007). However, sequencing of genes involved in DNA repair processes has allowed the identification of several intermediate-penetrance breast cancer susceptibility genes. In high-risk families, the direct sequencing of candidate genes for coding variants led to the identification of rare mutations affecting i) CHEK2, a checkpoint kinase involved in DNA repair, ii) BRIP1/BACH1, which has a BRCA1-dependent role in DNA repair and checkpoint control, and iii) PALB2, a protein that promotes localization and stability of BRCA2. Heterozygous mutation carriers of ATM, the key kinase involved in the response to double-strand breaks, and the gene responsible for the autosomal recessive disorder ataxia telangiectasia, were also identified carrying a moderate increase in breast cancer risk.

The risk for heterozygous mutation carriers of these genes lie between 2.0X and 4.3X, classifying them as genes with a moderate penetrance (Ripperger *et al.*, 2009). Carriers of these alleles have approximately a 6-10% risk of developing breast cancer by age 60, compared to 3% of the general population (Stratton and Rahman, 2008). The multiple pathogenic mutations of these four genes are individually very rare, and collectively account for only approximately 2.3% of the overall familial risk of breast cancer. Interestingly, biallelic germline mutations affecting BRIP1/BACH1, PALB2, as well as BRCA2, are responsible for different clinical subgroups of the inherited genomic instability disorder Fanconi anemia (FA) that are named FANCJ, FANCN, and FANCD1, respectively (Figure 3).

An association between moderate increase in breast cancer risk and RAD50 mutations was also identified (4.3X for the founder mutation 687delT). However, RAD50 involvement in breast cancer predisposition seems likely to be specific for the Finnish population (Ripperger *et al.*, 2009).

A feature of the intermediate-penetrance susceptibility genes is the incomplete cosegregation of mutant alleles with the disease. This is because most mutation carriers do not develop breast cancer because of the low risk, and more

variants in multiple genes have likely to occur in the same individual for disease to develop.

2.3. Common Low-Penetrance Breast Cancer Predisposition Alleles

A third class of genes in the landscape of breast cancer susceptibility has been recently found through association studies that identified low-penetrance breast cancer susceptibility alleles. In order to reach the adequate statistical power necessary to detect common single nucleotide polymorphisms (SNPs) associated with an expected small effect on individual risk, world-wide multigroup collaborations have combined data to achieve very large numbers of cases and controls. Three genome-wide association studies have reported seven new low-penetrance breast cancer susceptibility loci at highly stringent levels of statistical significance (Easton *et al.*, 2007; Stacey *et al.*, 2007; Cox *et al.*, 2007). The SNPs associated to date with low breast cancer risk are located within FGFR2 (Fibroblast Growth Factor Receptor 2), LSP1 (Lymphocyte-specific protein 1), MAP3K1 (Mitogen-Activated Protein Kinase kinase kinase 1), TNRC9/TOX3 (putative transcription factor), TGFB1 (Transforming Growth Factor Beta 1) genes, as well as in 8q24 and 2q35 loci where no known protein-coding genes reside. Moreover, Cox *et al.* reported on a common SNP in CASP8 (caspase-8) associated with a decreased risk of breast cancer. Although the relative risk of breast cancer associated with a single copy of each risk allele ranges between 1.07 and 1.26, the population frquency of each of these SNP is rather high (28%-87%). Thus, despite the low risk conferred by these alleles, their contribution to the population attributable risk of breast cancer could be relevant. The six loci identified by Easton *et al.* and Cox *et al.* are estimated to account for 3.9% of the excess familial risk of breast cancer in European population (Stratton and Rahman: 2008).

3. Somatic Alterations in Breast Cancer

Breast cancer is a highly genetically heterogeneous disease with respect to the loci involved and the specific somatic mutations that have been found. In the attempt to better characterize breast cancer at molecular and genomic level, a

large number of studies investigated the expression and genomic profiles of breast tumor cells. The development of solid tumors involves acquisition of genetic and epigenetic alterations, and concomitant changes in gene expression, which modify normal growth and survival pathways.

Array comparative genomic hybridization (CGH) is a technique for the detection of Copy Number Alterations (CNA) such as deletions and amplifications (Kallioniemi *et al.*, 1992); it is a powerful tool to detect genomic alterations that occur as a result of genomic instability, thus providing new insights in cancer genetics. The comparison among genomic profiles of different classes of breast tumors (hereditary and sporadic cancers, different histopathological subtypes) have been shown to be useful in identifying specific aberration patterns. Hereditary breast cancers show specific alterations that differentiate them from sporadic tumors, suggesting different tumor progression pathways (Palacios *et al.*, 2008). Overall, BRCA1-associated tumors have a higher frequency of CNA than BRCA2-associated and sporadic tumors (van Beers *et al.*, 2005). BRCA2-associated and sporadic breast cancer have more similar alteration patterns, even if few chromosomal regions more frequently altered in BRCA2 tumors have been identified. The genetic changes more commonly found in BRCA1 tumors are gains of 8q and 3q, as well as losses of 4p, 4q and 5q. BRCA2 tumors present gains of 8q, 19q and 20q and lose 8p, 13q and 11q (Palacios *et al.*, 2008). In familial non BRCA1/BRCA2 associated tumors a significant higher frequency of gain in 8q, 19p, and 19q, and loss on 8p was observed (Oldenburg *et al.*, 2007).

Chromosomal regions showing a higher frequency of genomic imbalance within specific class of breast tumors represent candidate loci for the search of genes involved in the aetiology and progression of the breast cancer. In particular, specifically amplified regions are useful in mapping oncogenes, being their increased expression one important mechanism of their activation. In breast cancer, several discrete amplicons have been identified including 17p12 (ERBB2/HER-2), 8q24 (c-Myc), 11q13 (CCND1), and 10q26 (FGFR2). Also the 8p11-p12 region has been reported amplified in 10-15% of human breast cancers (Ray *et al.*, 2004).

Moreover, somatic loss of heterozygosity (LOH), i.e. deletion of one allele at a heterozygous locus, has been widely reported as a means to identify putative

tumor suppressor genes. Miller *et al.* (2003) analyzed 151 published LOH studies on breast cancer, highlighting several small DNA regions of recurrent loss likely harbouring putative tumor suppressor genes.

The new technologies for the detection of copy number alterations (CNAs) are constantly improved, providing great advances in our understanding of the tumor cells genome. To date, the sensitivity and analytical resolution have been largely increased due to the high number of markers interrogated throughout the whole genome. In recent years, high-density single nucleotide polymorphisms (SNPs) mapping has provided a suitable means to obtain a detailed portrait of cancer genome alterations. The genotyping of hundreds of thousand of SNPs can allow the simultaneous identification of CNA and loss of heterozygosity (LOH), and, notably, can unveil copy-neutral changes (i.e. two copies of the same allele with lost of the other one), a condition referred to as uniparental disomy (UPD) (Tuna et al., 2009).

4. A Novel Integrated Analysis for Putative Tumor Suppressor Gene Identification

One of the major challenges in cancer genetics research is the identification of the genes that contribute to the development of the malignant phenotype. Noensie and Dietz (2001) developed a novel strategy for the identification genes harboring nonsense mutations that underline human diseases. The approach is based on the inhibition of one of the best characterized mRNA quality control pathways, the Nonsense-mediated mRNA decay (NMD). NMD triggers the degradation of mRNAs harboring a premature termination codon (PTC), avoiding the synthesis of potentially harmful truncated protein (see below for the molecular mechanism description). The strategy was named “Gene Identification by NMD Inhibition” (GINI): preventing NMD-mediated mRNAs degradation, PTC-containing transcripts are stabilized. Comparing NMD competent and NMD-inhibithed cells, mutant mRNAs could be identified, because of their higher expression levels in treated cells. At present, the most commonly used technology for a widespread evaluation of differences in mRNA levels is the microarray expression profiling. Microarray technology allows an

analysis at the transcriptome level, providing the tool for a complete monitoring of the differences in transcripts levels between two groups of cell populations. This approach can be readily applied to investigate genes involved in recessive diseases, in which there is a biallelic gene inactivation, because wild type allele would otherwise mask the increase of level of premature terminating mRNA (Noensie and Dietz, 2001).

Due to the strong interconnection between the NMD and translation pathways (see below), the NMD machinery can be inhibited by a pharmacological block of protein synthesis. Among different translation blocking drugs, emetine has been shown to be the most effective and selective in stabilizing nonsense reporter transcripts (Noensie and Dietz, 2001). Background emetine response, at least partially, was determined by treating normal cell lines and identifying *wild type* transcripts with an increased expression level.

The major drawback of the GINI method is the nonspecific increased expression of genes involved in emetine-induced stress response, making difficult to select candidate genes for sequencing analysis. To further improve the identification strategy of mutant genes, after emetine treatment, cells were also incubated with actinomycin D, a transcription blocking drug (Ionov *et al.*, 2004). In the absence of newly synthesized mRNA, more rapid degradation of stress response transcripts than mRNA truly harboring nonsense mutations allows a better identification of the latter.

The GINI approach has been used to identify nonsense mutations involved in several kind of tumors, such as mantle cell lymphoma (Pinyol *et al.*, 2007), melanoma (Bloethner *et al.*, 2008), prostate cancer (Rossi *et al.*, 2005; Kunnev *et al.*, 2009), and colon cancer (Ionov *et al.*, 2004; Ivanov *et al.*, 2007). These studies confirmed the efficiency of the GINI method in the identification of mutant genes in cancer cell lines (even though the relevance of the specific mutation identified has to be eventually demonstrated through functional analyses and mutational screening on a larger number of tumor samples).

Chances of identifying mutated genes could be increased by combining data on gene expression profiling with additional information on the disease under study. This information could consist, for example, either in the biological function of candidate genes or in a known genomic map position for a given phenotype. Huusko *et al.* combined GINI with high resolution data on deleted

genomic regions in cancer cell lines obtained with array-based comparative genomic hybridization (aCGH). In this case, the selection of candidate genes was achieved by identification of genes i) whose transcripts are stabilized following NMD pathway inhibition and ii) that have also lost one allele. Thus, the integration of these two array-based analyses allows a genome-wide identification of bi-allelic inactivation events. In cancer cells genomic alterations that prevent the synthesis of functional protein are often a hallmark of tumor suppressor genes, whose inactivation is advantageous for tumor cell growth and survival. In the DU 145 prostate cancer cell line this strategy led to the identification of a hemizygous nonsense mutation in EPHB2 gene, which encodes a tyrosine kinase receptor (Huusko *et al.*, 2004). Further evaluation of EPHB2 alterations in prostate tumor samples, led to identification of cancer-associated mutations in 8% of primary and metastatic specimens, and functional in vitro experiments unveil a role as a tumor suppressor gene in prostate cancer progression. Muggerud and co-workers (2009) applied this integrated genomic approach to identify putative tumor suppressor genes in breast cancer. Among cell lines analyzed, in ZR-75-I cells a truncating mutation affecting RIC8A gene (a guanine nucleotide exchange factor), was identified.

5. The Nonsense-Mediated mRNA Decay (NMD) Pathway

One of the cellular pathways that controls the fidelity and accuracy of gene expression is the Nonsense-Mediated mRNA Decay (NMD). The NMD pathway recognizes and degrades mRNA harboring PTC located at least 50-55 nucleotides upstream of the last exon-exon boundary (Nagy and Marquat, 1998), thus preventing the synthesis of potentially deleterious truncated proteins. Erroneous, PTC harboring transcripts can arise owing to genomic frameshift (insertions or deletions) or nonsense mutation, as well as a result of errors in pre-mRNA splicing, leading to intron derived stop codon or a nonsense codon downstream from the site of missplicing, and unproductive DNA rearrangements in T cell receptor and immunoglobulin loci.

Nonsense-mediated mRNA decay is an evolutionary conserved mechanism involved in the surveillance and regulation of gene expression in all eukaryotes

examined so far, even if the mechanism of PTC recognition differs among species (Behm-Ansmant *et al.*, 2007). During the pre-mRNA splicing, a multiprotein complex called Exon Junction Complex (EJC) is deposited 20-24 nt upstream of exon-exon junctions (Le Hir *et al.*, 2000), and during the first mRNA translation event the NMD machinery checks if an intron is present downstream of the stop codon (Figure 4). If an EJC is detected downstream of a stop codon, the transcript is recognized as premature terminating and triggered to degradation (Maquat, 2004).

5.1 NMD Core Proteins

UPF2 and UPF3 together with UPF1 are the three principal players of the NMD pathway. UPF1 is a phosphoprotein that localizes mainly in the cytoplasm, and contains domains with ATP-dependent 5' to 3' helicase activity and RNA-dependent ATPase activity. UPF1 interacts with eukaryotic translation release factors eRF1 and eRF3, and with UPF2 and UPF3, providing a link between the mRNA survelliance complex and the translation machinery (Conti and Izaurralde, 2005).

UPF2 localizes in the perinuclear region of the cytoplasm, and it is the molecular bridge that mediates the interaction between UPF1 and UPF3 to elicit NMD (Lykke-Andersen *et al.*, 2000).

UPF3 is a predominantly nuclear protein that shuttles to the cytoplasm. There are two UPF3 genes, UPF3a on chromosome 13, and UPF3b on the chromosome X. These homologous proteins are unlikely to be redundant, because depletion of human UPF3b, but not of UPF3a, impairs NMD of some mRNA containing nonsense codons (Behm-Ansmant *et al.*, 2007).

5.2 Molecular Mechanism of NMD

The recognition of PTC is a translation-dependent process that involves a cross-talk between the ribosome stalled at a stop codon and the downstream EJC *cis*-acting signal on the mRNA. To efficiently prevent the synthesis of erroneous truncated proteins, the NMD pathway acts to degrade premature terminating mRNAs as early as possible. In fact, experimental evidences shown that, in mammals, NMD occurs during the so called pioneer round of translation, i.e. the first event of neo-synthetized mRNA translation (Chang *et al.*, 2007).

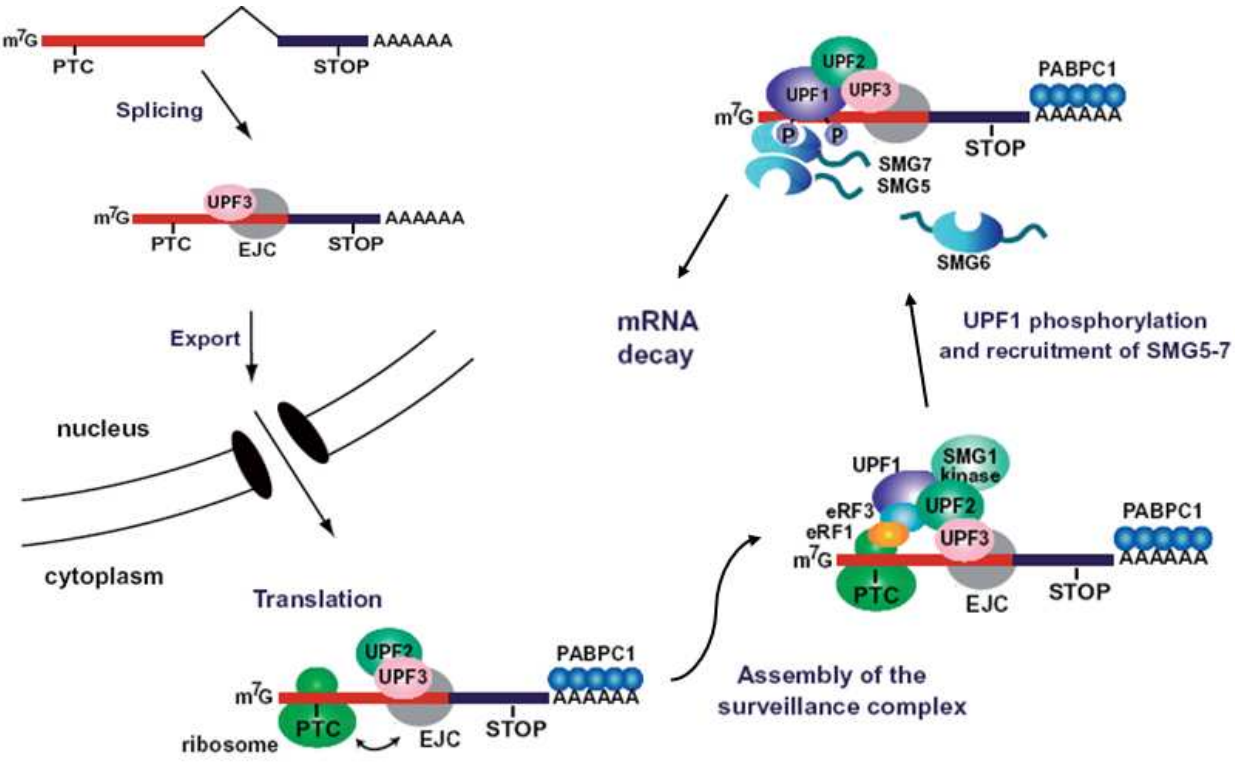

Figure 4. Nonsense-mediated mRNA Decay (NMD) pathway. Schematic representation showing the main steps in premature termination codon (PTC) recognition. The presence of an EJC (Exon Juntion Complex) downstream of a stop codon targets mRNA as a premature terminating: upon assembly of the surveillance complex (UPF1, SMG1, eRF1, eRF3), UPF1 can interact with downstream UPF2 and UPF3 EJC-associated proteins. NMD core proteins interaction triggers SMG1-mediated UPF1 phosphorylation, which in turn determines UPF1 interaction with SMG5-6-7 proteins. Nonsense mRNA degradation occurs via both 5' to 3' and 3' to 5' exoribonuclease pathways (modified from Behm-Ansmant *et al*., 2007).

The heterodimer CBP80:CBP20 (Cap Binding Protein) is bound at the 5' end of newly synthesized mRNAs, and this 5'-cap binding complex is later replaced by eIF4E, a translation initiation factor that support translation for the bulk of proteins production. It has been evidenced that EJC (and thus essential NMD factors) is present on mRNAs bound by CBP80:CBP20 but not on eIF4E bound transcripts (Isken and Maquat, 2007).

During the pioneer round of translation, two proteins involved in NMD pathway, UPF1 and the phosphoinositide-3-kinase-related protein SMG1, are recruited at the termination codon, together with the eukaryotic release factors eRF1 eFR3 involved in translation termination, to form the so called SURF complex (SMG1,

UPF1/2,eRF1/3) (Figure 4). When this complex is assembled on a premature termination codon located at least 50-55 nt upstream the last exon-exon boundary, UPF1 can interact with the UPF2 protein located on the downstream EJC. The establishment of this molecular cross-talk between SURF complex and EJC activates SMG1 to phosphorylate UPF1. Phosphorylated UPF1 is recognized by SMG5, SMG6 and SMG7 proteins that contain a 14-3-3-like domain through which they bind UPF1 at phosphoserine residues. SMG5-7 proteins promote UPF1 dephosphorylation by the recruitment of the protein phosphatase 2A (PP2A) (Behm-Ansmant *et al.*, 2007). The UPF1 phosphorylation and its subsequent dephosphorylation are essential steps to drive mRNA decay. SMG7 is thought to be the terminal effector in NMD, providing the link between the PTC recognition machinery and the mRNA degradation process. Through its C-terminal region, SMG7 localizes in cytoplasmatic foci called P-bodies (mRNA Processing bodies), a dynamic compartment of eukaryotic cells that harbors high concentrations of RNA decay factors (Chang *et al.*, 2007). Both the decapping-dependent 5'-to-3' exoribonuclease pathway and the exosome-mediated 3'-to-5' exoribonuclease pathway rapidly degrade mRNAs harboring PTCs.

Recently, it was also suggest that NMD can occur even in the absence of a downstream EJC, an alternative mechanism of PTC recognition that mainly relies on the capability of UPF1 to interact with Poly(A) Binding Protein C1 (PABPC1), which binds to poly(A) mRNA tail (Stalder and Muhlemann, 2008). Indeed, in human cells artificial tethering of PABP into close proximity of an otherwise NMD-triggering PTC efficiently suppresses NMD (Muhlemann *et al.*, 2008).

6. NMD in mRNA Decay and Gene Expression Regulation

Gene expression profiling experiments in *S. cerevisiae*, *D. melanogaster* and human cells depleted of essential NMD factors have revealed that 3-10% of the transcriptome is regulated by NMD factors, either directly or indirectly. These data indicate an important role of NMD in gene regulation that extends beyond mRNA quality control (Mühlemann *et al.*, 2008). These studies on the physiological role of NMD revealed that the repertoire of genes regulated by this

surveillance mechanism is not conserved among different species (Behm-Ansmant *et al.*, 2007).

Among the NMD targets there are also naturally occurring wild type mRNA with features recognized by the surveillance machinery. These physiological substrates of NMD pathway are i) transcripts with upstream Open Reading Frame (uORF), that is, mRNA with alternative AUG translational initiation codon in the 5'-UTR leading to premature termination; ii) genes with introns in the 3'-UTR at a distance greater than 50-55 nt from the termination codon; iii) nonsense-containing transposon, retroviral sequences or pseudogenes; and iv) selenoprotein mRNAs, which harbor UGA codons encoding for selenocysteine residues instead of a stop codon (Mendell *et al.*, 2004).

In mammalian cells, a primary source of nonsense codons that elicit NMD appears to consist of alternative pre-mRNA splicing, which affects the expression of about 60-75% of human genes (Lejeune and Maquat, 2005). In fact, up to one-third of human alternative splicing events create a premature termination codon (McGlincy and Smith, 2008). Moreover, a coupling between NMD and alternative splicing could lead to a quantitative gene expression regulation, fine-tuning the amount of mRNA that will be translated. Interestingly, certain gene families expression seems to be regulated by a mechanism that couples alternative splicing and NMD. Indeed, NMD is involved in an auto-regulatory gene expression mechanism termed 'Regulated Unproductive Splicing and Translation' (RUST) (Mühlemann *et al.*, 2008). The splicing-promoted inclusion or exclusion of specific sequences in the mRNA results in the generation of a premature termination codon that triggers NMD. Experimental evidences suggest that alternative splicing-NMD coupling might be a conserved mechanism in serine/arginine-rich proteins (SR-proteins) level regulation (Lareau *et al.*, 2007), as well as in the expression regulation of hnRNPs (heterogeneous nuclear ribonucleoproteins) and many core spliceosomal proteins. Thus, alternative splicing-NMD-mediated negative feedback could be an important mechanism for the homeostatic regulation of splicing factors (McGlincy and Smith, 2008).

Among NMD targets, genes involved in amino-acid homeostasis seem to be a class of NMD-regulated transcripts. Indeed, Mendell *et al.* (2004), analyzing gene expression profiling of UPF1 silenced human cells, found upregulation of

15% of the interrogated transcripts that belong to the ontologic category of ‘amino-acid transport’, ‘amino-acid biosynthesis’ and ‘amino-acid activation.

Materials & Methods

1 Families Recruitment

Families were identified through the Hereditary Breast/Ovarian cancer Center of the Istituto Oncologico Veneto (Padova), and through Istituto Nazionale Tumori and Istituto Oncologico Europeo (Milan). Recruitment was based on published operational criteria (Federico *et al*., 1999). Briefly, families considered carriers of breast cancer predisposing germline mutations are those that have at least one of the following criteria: i) a member affected by both breast and ovarian cancer, ii) early onset breast cancer (<36 years old), iii) bilateral disease, iv) males affected by breast cancer, v) and breast cancer affected individuals with at least 2 first-degree relatives affected by breast (<50 years, or bilateral disease, or a male affected) or ovarian cancer. The presence of at least 3 affected first-degree relatives is necessary for the recruitment of individuals affected by breast cancer older than 50.

2 DNA extraction

DNA purification was performed using QIAamp DNA Mini Kit (Qiagen). Cells were resuspended with 200 µl PBS 1X and were then lysated adding 200 µl AL buffer and 20 µl proteinase K. To obtain a complete cell lysis, samples were vortexed 15 sec and then incubated 30 min at 56℃. 200 µl of ethanol 96-100% were added to samples to allow optimal binding conditions of DNA to column membrane. Samples were vortexed for 15 sec and through a centrifugation step (1 min at 6000 x g) DNA was adsorbed onto the silica membrane. DNA was then washed by two centrifugation step to remove any residual contaminant. 500 µl AW1 buffer were pipetted into spin column and eluted by 1 min centrifugation at 6000 x g; then 500 µl AW2 buffer were pipetted into spin column and eluted by 3 min centrifugation at 20000 x g. To completely eliminate the chance of possible buffer AW2 carryover, samples were further centrifuged 1 min at 20000 x g. Purified DNA was eluted pipetting directly onto the silica membrane 100 µl of water, incubating the spin column 20 min at room temperature, and centrifuging samples 1 min at 6000 x g. Elution step was repeated twice to increase DNA yields.

3 Denaturing High Performance Liquid Chromatography (DHPLC)

Denaturing High Performance Liquid Chromatography (DHPLC) is a Transgenomic technology that allows the identification of single nucleotide polymorphisms or point mutation as well as small insertions and deletions. DHPLC was used to investigate the presence of deleterious mutation in BRCA1 and BRCA2 genes.

BRCA1 and BRCA2 genes were entirely amplified through 35 and 47 PCR reactions, respectively. After PCR reactions, amplified fragments were heated at 96℃ for 5 min, and then slowly cooled to 10℃. This decreasing temperature ramp allows the formation of the so called heteroduplexes, that is double-stranded PCR products made up by the two nearly complementary strands derived from two heterozygous alleles. Thus, there are four possible configurations for heterozygous samples (Figure 5). Since in HCC1500 cells BRCA2 gene has only one allele, to allow the formation of heteroduplexes cell line genomic DNA was mixed with genomic DNA of a wild type control sample (2:1 ratio).

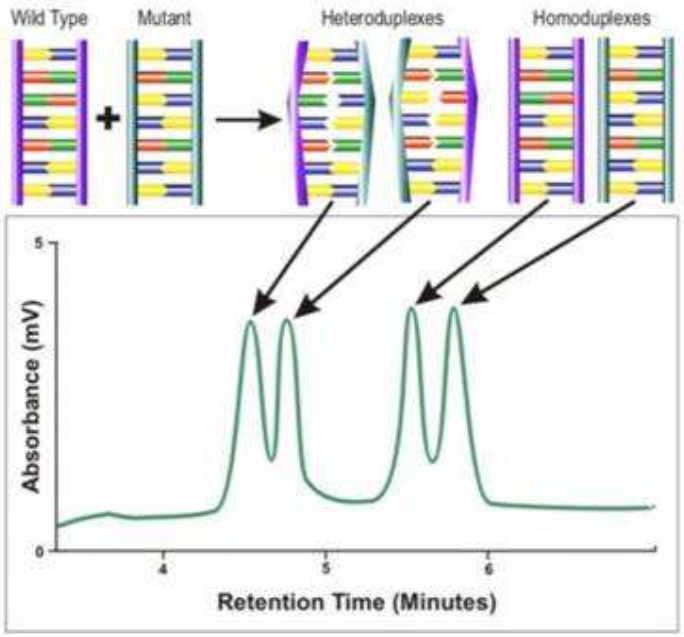

Figure 5. DHPLC homoduplexes and heteroduplexes elution pattern. The slow PCR product renaturation allows the formation of homoduplexes and heteroduplexes molecules. The differential interactions with the chromatographic column allow the discrimination of all four possible double-stranded molecules configurations (from Transgenomic website).

The mutational detection analysis is performed at partially denaturating temperature condition, which favors a differential interaction of homoduplexes

and heteroduplexes fragments with the chromatographic column. The samples are injected into the buffer flow path containing triethylammonium ion (TEA+) and acetonitrile (ACN). The positively charged portion of TEA+ forms an association with the negatively charged phosphate backbone of the DNA; therefore molecules bind to the hydrophobic chromatographic column. Based upon specific DNA fragment features, oven temperatures that allow the partial denaturating status of amplified products and ACN gradient conditions are set up (Navigator™ Software). Over time, the increasing ACN concentration of the solution injected into the chromatographic column reduces the DNA-binding capabilities of the TEA+ ions, and the DNA fragments are released from the column. Heteroduplexes, due to their lower column binding affinity, elute earlier than homoduplexes.

The PCR fragments pass through the UV detector, which detects the nucleotide absorbance (260 nm) over time, and information is recorded as a chromatogram. If no mutations are present, all of the homoduplex DNA fragments elute at the same time, producing a single peak on the chromatogram. If a mutation is present, two to four peaks will be visible.

To assess the sensitivity of DHPLC analysis, for each PCR product positive controls, that is heterozygous samples, are included.

4 Multiplex Ligation-dependent Probe Amplification (MLPA)

PCR and DHPLC analysis do not evidence large heterozygous genomic rearrangements; thus, to investigate large deletions and amplifications wthin BRCA1 and BRCA2 genes, the multiplex PCR method ‘’Multiplex Ligation-dependent Probe Amplification” (MLPA, MRC-Holland) was used.

The MLPA reaction can be divided in four steps (Figure 6). 1) DNA denaturation and hybridization of a mixture of MLPA probes. 100 ng of DNA sample in 5 µl are denatured for 5 min at 98℃. Subsequently, 1.5 µl of buffer and 1.5 µl of MLPA probes mix are added at 25℃. Each MLPA probes pair consist of two separate oligonucleotides, designed to hybridize to immediately adjacent target sequences, and both contain one of the two universal PCR primer sequences. In fact, a single PCR primers pair is used for a subsequent amplification step. After 1 min at 95℃ following probe addition, samples are incubated over night (at least 16 hrs) at 60℃ allowing hybridization to occur. 2) Ligation reaction.

Only when both the two probes hybridized to their adjacent targets, a ligation reaction could occur. 1 µl Ligase-65 is mixed with 3 µl Ligase-65 buffer A, 3 µl Ligase-65 buffer B, and 25 ml H_2O. This ligation buffer mix was added to samples at 54℃, and the ligation reaction occurred for 15 min at 54℃. Ligase inactivation was obtained incubating samples 5 min at 98℃. 3) PCR reaction is performed by mixing 4 µl SALSA PCR buffer, 31.5 µl H_2O, 10 µl MLPA ligation reaction, 2 µl SALSA PCR-primers, 2 µl SALSA Enzyme Dilution buffer, and 0.5 µl SALSA Polymerase, through 35 cycles at: 30 sec at 95℃, 30 sec at 60℃, 1 min 72℃. Because only ligated probes will be expon entially amplified during the subsequent PCR reaction, the number of probe ligation products is a measure of the number of target sequences in the sample.

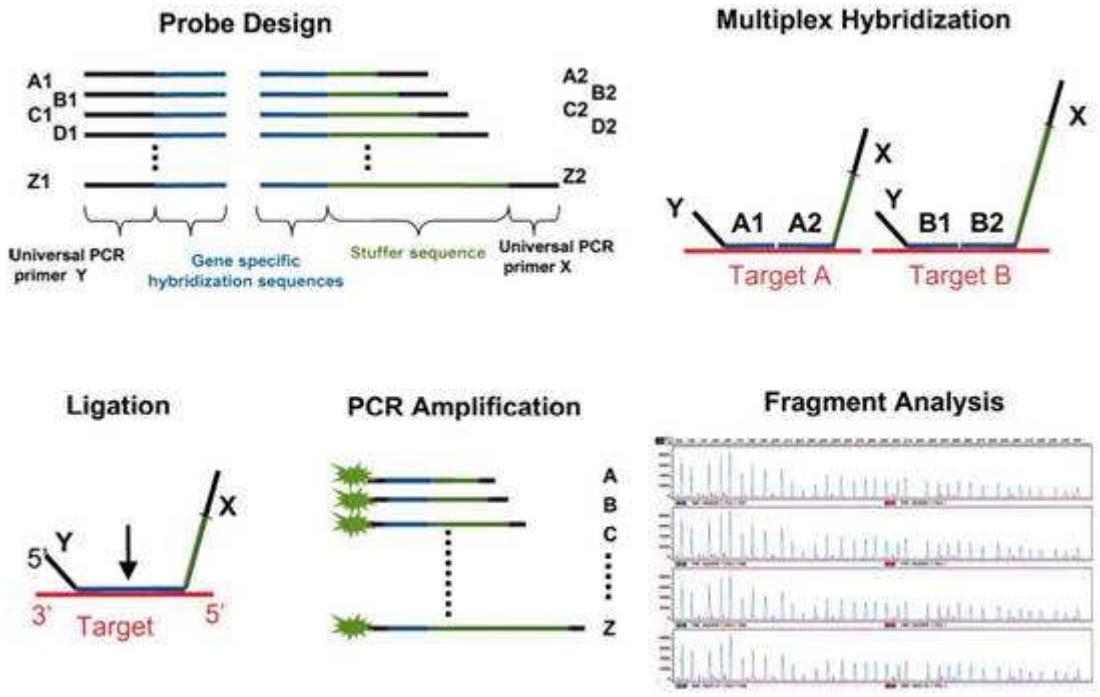

Figure 6. MLPA procedure. Whenever target sequences are present, the two probes of each MLPA pair hybridize to adjacent sequences and are subsequently ligated. All ligated MLPA probes are PCR amplified through universal fluorescent-labelled primers, and then separated by capillary electrophoresis, due to the different size of stuffer sequences (from MRC Holland website).

4) Separation of amplification products, which range between 130 and 480 nt in length, by capillary electrophoresis (ABI-PRISM 3130xl, Applied Biosystems). Within the multiplex PCR, one probe in each oligonucleotide pairs contains a stuffer sequence, different in length between the different probes.

Differences in probes amplification are detected by comparing the sample peak pattern to that of the reference sample. Probes that have different amplification signal in a sample relative to the reference infer the presence of sequences with aberrant copy numbers. Probe oligonucleotides that are not ligated cannot be amplified and will not generate a signal. On the other hand, gene amplification give to MLPA probes more template to generate more PCR product relative to normal reference samples.

Within the MLPA probe mix there is a pool of internal controls for the evaluation of the accuracy of the reaction, i.e. proper DNA amount, denaturation, probe-to-target hybridization, and ligation efficiency.

After the initial evaluation of raw data, electrophoresis results are normalized. Statistical analysis is performed using the Coffalyser software (MRC-Holland). Normalization of MLPA data is achieved dividing the peak area of each amplification product by the total peak area of the reference probes in the probemix. Furthermore, an additional correction could be made for a possible length-dependent bias in fragment amplification (the decrease of peak high for longer fragments).

5.1 Co-Segregation Analysis

To assess the co-segregation of p.Val1688del BRCA1 variant with the disease, patients' peripheral blood was processed to isolate PBMCs and DNA was extracted through the phenol-chloroform method. To identify the BRCA1 variant, exon 17 was PCR amplified (17F 5'-GGGTTTCTCTTGGTTTCTTT-3'; 17R 5'-TCGCCTCATGTGGTTTTA-3') and sequenced (ABI-PRISM-3130XL, Applied Biosystems).

Co-segregation analysis was performed using the Linkage package (http://capella.uni-kiel.de/gmc2006/linkage/user.pdf), assuming an allelic frequency of 0.0001 and a 75% penetrance, according to the literature for pathogenetic BRCA1 mutations.

5.2 Co-Occurrence Analysis

The presence of a well established pathogenetic BRCA1 mutation in p.Val1688del carriers was investigated through a mutational screening of the entire coding sequence of BRCA1 gene. The screening for point mutations and small insertions or deletions identification was achieved by DHPLC analysis (Denaturing High-Performance Liquid Chromatography), whereas the absence of BRCA1 large genomic rearrangements was assessed by MLPA (Multiplex Ligation-dependent Probe Amplification).

To integrate the co-occurence data in the multifactorial likelihood ratio model proposed by Goldgar *et al.* (2004), an overall frequency of deleterious BRCA1 mutations of 10% was used. This percentage is based on the ratio of pathogenic BRCA1 mutation carriers in our cohort of about 800 high risk breast/ovarian cancer families.

5.3 Loss of Heterozigosity (LOH) Analysis

Paraffin embedded tumor samples from patients were processed to isolate DNA by QiAmp DNA Mini Kit (Qiagen) while DNA from PBMC was extracted by phenol-chloroform standard procedure.

The BRCA1 gene fragment containing p.Val1688del was PCR amplified from both PBMC-derived and tumor-derived genomic DNA, and the genotype was investigated by DHPLC. To discriminate by sizing the two alleles in each sample, PCR products were injected into the chromatographic column in non denaturing conditions (50°C), and gradually eluted by an acetonitrile increasing concentration gradient.

Likelihood ratios were calculated using the probability distribution for the three possible LOH results (no LOH, loss of the wild type allele, or loss of the variant allele), as previously described by Osorio *et al.* (2007).

5.4 Short Tandem Repeat (STR) Genotyping

Four microsatellites (D17S855, D17S1322, D17S1323, and D17S1327) spanning BRCA1 locus were genotyped to determine a possible common haplotype among p.Val1688del carriers.

Each microsatellite was independently amplified using a FAM-labeled primer. PCR products were analyzed by capillary electrophoresis (ABI-PRISM 3130xl, Applied Biosystems), and the size of microsatellites determined by Gene Mapper v3.7 software (Applied Biosystems). The allelic phase was determined by genotyping parents of index cases and identifying the haplotype of the available tumor samples with somatic loss of the *wild type* BRCA1 allele.

5.5 Philogenetic Sequence Conservation

The constrain position likelihood for amino acids Val1687 and Val1688 was calculated using parameters previously defined by Tavtigian *et al*. (2006). In particular, based on a 12-sequence alignment, that contains full-length *BRCA1* sequences from eight placental mammals plus *M. domestica*, *G. gallus*, *X. laevis*, and *T. nigroviridis*, a likelihood ratio of 27:1 was derived for positions that were shown to be invariant through evolution. This ratio dropped to 2.6:1, 1:2.25,1:0.025 and 1:0.003 at positions including 1,2,3 and 4 substitution in the sequence alignment.

6.1 Cell Cultures

Breast cancer HCC1937 BRCA1 defective cell line (homozygous for the 5382insC mutation) was obtained from ATCC (American Type Culture Collection) and maintained in RPMI-1640 medium supplemented with 10% FBS (Fetal Bovine Serum), and L-glutamine 2 mM.

HCC1500 cell line has been a kind gift from Alfred O. Mueck (University Women's Hospital, Tuebingen, Germany). It is a breast cancer cell line established from a 32 years old female with a family history of early-onset colon cancer, as well as a sister with breast cancer. Cells were maintained in RPMI-1640 medium supplemented with 10% FBS, 2 mM L-glutamine, 1.5 g/L sodium bicarbonate, 4.5 g/L glucose, 10 mM Hepes, 1 mM sodium pyruvate, and antibiotics. Both cell lines were maintained at 37°C in a 5% CO_2 atmosphere.

6.2 Pharmacological treatment

HCC1937 cells were seeded in 6 wells plate at a density of 10^5 cells/well 24 hours before chemical treatment. For NMD pathway inhibition, emetine (100

μg/mL) (Calbiochem) was added to the culture medium. Emetine is an effective inhibitor of protein synthesis in eukaryotes, stabilizing ribosomes on polysomes. The 80S ribosome subunit-emetine complexes cannot carry out aminoacyl-tRNA binding and translocation. Untreated cells were used as control.

Inhibition of translation triggers a cellular stress response that leads to increased levels of wild type transcripts, masking the stabilization of premature terminating mRNAs. In order to reduce nonspecific drug response, after an initial 12 hours emetine treatment cells were additionally treated with actinomycin D (5 μg/mL) (Sigma). Actinomycin D is a polypeptide antibiotic which inhibits translation by forming a stable complex with double-stranded DNA, thus inhibiting DNA primed RNA synthesis. Cells were incubated with actinomycin D for 30 min, 2 hrs or 4 hrs. Control cells were incubated with the ethanol vehicle.

After a wash with PBS (Phosphate Buffered Saline), cells were lysated for RNA isolation. Lysates were stored at -80℃.

6.3 Gene silencing

Transient gene silencing was achieved through the method of small interfering RNA (siRNA), small RNA molecules (21-23 nt) that target specific transcripts to degradation.

24 hrs before transfection, 200,000 HCC1937 cells/well were seeded in a 6-well cell culture plate using cell culture medium without antibiotics. siRNA transfection was performed preparing in OptiMEM medium (Invitrogen) the proper dilution of DharmaFECT #2 transfection reagent (Dharmacon), UPF1 and UPF2 pools of four siRNAs (ON-TARGET Plus, Dharmacon), and a scramble siRNA as a control of transfected cells. After an incubation of 5 min at room temperature, siRNAs were mixed with liposomes, and the mixture was incubated at room temperature for 20 min for the formation of siRNA-liposome complexes; 400 μl/well of transfection medium was then added to cells. UPF1 and UPF2 silencing was performed with 50 μM siRNA. For simultaneous UPF1 and UPF2 knock-down cells were transfected with 50 μM si-UPF1 and 25 μM si-UPF2.

Gene silencing in HCC1500 cells was performed seeding in a 6-well cell culture plate and in antibiotic-free medium 700,000 cells/well 24 hrs before transfection. The transfection protocol is the same of the one used with HCC1937 cells, except for the transfection reagent used, lipofectamine 2000 (Invitrogen), and the transfection medium volume (500 μl/well), following liposome's manufacturer instruction. For simultaneous UPF1 and UPF2 knock-down cells were transfected with 75 μM of each siRNA. After 24 h, cell culture medium was replenished to reduce liposome-mediated cytotoxicity. To maintain a good knockdown of UPF1 and UPF2 genes up to 6 days, at 72 hrs from transfection cells were transfected again, following the same protocol.
For Western Blot and real-time PCR assessment of gene silencing efficiency, cells were washed with PBS 1X and lysated directly into the cell culture well.

7 Gel Electrophoresis and Western Blot

Cell lysates (lysis buffer: 20% SDS, 1% □-mecaptoethanol, 10% glicerol, 50 mM Tris-HCl pH 6.8, 0.05‰ bromophenol blue) were homogenized by incubating samples 10 min at 95°C and passing them through a fine gauge needle. Samples were run on 10% polyacrylamide gel at 14-25 mA (gel electrophoresis buffer: 2M glycine, 250mM Tris-HCl, 1% (w/v) SDS).
Proteins transfer was performed onto PVDF membrane (Hybond-P, Amersham Bioscience). PVDF membranes were previously equilibrated 10 sec into 100% methanol solution, 5 min into milliQ water, and finally 15 min into transfer buffer (2M glycine, 250mM Tris-HCl, 1% (w/v) SDS, 20% (v/v) methanol). Electroblotting was performed at 4℃ at 350 mA for 2.5 hrs.
PVDF membranes were subsequently dried, and then incubated 10 min in a 100% methanol solution and washed 5 min with milliQ water. In order to block nonspecific antibody binding sites, PVDF membranes were incubated 1 hr in buffer A: 2% blocking solution (Roche) diluted in PBS 1X, added with 0.05% (v/v) Tween-20. Protein staining was performed incubating membranes for 1 hr with the primary antibody anti-UPF1 (P14: sc-18260, Santa Cruz Biotechnology), anti-UPF2 (N-16: sc-20225, Santa Cruz Biotecnology), or anti-actin (20-30, Sigma) diluted in PBS 1X with addition of 3% (w/v) BSA, following manufacturer's instructions. After several washes with 0.05% (v/v) Tween-20

diluted in PBS 1X, the membranes were incubated 1 hr with the secondary antibody HRP-conjugated (anti-goat or anti-rabbit IgG, Santa Cruz Biotechnology), properly diluted in buffer A, following manufacturer's instruction. After washes to remove the secondary antibody excess, proteins detection was performed on a Hyperfilm MP photographic plate (Amersham Bioscience), using a chemoluminescent reagent (ECL Plus Western Blotting Detection Reagents, GE Healthcare)

8.1 Total RNA isolation

Total RNA was extracted using RNeasy MiniKit (Qiagen) following manufacturer's instructions. This technique is based on affinity purification, combining RNA binding properties of a silica-gel-based membrane with centrifugation, to efficiently isolate RNA molecules longer than 200 nt.

Cells were lysated into wells of cell culture plates with 350 µl of RLT buffer, containing guanidine-isothiocyanate, a chaotropic agent for RNases inhibition, supplemented with 1% (v/v) β-mercaptoethanol. Cell lysates were also homogenized through a 0.9 mm needle. When stored at -80℃, cell lysates were first incubated for 20 min at 37℃ to dissolve completely salts. To optimize RNA binding conditions to column membrane, 350 µl of 70% (v/v) ethanol was added to samples before their transfer into RNeasy column. After 15 sec of centrifugation at ≥8000 x g, 350 µl of RW1 buffer was added into column, and then removed by a 15 sec centrifugation at ≥8000 x g. DNA digestion was also carried out, incubating samples bound to column membrane 15 min at room temperature with DNase I. To remove DNase I, the silica-gel-based membrane was wash with RW1 buffer, as previously described. Other two washes were performed with 500 µl RPE buffer (an ethanol containing solution) to remove contaminants. In order to discard RPE buffer and dry the silica-gel membrane, columns were centrifuged a first time for 15 sec at ≥8000 x g, and then for 2 min at ≥8000 x g. To ensure the complete elimination of ethanol, a further centrifugation was made in a microcentrifuge at full speed for 1 min. To elute RNA, 30 µl of RNase-free water was pipetted directly onto the RNeasy silica-gel membrane and the column transferred into a new collection tube was centrifuged for 1 min at ≥8000 x g. RNA was stored at -80℃.

8.2 cDNA synthesis

RNA samples were reverse transcribed to obtain PCR templates. In a first step, 1 μg of total RNA, mixed with 3 mM dNTPs and 10 μM random hexamers, was heated for 5 min at 65℃, and, soon after, samples were incubate on ice for at least 1 min to resolve RNA secondary structures. Thereafter, RNA was reverse transcribed using 10 U/μl of SuperScript III reverse transcriptase (RT) (Invitrogen) in presence of 50 mM Tris-HCl pH 8.3, 75 mM KCl, 3 mM $MgCl_2$, 5 mM dithiothreitol (DTT), and 2 U/μl of RNaseOUT (Invitrogen), a recombinant RNase inhibitor. After 5 min incubation at 25℃, single-strand cDNA is synthesized during 1 hr incubation at 50℃. The RT enzyme was inactivated by heating samples 15 min at 70℃.

8.3 Quantitative PCR

To evaluate the degree of NMD core factors' silencing, and also to estimate the increase in PTC-containing transcript, quantitative PCR was used. Currently, real time PCR is broadly used to quantify the level of a transcript of interest. Real time PCR allows detection and measurement of the amplified product while in progress. During the PCR exponential phase, when no reaction components are limiting, the amount of PCR product approximately doubles in each cycle. Thus, a real time measurement of the amplicon amount in the exponential phase is directly correlated to the initial quantity of template present in the reaction. In the exponential phase, the cycle number at which samples reach a determined, well detectable fluorescent signal as defined by the threshold line is called threshold cycle (Ct).

Each sample was analyzed in triplicate. PCR product detection was performed with the SYBER Green method. SYBER Green is a fluorescent dye that specifically binds to double-strand DNA. PCR reactions were performed with 10 μl of Platimun SYBR-green qPCR SuperMix-UDG with ROX (Invitrogen), 0.5 μM forward and reverse primers, 3 μl H_2O. To assess the amplification efficiency for each primer pair, serial dilutions of reference cDNA (each point as a duplicate) were analyzed for all genes tested within the plate, to generate a standard curve (log of the starting quantity of template against the Ct value). Amplification was performed as follows: 50℃ for 2 min, 95℃ for 2 min, and 40

cycles of three step denaturation/annealing/extension reactions: 15 sec at 95℃, 30 sec at 60℃, 30 sec at 68℃. To assess the absen ce of nonspecific PCR products, amplification reaction was followed by a melting curve: 15 sec at 95℃, 20 sec at 60℃, and with a slower ramp rate, the temperature reached again 95℃.

Real-time PCR data were analyzed following the □□Ct method, to obtain a relative quantification that compares specific mRNA amount in test and control cells. The amplification efficiency (E) was calculated through the relation:

$E\ (\%) = (10^{-(1/S)} - 1) \cdot 100$, where S is the standard curve's slope. The real-time PCR reaction was considered successful when amplification efficiency was ≥85%. Relative quantifications were obtained normalizing target gene expression levels to average expression levels of two reference genes: lamin A/C (Fw 5'-GTACGGCTCTCATCAACTC-3', Rv 5'-AGGTCATCTCCAT CCTCATC-3') and GAPDH (Fw 5'-GAAGGTGAAGGTCGGA GT-3'; Rv 5'-CATGGGTGGAATCATATTGGAA-3'). To calculate relative gene expression, the Pfaffl method was applied, which takes into account a possible different amplification efficiency of target and reference genes. The expression ratios (R) between the test sample (silenced cells) and calibrator (control cells) were calculated using the following formula:

$R = (E_{target})^{-\square Ct,\ target\ (test-calibrator)} / (E_{reference})^{-\square Ct,\ reference\ (test-calibrator)}$.

By real-time PCR, relative expression levels of UPF1 (Fw 5'-GTGTCACTGCAGC GGATCGT-3', Rv 5'-GCCTCGGTCCTGTTCAGGTA-3'), UPF2 (Fw 5'-GCGGTGGACTTAAGCATGTA-3', Rv 5'-CACCACTTCGTTGC TCTAGA-3'), and p53 (Fw: 5'-AGTCTGTGACTTGCACGTACTC-3', Rv: 5'-GTG CTGTGACTGCTTGTAGATG- 3') were monitored.

9 Candidate Genes Selection

To reduce the number of putative genes affected by nonsense mutations, microarray data were overlapped with information that provided inclusion and exclusion criteria. Focusing our attention on bi-allelic inactivation events, a first selection was made by the identification of up-regulated genes that were located within genomic regions with LOH (Loss of Heterozygosity). The combination between microarray data on NMD sensitive genes and genomic profile information of HCC1500 cell line could highlight putative tumor

suppressor genes. According to Knudson's two hits hypothesis, a complete loss of gene function could be achieved by the truncation of one allele and the loss of the wild type one. Thus, taking advantage of the reports on *bona fide* NMD targets (Mendell *et al.*, 2004; Wittmann *et al.*, El-Bchiri *et al.*, 2008), genes already described as up-regulated after NMD inhibition were excluded. To further reduce the candidate gene number, genes within loci found to be associated to breast cancer predisposition were prioritized (Wiechec *et al.*, 2008; Argos *et al.*, 2008; Huusko *et al.*, 2004b; Bergman *et al.*, 2007; Gonzalez-Neira *et al.*, 2007; Zheng *et al.*, 2009; Gronwald *et al.*, 2005; Seitz *et al.*, 1997; Easton *et al.*, 2007; Oldenburg *et al.*, 2008; Rosa-Rosa *et al.*, 2009a; Kainu *et al.*, 2000; Thompson et al, 2002; Meguire *et al.*, 2005; Rosa-Rosa *et al.*, 2009b). Finally, genes harboring features of NMD targets, that is, genes involved in RNA processing, introns within 3'-UTR (http://genome.ucsc.edu), and genes with putative 5' uORF (http://flj.hinv.jp/ATGpr/atgpr/index.html) were excluded.
Due to the strong relation between DNA repair processes and breast cancer susceptibility, among genes up-regulated after NMD inhibition, those involved in DNA repair were chosen for full sequencing.

10.1 Direct Sequencing

Genomic DNA (50-25 ng) was PCR amplified with 0.05 U/ml of AmpliTaq Gold DNA polymerase (Applied Biosystems) in the presence of PCR buffer 1X, 1.5-2.5 mM $MgCl_2$, 480 μM dNTPs, 0.6 μM for each M13-tailed primer. The larger part of primer sequences to amplify the entire coding sequence of candidate genes was obtained from Probe database (http://www.ncbi.nlm.nih.gov/probe/). Standard PCR condition were: 8 min at 95℃, followe d by 40 cycles of 30 sec at 95℃, 40 sec at melting temperature of primer pair, 40 sec at 72℃. PCR reaction ended with 10 min at 72℃.
To perform DNA sequencing, PCR products were subsequently cleaned-up to remove excess of primers and unused nucleotides using an enzymatic method. 5 μl of PCR product were mixed with 2 μl of ExoSAP-IT (USB corporation), which contains Exonuclease I and Shrimp Alkaline Phosphatase enzymes for degradation of unused PCR reaction components. Samples were incubated 15 min at 37℃; then ExoSAP-IT inactivation was achiev ed with 15 min sample incubation at 80℃.

Sequencing reactions were performed with BigDye terminator v1.1 Cycle Sequencing Kit (Applied Biosystems). 1 µl of purified PCR products properly diluted (1-5 ng/µl) was mixed with sequencing buffer at the final concentration 1X, BigDye Terminator solution (containing dNTPs, fluorophore-conjugated ddNTPs, and Taq polymerase) at 1:10 final dilution, 0.16 µM M13 universal primer (M13_Fw: 5' TGTAAAACGACGGCCAGT 3'; M13_Rv: 5' CAGGAAACAG CTATGACC 3'), or amplicon-specific primer. Sequencing reaction was performed as follows: 1 min at 96℃, and 25 cycles of 10 sec at 96℃, 4 min at 60℃.

Fluorescent-labeled sequences were then cleaned-up to remove unincorporated dye terminators, dNTPs, and free salts using BigDye XTerminator Purification Kit (Applied Biosystems) following manufacturer's instructions. Within the reaction plate, 10 µl of sequencing reaction products were added to the XTerminator solution: 45 µL SAM solution and 10 µL XTerminator solution containing beads that sequester unused cycle-sequencing components. The reaction plate was subsequently vortexed at 2000 rpm for 30 min and then centrifuged at 1000 x g for 2 min. Samples were analyzed by capillary electrophoresis using ABI-PRISM 3130xl genetic analyzer (Applied Biosystems).

10.2 Sequences Analysis

Sample sequences were automatically aligned with GeneBank reference genomic sequences (build 37, assembly GRCh37), using SeqScape v2.5 software (Applied Biosystems). Exonic sequence variants were then searched in SNP database (http://www.ncbi.nlm.nih.gov/snp) to verify if they were already known polymorphisms. Intronic sequence variants located less than 30 bp from exon boundaries were *in silico* evaluated for their effect on splicing process (http://www.fruitfly.org/seq_tools/splice.html). I*n silico* predictions of functional effects of not previously reported sequence variants were performed by Polyphen prediction tool (http://genetics.bwh.harvard.edu/pph/).

11.1 Microarray Gene Expression Profiling

Genome-wide expression profiling was performed with Human-Genome U133 Plus 2.0 platform (Affymetrix). This array allows the measurement of expression level of over 47,000 transcripts, interrogating each gene with multiple independent probes.

Three biological replicates of paired UPF1 and UPF2 silenced and control cells at 120 hrs from transfection were lysated and total RNA isolated with RNeasy Mini Kit (Qiagen). RNA was then quantified through determination of its absorbance at 260 nm. Absorbance ratio 260nm/280nm was also evaluated. Integrity of RNA samples was assessed by Agilent 2100 Bioanalyzer.

Samples were carefully processed following the one-cycle target labeling protocol provided by the manufacturer (Figure 7). Initially, 4 μg of RNA were retrotranscribed using One-Cycle cDNA Synthesis Kit (Invitrogen). RNAs mixed with 8.3 μM T7-oligo(dT) primer (5' – GGCCAGTGAATTGTAATACGACTCA CTATAGGGAGGCGG – (dT)24 – 3') and a pool of properly diluted eukaryotic polyA RNA controls, as exogenous positive controls (Bacillus transcripts dap, lys, phe, thr, trp modified by the addition of a Poly(A) tail) to monitor the entire target labeling process. Samples were incubated at 70℃ for 10 min, followed by at least 2 min incubation at 4℃. The first-strand cDNA synthesis reaction was performed adding to samples 1^{st} strand reaction mix at a final concentration 1X, 10 mM DTT, and 0.5 mM dNTPs. Following 2 min pre-heating at 42℃, 1 μl SuperScript II reverse transcriptase enzyme was added to the mixes. Samples were incubated 1 hr at 42℃, and then cooled at 4℃ for at least two min to avoid the inactivation of the enzyme needed for the second-strand cDNA synthesis. Double-stranded cDNA synthesis was performed adding to single-strand cDNA samples 2^{nd} strand reaction mix at 1X final concentration, 0.2 mM dNTPs, 1 μl RNase H, 4 μl *E. coli* DNA polymerase I, and 1 μl *E. coli* DNA ligase. Samples were incubated 2 hrs at 16℃, followed by 5 min incubation at 16℃ after the addition of 2 μl of T4 DNA polymerase. Enzyme inactivation was achieved by the addition of 40 mM EDTA.

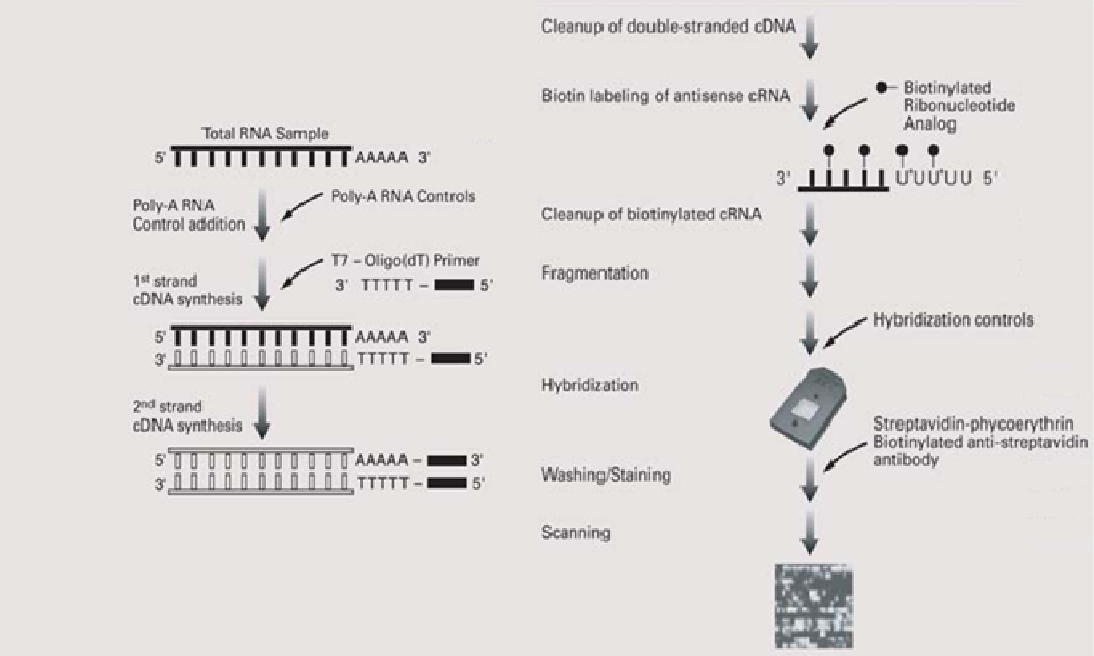

Figure 7. Sample processing for microarray gene expression analysis. Total RNA is first reverse transcribed using a T7-oligo(dT) primer in the first-strand cDNA synthesis reaction. Following RNase H-mediated second-strand cDNA synthesis, the double-stranded cDNA is purified and serves as a template in the subsequent *in vitro* transcription reaction (IVT). The IVT reaction is carried out in the presence of T7 RNA Polymerase and a biotinylated nucleotide analog/ ribonucleotide mix for cRNA amplification and biotin labeling. The biotinylated cRNA targets are then cleaned, fragmented, and hybridized to GeneChip expression arrays.

Double stranded cDNA samples were subsequently cleaned-up with Sample Cleanup Module (Qiagen): 600 µl of cDNA Binding Buffer was added to cDNA and vortexed 3 sec. 500 µl of cDNA was transferred onto a cDNA Cleanup Spin Column and centrifuge 1 min at ≥8,000 x g. After discarding the flow-through and repeating a second time these initial steps to collect within the column all the synthesized cDNA, a first wash step was performed with 750 µl of cDNA Wash Buffer. After buffer addition, columns were centrifuged 1 min at ≥8,000 x g and, to ensure complete ethanol removal, a further centrifugation for 5 min at ≤25,000 x g was performed. 14 µl of cDNA Elution Buffer were pipetted directly onto spin columns membrane and, after 1 min of incubation at room temperature, cDNA was eluted into a new collection tube by 1 min centrifugation at ≤25,000 x g.

cDNA samples were subsequently processed with GeneChip IVT Labelig Kit (Invitrogen) for in vitro transcription (IVT), to amplify and, at the same time, label the initial population of RNA molecules. At room temperature, 12 μl of purified cDNA samples were mixed with 8 μl of RNase-free water, 4 μl 10X IVT labeling buffer, 12 μl IVT labeling NTP mix with biotinylated dNTPs, and 4 μl IVT labeling enzyme mix, which contains T7 RNA polymerase. To obtain biotin-labeled cRNA (complementary RNA), samples were incubated 16 hrs at 37℃.

Biotinylated cRNA samples were then purified with Sample Cleanup Module (Qiagen). 60 μl of RNase-free water were added to biotinylated cRNA and vortexed for 3 sec; then 350 μl IVT cRNA binding buffer was added and samples again vortexed for 3 sec. After addition of 250 μl of ethanol 96-100%, samples were transferred onto cRNA cleanup spin columns and were centrifuged for 15 sec at ≥8,000 x g. A wash step was performed pipetting 500 μl of IVT cRNA wash buffer onto spin columns and centrifuging for 15 sec at ≥8,000 x g. Discarded the flow-through, 500 μl of ethanol 80% (v/v) were added onto spin columns, then removed by 15 sec centrifugation at ≥8,000 x g. To ensure a complete ethanol removal, a further 5 min centrifugation at ≤25,000 x g was performed. Finally, purified biotinylated cRNA samples were eluted pipetting directly onto spin column membranes 11 μl of RNase-free water and centrifuging 1 min at ≤25,000 x g. This elution step was repeated once with 10 μl of RNase-free water. Biotin-labeled cRNAs were quantified determining samples absorbance at 260 nm, and cRNA quality was assessed by Agilent 2100 Bioanalyzer.

Before hybridization onto GeneChip probe array, biotin-labeled cRNA was fragmented: 20 μg of cRNA were mixed with fragmentation buffer and samples were incubated 35 min at 94℃. cRNA fragments (with in the range size of 35-200 bp) were qualitatively analyzed by Agilent 2100 Bioanalyzer.

At the end of the procedure, samples were processed for hybridization: 0.05 μg/μL of fragmented biotin-label cRNA were mixed with 50 pM control oligonucleotide B2, 1X eukaryotic hybridization control (*E. coli* transcripts bioB, bioC, bioD, and cre, at 1.5 pM, 5 pM, 25 pM, and 100 pM, respectively), 0.3 mg/mL herring sperm DNA, 0.5 mg/mL BSA, 10% DMSO, and 1X hybridization buffer. Hybridization cocktails were incubated 5 min at 65℃, 5 min at 98℃, and

5 min at 45℃, and then centrifuged 10 min at ≤25,000 x g. Meanwhile, GeneChip arrays were filled with 1X pre-hybridization buffer and were kept in the hybridization oven 640 (Affymetrix) for 10 min at 45℃, with rotation. After removal of pre-hybridization buffer, arrays were re-filled with 200 µL of clarified hybridization cocktail and were placed into the hybridization oven at 45℃, with rotation at 60 rpm for 16 hrs.

Standard washing and staining procedures were performed with the Fluidics Station 450 (Affymetrix) and with the assistance of the Affymetrix GeneChip Command Console Software. The amplification of biotinylated target signal was then achieved with subsequent cycles of incubation with streptavidin conjugated with the fluorescent dye phycoerythrin, biotinylated antibody, and streptavidin-phycoerythrin. GeneChip probe arrays were scanned using GeneChip Scanner 3000 (Affymetrix).

11.2 Microarray Data Analysis

Quality controls, which assess the reliability of RNA samples and of the intermediate stages of sample preparation and hybridization, were performed with Simpleaffy package (BioConductor). Data normalization was performed using the RMA algorithm (Robust Multi-array Analysis), which corrects background, normalizes, and summarizes probe level information.

To detect differential gene expression, microarray data were analyzed with SAM algorithm (Significance Analysis of Microarray), with an FDR (False Discovery Rate) of 1% that means setting the number of false positives ≤ 1% among genes identified as differential expressed between treated and control sample. Finally, reconstruction of genes expression levels was performed using a set of custom Chip Definition Files (CDF) that are composed of a unique probeset for each gene, containing only highly specific probes (Ferrari *et al.*, 2007).

12.1 Copy Number and LOH Analysis

Copy number and LOH analysis was performed using Genome-Wide Human SNP Array 6.0 (Affymetrix). This array platform contains about 1,850,000 SNP and non-SNP oligonucleotide probes, which allow an average resolution of 700 bp. To prepare genomic DNA for hybridization on chip, the Genome-Wide

Human SNP Nps/Sty Assay Kit 5.0/6.0 (Affymetrix) was used, carefully following manufacturer's instructions (Figure 8).

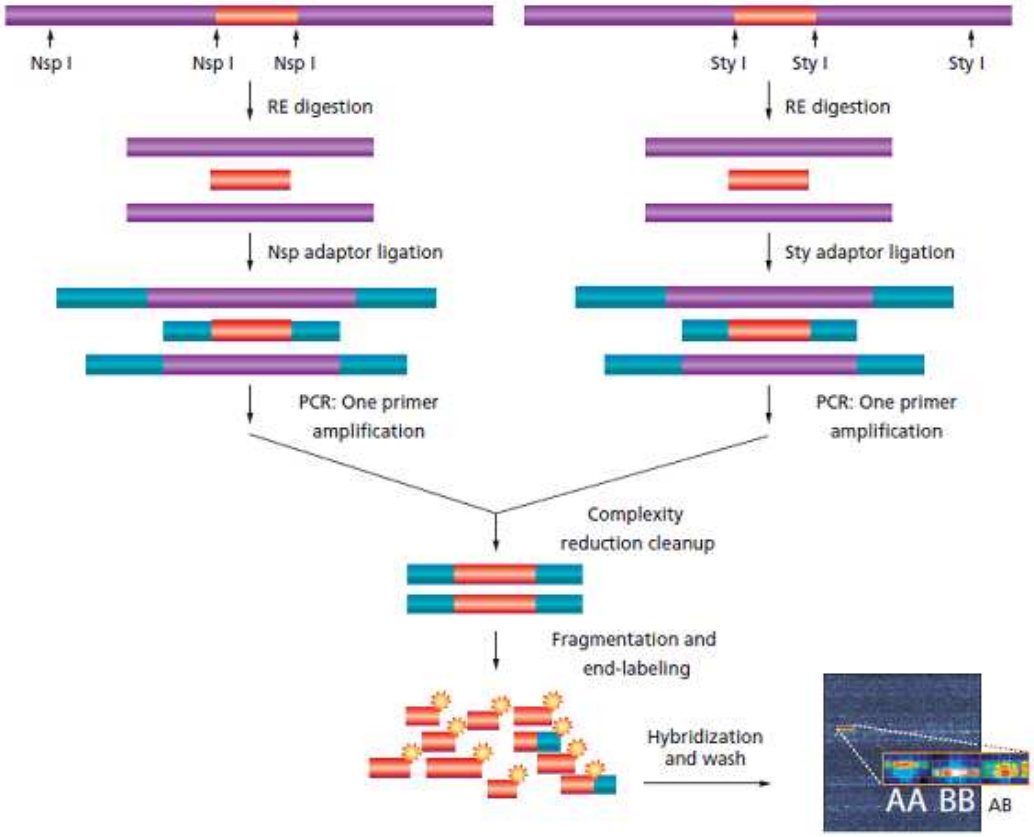

Figure 8. Sample processing for copy number and LOH analysis. Genomic DNA is digested with two restriction enzymes and ligated to adaptors that recognize the cohesive 4 bp overhangs. Universal primer pairs that recognizes the adaptor sequences are used to amplify adaptor-ligated DNA fragments. PCR conditions have been optimized to preferentially amplify fragments in the 200 to 1,100 bp size range. PCR amplification products for each restriction enzyme digest are combined and purified using polystyrene beads. The amplified DNA is then fragmented, labeled, and hybridized to a SNP Array 6.0.

Two aliquots of 250 ng each of HCC1500 genomic DNA were digested with 0.5 U/µl NspI and StyI restriction enzymes, respectively (digestion master mix: NE buffer 1X, BSA 2 µg). Together with test DNA, also a negative (H_2O) and a positive control (Ref103) were processed. Digestions were performed at 37°C for 2 hrs, followed by 20 min at 65°C.

Genomic DNA was then ligated to adaptors that recognize the cohesive 4 bp overhangs. The two ligation master mixes were set up with 1.5 μM NspI or StyI adaptors, 32 U/μl T4 DNA ligase, and T4 DNA ligase buffer 1X. Ligation reactions were performed at 16°C for 3 hrs, followed by 20 min at 70°C. To obtain the proper template concentration for the following PCR reactions, ligated samples were diluted 1:4.

10 μl of NspI and StyI ligated samples were used to PCR amplify genomic DNA; four and three PCR reactions were set up for NpsI and StyI digested DNA, respectively. Each reaction contains TITANIUM Taq PCR buffer 1X, 1 M GC-Melt, 1.4 mM dNTPs, 4.5 μM primer, and TITANIUM Taq DNA polymerase 1X. PCR reactions were performed as follows: 3 min at 94°C, 30 cycles of denaturation at 94°C for 30 sec, annealing at 60°C for 45 sec, extension at 68°C for 15 sec; and finally 7 min at 68°C. To verify the positive result of PCR reactions, which would produce a pool fo amplified genomic fragment within 200-1100 bp size range, amplified samples were run on a 2% agarose gel.

Subsequently, the seven NspI and StyI PCR reactions were pooled together within a unique tube and purified to remove unused primers and dNTPs by magnetic beads. 1 ml of magnetic beads solution was added to pooled sample, and, after well mixing, it was incubated 10 min at room temperature, for DNA binding to the magnetic beads. After 3 min centrifugation at maximum speed in a microcentrifuge, tubes were placed on a magnetic stand, which holds the bead pellet at the bottom of the tube. Once removed the supernatant, 75% ethanol was added and the samples vortexed for 2 min. To complete this wash step, samples were centrifuged 3 min at maximum speed and tubes paced on the magnetic stand to remove the supernatant. This procedure was repeated twice to ensure a complete ethanol removal. Finally, 55 μl of EB elution buffer were added and samples vortexed for 10 min, to ensure a homogeneous bead resuspension and thus an effective DNA release. After 5 min centrifugation at maximum speed, tubes were place on the magnetic stand for at least 5 min, and 47 μl of supernatants transferred in new tubes. PCR purified products were quantified at 1:100 dilution using spectrophotometer. Adequate DNA yield ranges between 4.5 and 7 μg/μl.

Then, the amplified DNA, after adjusting its volume to 50 µl by adding the 10X fragmentation buffer, was fragmented: working rapidly on ice to avoid loss of enzyme activity, 0.5 U/ml of fragmentation enzyme (DNase I) and fragmentation buffer 1X were added to the samples. DNA fragmentation take place during 35 min incubation at 37°C, followed by enzyme inactivation step at 95°C for 15 min. The fragmentation reaction, which should lead to an average fragment size <180 bp, was checked on 4% agarose gel.

At the end of genomic DNA processing for hybridization, fragmented samples were labeled using the activity of Terminal Deoxynucleotidyl Transferase (TdT), which catalyses, without a template, the addition of nucleotides to the 3' terminus of DNA molecules. Labeling reaction was performed adding to fragmented sample TdT buffer at the final concentration of 1X, 0.83 mM DNA labeling reagent (biotin-labeled ddNTPs), and 1.5 U/µl TdT enzyme. Labeling nucleotides addition was been achieved by 4 hrs incubation at 37°C, following by an inactivation step at 95°C for 15 min.

Finally, sample was hybridized on the array chip. To this aim, a hybridization master mix was prepared: 80 mM MES (2-(N-Morpholino)ethanesulfonic acid hydrate, 4-Morpholineethanesulfonic acid; $C_6H_{13}NO_4S$), 3.4X Denhardt's solution, that is a mixture of high-molecular weight polymers capable of saturating non-specific binding sites and artificially increasing the concentration of available probe, 8 mM EDTA, 0.16 mg/mL herring sperm DNA, oligo control reagent diluted 1:95, 16 µg/mL Human Cot-1 DNA, 0.016% Tween-20, 6.8% DMSO, 3.7 M TMACL (tetramethylammonium chloride). 190 µl of hybridization master mix were added; the samples were denatured at 95°C for 10 min, and then manteined at 49°C. 200 µl of hybridization cocktail were injected into the array. Hybridization took place in a hybridization oven (GeneChip Hybridization Oven 645), with the array chips rotating at 60 rpm and 50°C for 16 to 18 hrs.

After hybridization, arrays were washed and stained following manufacturer's instructions. After removal of hybridization cocktail arrays were filled with 270 µl of array holding buffer (100 mM MES, 1 M NaCl, 0.01% Tween-20), and inserted into one module of GeneChip Fluidics Station 450 (Affymetrix). The standard protocol for array wash and staining was used (GenomeWideSNP6_450). Chip arrays were subjected to subsequent cycles of

washing, staining with 10 ng/mL streptavidin molecules conjugated with phycoerythrin (SAPE), post staining washes, staining with 5 ng/mL biotinylated antibody, a further staining step with SAPE solution, and finally a washing step and complete filling of the array with array holding buffer. Genome-Wide Human SNP 6.0 Array were scanned using GeneChip Scanner 3000 (Affymetrix).

12.2 SNP Array Data Analysis

The genome-Wide Human SNP 6.0 Array file was analyzed using the Genotyping Console (GTC) 3.0.1 (Affymetrix). Genomic profiles of analyzed sample were compared to the reference model file, supplied by Affymetrix, that was generated upon the analysis of 270 HapMap individuals. Affymetrix software allows a simple graphical overview of i) allele difference, that is the difference between signals of the two possible genotypes each standardized with respect to their median values in the reference dataset, ii) copy number state (CN 0: homozygous deletion; CN 1: heterozygous deletion; CN 2: normal diploid; CN 3: single copy gain; CN 4: amplification with copy number ≥4), determined using the HMM model (Hidden Markov Model), iii) LOH, iv) log2 ratio value, and v) smoothed log2 ratio. Moreover, for each genomic locus GTC 3.0.1 visualizes reference genes. GTC 3.0.1 segment reporting tool (SRT) parameters were set to identified copy number alterations greater than 500 Kb and that contain at least 450 markers. Considering that the average spacing of the targets of this array is 0.7 Kb, the smallest genomic altered region is thus expect to be supported by 714 SNP markers.

Copy number data were also analyzed with Partek Genomic Suite Software, considering significant a copy number alteration detected in a genomic segment containing at least 30 markers.

Aim of the study

The aim of the projects described in this thesis deals with the identification of genes and variants that might explain the genetic factors that underline tumor occurring in families at high risk for breast and ovarian cancer but without detectable pathogenetic alterations in the BRCA1 or BRCA2 genes.

The first project focuses on the characterization of a BRCA1 unclassified variant that recurs in Veneto and that cannot be used for the clinical genetic testing of family members (i.e. for their inclusion or exclusion from surveillance programs) because of its unclear on the protein coded. So far, many types of studies have been used to classified BRCA1 and BRCA2 variants, including functional, genetic-epidemiological and *in silico* analyses, among others. In most cases, however, none of these methods provide indications strong enough to unambiguously classify the variant. We therefore aimed at integrating as many different evidences as possible, to obtain a final response that could support clinical decisions for individuals at high risk of breast cancer and ovarian cancer. We took advantage of a method recently described that maximize the power of this kind of studies and could likely be extended to the analysis of the other genetic disease.

The purpose of the second project relies on the identification of genetic alterations that might be relevant for tumor predisposition and/or development in patients without highly penetrant mutations in the major breast cancer predisposing genes. To this purpose, we decided to perform a comprehensive genomic analysis in the search of point mutation as well as major genomic rearrangements in a breast cancer cell line selected for a probability of being derived from a hereditary breast cancer. We employed a multi-strategy approach based on the GINI (Gene Identification by NMD Inhibition) method as well as high-throughput technologies for the sensitive detection of chromosomal imbalances.

Project I:
Results & Discussion

Results

1. Integrated Evaluation of the BRCA1 p.Val1688del variant

In the last ten years, nine families carrying BRCA1 p.Val1688del sequence variant were recruited by the Hereditary Breast/Ovarian Cancer Center at Istituto Oncologico Veneto .Three additional families with the same BRCA1 variant were identified by the Istituto Nazionale Tumori (INT, Milan, Italy) and the Istituto Oncologico Europeo (IEO, Milan, Italy).

All families were from the Veneto Region, but, we were unable to connect any of the pedigrees despite exhaustive family history on three or four generation. The presence of a common ancestor for a putative p.Val1688del founder mutation effect in the Veneto region was investigated through genotyping of microsatellite markers spanning the BRCA1 locus. This analysis identified a common haplotype shared by all index cases, thus confirming a common origin of this variant.

To rule out the possibility that p.Val1688del represented a common polymorphic variant, this three nucleotides deletion (5181-1583 of exon 17 of BRCA1) was analyzed in 200 healthy individuals from Veneto. Mutational screening did not reveal any p.Val1688del in this geographically matched control population.

In the attempt to investigate the pathogenetic role of the p.Val1688del BRCA1 variant, different type of evidences, deriving from segregation, tumor pathology, evolutionary conservation, and epidemiological data, were integrated in a multifactorial model proposed by Goldgar and co-workers (2004). The aim of this approach is to combine the odds of causality of the variant derived from multiple independent sources. This model properly weighted different evidences related to the variant under study to finally determine the ratio:

$$\frac{\text{likelihood of the observed data under the hypothesis of causality}}{\text{likelihood of the observed data under the hypothesis of neutrality}}$$

1.1 Co-occurrence

Clinical data of the 12 index cases (Table 1) were similar to those usually associated to BRCA1 gene defects. In fact, i) four of the 12 patients had bilateral breast cancer, ii) three probands developed both breast and ovarian cancer iii) and, overall, patients were characterized by an early age at cancer diagnosis (mean values 39.4 and 50.3 years for breast and ovarian cancer, respectively). In agreement with these data, the BRCAPRO analysis of *a priori* probability of BRCA1/BRCA2 mutation scored higher than 50% in all but one patient (Table 1). Despite exhaustive BRCA1/BRCA2 mutational screening of all the samples, the BRCA1 p.Val1688del was the only detectable anomaly. Supporting this data, also four additional p.Val1688del carriers were reported by Myriad Genetics at the BIC database (Breast Cancer Information Core, http://research.nhgri.nih.gov/bic) without any other detectable BRCA1/BRCA2 deleterious mutation. Due to the inferred lethality of BRCA1 homozygotes, the absence of co-occurence of p.Val1688del with other BRCA1/BRCA2 deleterious mutations is slightly in favor of causality by 1.85:1 (Goldgar *et al*., 2004).

Table 1. Phenotypic characteristics of 12 families carrying the p.Val1688del. (Malacrida *et al*., 2008)

			Family								
			Proband Age (years)		Breast Cancer Patients		Ovarian Cancers Patients				
Proband	Family	Tumor Type	Breast Cancer	Ovarian Cancer	No.	Age (years)	No.	Age (years)	Breast/ Ovary Ratio	Other Tumors	BRCAPRO
B027	14	Ovary	—	53	5	48, 41/49, 47, 44, 46	1	57	6/2	—	0.927
B083	61	Bilateral breast	27/40	—	3	35, 35, 30	—		5/0	Colon (n = 1)	0.997
B105	78	Bilateral breast/ovary	44/50	53	—		—		2/1	Lung (n = 2), prostate (n = 2), gastric (n = 2)	0.889
B133	100	Breast	35	—	4	45, 55, 45, 42	—		5/0	—	0.822
B322	232	Breast	68	PO	2	29, 58	—		3/0	—	0.093
B417	305	Bilateral breast/ovary	31/45	53	1	40	1	52	3/2	Lung (n = 4), prostate (n = 1), osteosarcoma (n = 1), gastric (n = 1)	0.999
B456	342	Breast	31	—	3	50, 35, 47	—		4/0	Thyroid (n = 1)	0.947
B656	500	Ovary	—	52	1	42	2	52, 58	1/3	—	0.957
B834	624	Ovary	—	43	—		1	45	0/2	Colon (n = 1)	0.503
MON6	MON	Breast	43	Prophylactic oophorectomy	2	58, 61/80	1	61	4/1	NHL (n = 1), gastric (n = 1), uterus (n = 1), peritoneal (n = 1), colon (n = 1)	0.517
BOT	BOT	Breast	39	—	—		2	54, 54	1/2	—	0.866
B002	B002	Bilateral breast/ovary	37/37	48	3	63, 50, 46	—		5/1	—	1.000
Mean values			39.4*	50.3	1	45.5*	54.1		2.8: 1		

Abbreviation: NHL, non-Hodgkin's lymphoma.
*Age at first tumor in bilateral cases was used for mean age calculation.

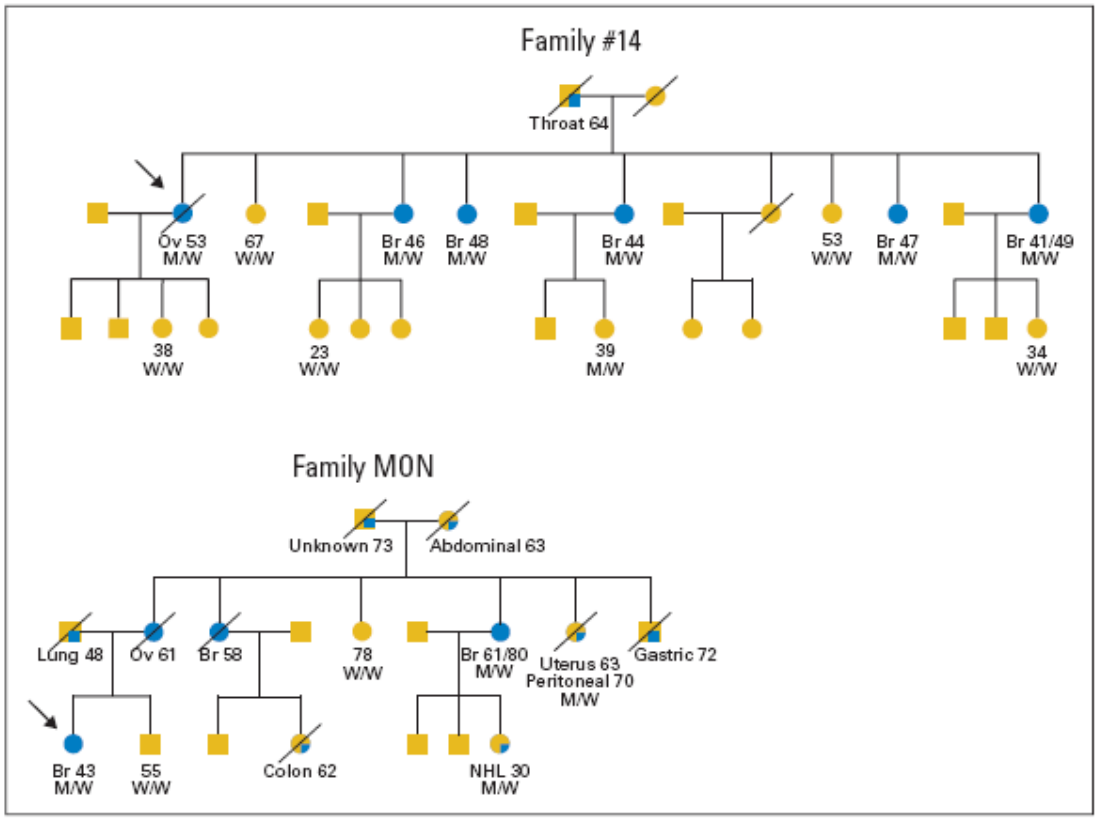

Figure 9. Pedegrees of two families carrying the p.Val1688del. The presence (M) or absence (W) of p.Val1688del is indicated under each tested individual. Numbers refer to current age and age at diagnosis for healthy and affected subjects, respectively. Probands are marked by arrows (from Malacrida *et al*., 2008).

1.2 Co-segregation

To assess co-segregation of the p.Val1688del with disease, 16 additional family members were investigated from two of the larger families. As shown in figure 9, six of six affected sisters in family #14 were carriers of p.Val1688del, whereas two healthy subjects did not inherit the variant. Similarly, in family MON all affected members analyzed carried the deletion, including a maternal aunt who developed cervical cancer at age 63 years and a peritoneal cancer 7 years after removal of the ovaries, in line with data suggesting a substantial residual risk for peritoneal cancer even after salpingo-oophorectomy (Finch *et al*., 2006). The mutant allele was also inherited by a proband's cousin affected by a non-Hodgkin's lymphoma at age 30, thus suggesting a nearly complete co-segregation. Assuming an allele frequency of 0.0001 and a 75% penetrance, and considering the two families as independent, co-segregation analysis was

performed with the Linkage package software. From the combined values of lod-score (log_{10} odds) of the two families, a odds ratio of 661:1 in favor of causality we obtained.

1.3 LOH at BRCA1 Locus

To investigate somatic LOH at the BRCA1 locus, DHPLC-based identification of the *wild type* and p.Val1688del allele was performed in two tumor samples, one breast and one ovarian cancer. In contrast to a 30% expected deletion frequency, randomly affecting either one of the two alleles in non-BRCA tumors (Osorio *et al*., 2007), both sample analyzed showed loss of the *wild type* allele. This data provided a causative odds for p.Val1688del of 27.8:1.

1.4 Tumor Histopathology of p.Val1688del Carriers

Histopatologic reports of five breast cancers (one medullary, one tubular, two ductal, and one of unknown histotype) from affected carriers disclosed features characteristic of BRCA1 tumors. Indeed, four out of five samples showed high grade (G3) and negativity for estrogen and progesterone receptors. These qualitative evidences further supported a pathogenic role of pVal1688del variant in a causative odds ratio of 3.8:1.

1.5 Phylogenetic Conservation of BRCA1 Valine1688 Residue

The deletion of the GTT that causes p.Val1688del falls within a trinucleotide repeat. According to consensus guidelines on mutation nomenclature, the sequence variant is named from the most 3' GTT repetition. However, deletion of Valine1687 or Valine1688 are synonymous events for the impairment of protein function. In fact, both these amino acid residues fall within a well conserved region of the BRCT domain.

To investigate the severity of the amino acid deletion, we used a multiple alignment of 12 full-length BRCA1 sequences from different species with a probability of pathogenecity derived from the number of substitutions occurring during evolution in these positions, as previously determined (Tavtigian *et al*., 2006). Valine1688 is substituted by an isoleucine in half of the 12 BRCA1

orthologs, whereas Valine1687 is one of the 130 invariant positions out of 1863 amino acids throughout the 12 species considered. This evidence suggests that Val1687 is under a strong functional constrain. Considering the constrain position likelihood as a predictor of variant pathogenicity, two scores of 27 and 2.7 were obtained for Val1687 and Val1688, respectively. In order to reduce the risk of false positive results, the lowest score (odds ratio 2.7:1) was used for the subsequent integration with the odds obtained from the other independent evidences.

1.6 Integration of Independent Odds Ratio for p.Val1688del Classification

Considering all the evidences obtained, we derived a combined odd of causality under the assumption that all data sources are independent of each other, accordingly to the approach proposed by Goldgar and co-workers. Under this model, the likelihood ratios obtained were multiplied to obtain a final score of 349,000:1 (1.85 x 661 x 3.8 x 27.8 x 2.7) in favor of causality, thus convincingly exceeding the minimal ratio of 1,000:1 for classification of a variant as pathogenetic.

Discussion

Even though great efforts have been made since BRCA1 and BRCA2 discovery, the genetic factors involved in breast cancer development in the majority of individuals who show positive family history for the disease, or have other features strongly suggestive of predisposing germline alterations, remain largely unknown.

A relevant fraction of families can likely be attributed to BRCA1 and BRCA2 variants that, because of their non obvious effect on the coded protein, cannot be used for identification of predisposed individuals within families. BRCA1 and BRCA2 unclassified variants (UVs) account for a large number of gene sequence changes and can represent the result of up to 13% of the genetic tests (Tavtigian *et al*., 2008). Therefore, the understanding of UVs pathogenetic role is highly valuable with respect to all high-risk individuals that could take advantage from a proper clinical management.

Among our cohort of patients, BRCA1 p.Val1688del variant occurs in 1.5% of families and was identified in a total of 12 families from Northern Italy. Haplotype analysis confirmed that this variant has a common founder, whose descendants live in a geographical region located among Padua, Vicenza, and Rovigo (even if none of the p.Val1688del carriers were from any of these cities).

Using a number of independent evidences, we showed that p.Val1688del is a pathogenetic mutation affecting the BRCA1 function, likely by altering the functionality of the BRCT domain. Indeed, Val1687 and Val1688, both of which can be considered the target of the deletion, lie in a highly conserved region of the BRCT domain. Accordingly, one of the first studies based on the intrinsic transactivation activity of BRCT domain reported an impaired function for mutants carrying the p.Val688del: this BRCA1 allele was unable to transactivate the transcription of different target genes when fused to GAL4 or LexA DNA-binding domains in yeast and mammalian-based assays (Vallon-Christersson *et al*., 2001). In addition, very recently De Nicolo and co-workers (2009) published a comprehensive functional characterization of BRCA1 p.Val1688del mutation. They confirmed p.Val1688del pathogenic effect and showed that this sequence

alteration profoundly destabilizes the BRCT hydrophobic core and compromises protein stability and function.

The clinical data of the 12 families analyzed in our study are consistent with a highly penetrant pathogenetic effect of p.Val1688del. Tumors phenotypic data highly resembled those of BRCA1-associated cancers, and ages at disease diagnosis in all affected carriers were not statistically different from those calculated from Italian carriers of pathogenic BRCA1 mutations (Aretini *et al*., 2003). Moreover, p.Val1688del showed a nearly complete co-segregation with the disease.

To obtain a final assessment of p.Val1688del pathogenicity, we used as many different approaches as possible. Notably, most of the approaches employed, such as co-occurrence with pathogenetic mutations, LOH, histopathology data, and conservation analysis are much more powerful to obtain evidence in favor of neutrality rather than causality. Therefore, in general, the higher the number of independent approaches employed the more reliable are the results of the analysis.

Using the approach described by Goldgar and co-workers (2004), we were able to integrate the results from five independent evidences. Under this model, we reached a highly significant result that, as such, can be readily translated in an unequivocal pathogenetic role of BRCA1 p.Val1688del.

If we consider the whole cohort of our patients with BRCA1 pathogenic mutations, p.Val1688del represents the most frequent mutation in Northeast Italy. This will allow implementation of pre-screen tests, at least in those families coming from specific geographical subregions. p.Val1688 accounts for 15% of our families with BRCA1 point mutation or few nucleotide deletions or insertions. Therefore, a single variant accounts for a relevant proportion of informative families comparable, for instance, with the one of major genomic rearrangements in the same gene (Agata *et al*., 2006). Thanks to this study, the genetic test of p.Val1688del can now be offered to probands' family members to discriminate subjects at increased risk for the disease, who can be enrolled in the surveillance programs from non-carriers who can be spared anxiety and unnecessary medical interventions. These data should prompt further research in the field of BRCA1 and BRCA2 variants characterization of and can

contribute to fill in the gap between families predicted to carry pathogenetic mutations and those in which a molecular defect can actually be demonstrated.

Project II:

Results & Discussion

Results

2. Genomic Characterization of the Hereditary Breast Cancer Cell Line HCC1500

In the attempt to shed some light upon molecular alterations that affect hereditary non-BRCA1/BRCA2 breast cancer, we performed a comprehensive genomic analysis of the HCC1500 cell line, which was established from a breast cancer occurred at age 32 in a patient with a family history of breast and colon cancer.

After the exclusion of mutations in either of the major predisposing breast cancer genes, we used two analytical strategies: First, we optimized a recently described approach, named GINI (Gene Identification by NMD inhibition) (Noesie and Dietz, 2004), to identify nonsense mutations affecting putative tumor suppressor genes. Secondly, we performed a profile of the genomic alterations by high resolution copy number and LOH analysis.

2.1 A Revised Strategy for Nonsense Mutation Detection

2.1.1 Evaluation of the GINI Method in the HCC1937 Cell Line

To evaluate the efficacy of the recently described GINI approach (Gene Identification by NMD Inhibition), the HCC1937 cell line was pharmacologically treated to inhibit the NMD pathway (Noensie and Dietz, 2001; Ionov *et al.*, 2004). By blocking the cellular translational machinery with emetine, in the absence of ribosome recognition of nonsense codons, UPF1 is not recruited onto termination complex and the mRNA quality control does not occur.

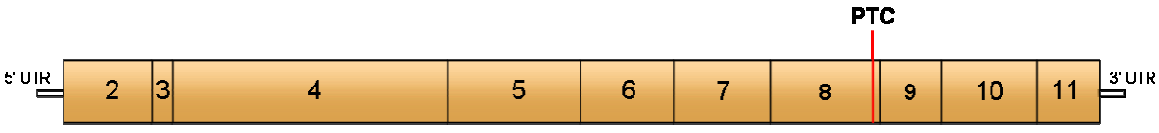

Figure 10. Nonsense TP53 mRNA. HCC1937 cells have lost the TP53 wild type allele while the remaining allele harbors the p.R306X nonsense mutation recognizable by the NMD pathway. PTC: premature termination codon.

HCC1937 cells have only one TP53 allele harboring p.R306X truncating mutation (c.916 C>T) (Figure 10). Thus, the TP53 gene was used as an endogenous control for the evaluation of the stabilization of premature terminating transcripts after NMD inhibition. HCC1937 cells were treated with emetine alone, blocking translation, or with emetine and actinomycin D, leading to a sequentially impairment of both translation and transcription. As shown in figure 11, pharmacologically treated cells showed a high degree of stabilization of truncated TP53 transcript compared with the untreated counterpart. Indeed, emetine treatment alone increased up to 9-fold the presence of TP53 mutant transcript. The combination of emetine with actinomicyn D, which was reported to reduce the nonspecific increase of stress-induced cellular transcripts, affected only slightly the accumulation of the reporter mRNA, leading to a 9-fold increase of the level of premature terminating TP53 transcript (Figure 11).

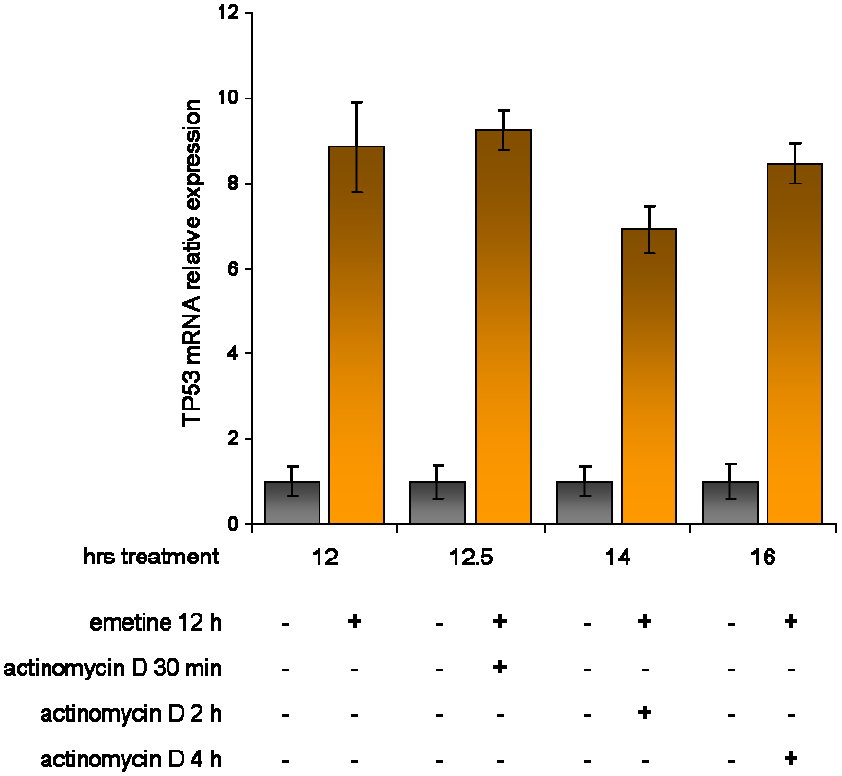

Figure 11. Real-time PCR evaluation of nonsense TP53 mRNA stabilization by pharmacological cell treatment. NMD-mediated degradation of mutated TP53 mRNA was inhibited by HCC1937 treatment with emetine subsequently followed or not by actinomycin D incubation. Non treated cells were incubated with ethanol vehicle.

However, the pharmacological treatment for NMD inhibition as originally described by Noesie and Dietz, and subsequently modified, has a major drawback causing a high background noise due to up-regulation of genes involved in the cellular stress-response, and actinomycin D compromises subsequent gene expression profiling due to an overall reduction of transcripts level (Ivanov *et al*.,2007).

To target more specifically the NMD pathway, the expression of two core proteins involved in nonsense transcript decay, UPF1 and UPF2 were transiently silenced by siRNAs. HCC1937 cells were transfected with pools of four siRNA targeting UPF1 or UPF2, and the decrease of mRNA and protein levels was assessed at different time points (48h, 72h, and 96h). Control cells were transfected with a pool of non-targeting siRNAs.

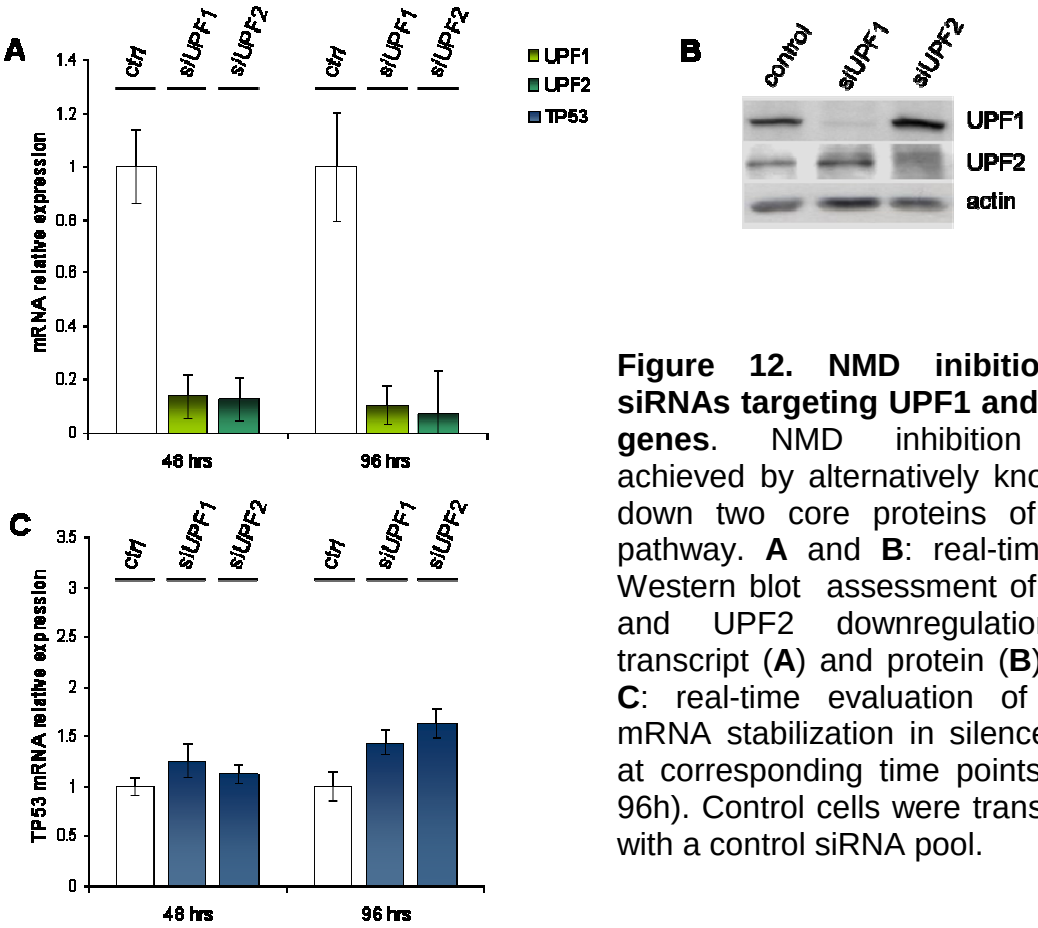

Figure 12. NMD inibition by siRNAs targeting UPF1 and UPF2 genes. NMD inhibition was achieved by alternatively knocking-down two core proteins of NMD pathway. **A** and **B**: real-time and Western blot assessment of UPF1 and UPF2 downregulation, at transcript (**A**) and protein (**B**) level. **C**: real-time evaluation of TP53 mRNA stabilization in silenced cell at corresponding time points (48h, 96h). Control cells were transfected with a control siRNA pool.

Both real-time PCR and Western blot experiments showed a good level of gene silencing (more than 80%) and protein amount reduction (Figure 12A and 12B).

The degree of stabilization of the mutant TP53 transcript was then verified by real-time PCR. Despite the low level of expression of UPF1 and UPF2 proteins, TP53 transcript did not increased significantly (Figure 12C), and would been probably insufficient for a microarray-based detection.

To improve nonsense transcripts up-regulation following NMD inhibition, HCC1937 cells were co-transfected with α-UPF1 and α-UPF2 siRNAs pools and the increase of premature terminating TP53 transcript was assessed at longer time points (96h, 120h, and 144h). Cells were first transfected with different concentrations of both siRNA pools, to identify the best condition. This approach led to an appreciable improvement in mutant TP53 transcript stabilization. Compared to control cells, downregulation of both UPF1 and UPF2 genes (Figure 13A) led to a 3.5-fold increase in TP53 mRNA level (Figure 13B). Supporting the efficacy of our method, the siRNA-mediated downregulation of UPF1 alone led to a lower increase in TP53 transcript level as shown in figures 13C and 13D.

2.1.2 High-Penetrance Genes in the HCC1500 Cell Line

Looking for new genes involved in the aetiology and progression of hereditary breast cancer, the HCC1500 cell line, was first analyzed to exclude possible pathogenetic mutations in BRCA1, BRCA2, and TP53 high-penetrance genes.

DHPLC and sequencing based mutational screening of BRCA1 and BRCA2 detected only an intronic variant (IVS 24-83 G>A) and a c.10462 A>G substitution (p.I3412V) in the BRCA2 gene. This missense sequence alteration occurs in the last exon and is reported at BIC website as UV, but both its high allelic frequency and its *in silico* evaluation suggest that it most likely represents a neutral polymorphism.

The MLPA analysis did not detect any large genomic rearrangements within both BRCA1 and BRCA2 loci, even though the somatic loss of one BRCA2 allele was found.

Also the analysis of the most frequently mutated exons of TP53 gene did not unveil any sequence alteration. Thus, HCC1500 cells were considered representative of a non-BRCA1/BRCA2 familial cancer,

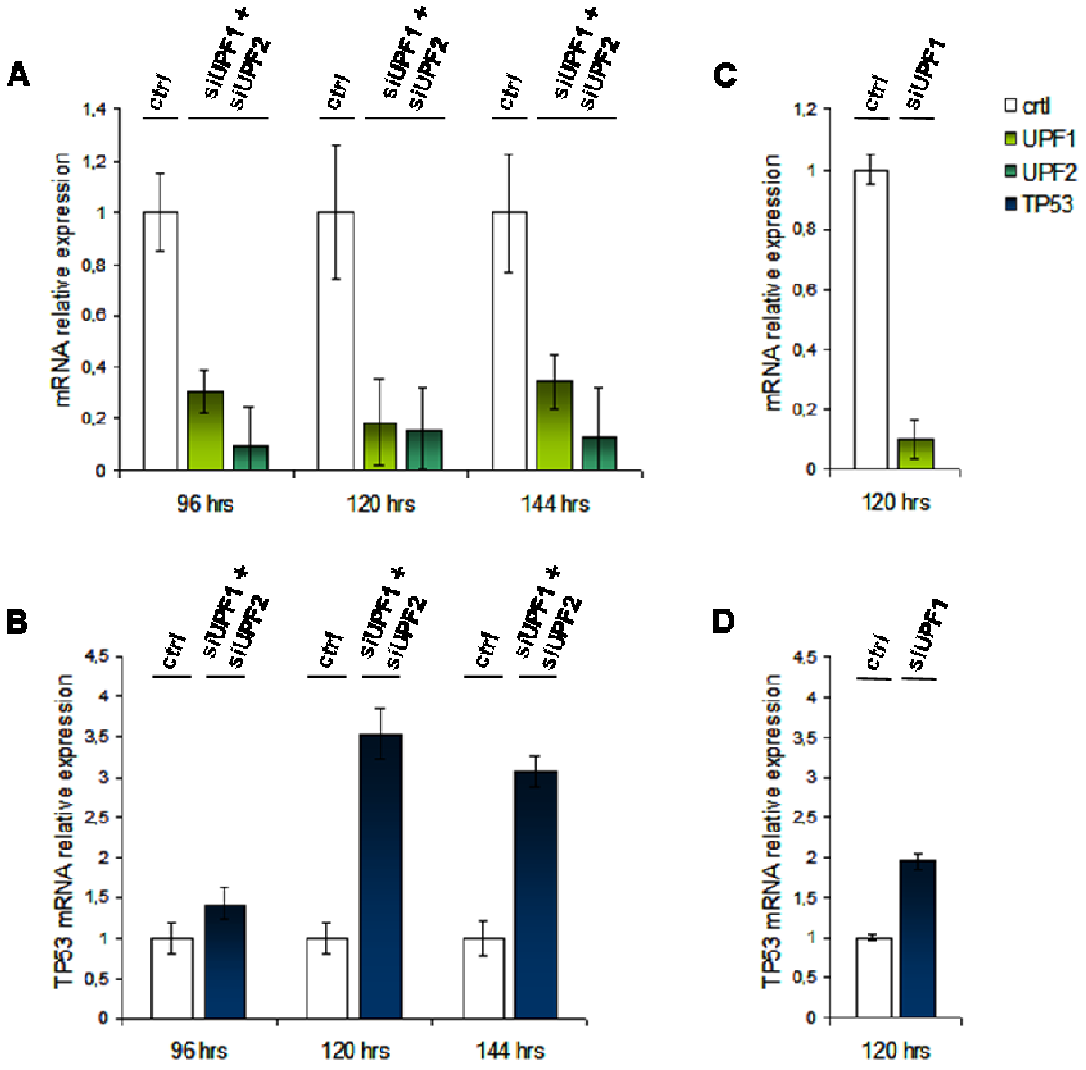

Figure 13. NMD-inhibition by simultaneous silencing of both UPF1 and UPF2. To improve NMD inhibition, HCC1937 cells were transiently silenced to downregulate both UPF1 and UPF2 up to 6 days. **A**: real-time PCR assessment of UPF1 and UPF2 mRNA level in silenced and control cells. **B**: relative expression of mutated TP53 mRNA at corresponding time points (96h, 120h, 144h). **C** and **D**: comparatively, HCC1937 cells were silenced only for UPF1 gene and the degree of NMD pathway inhibition was assessed by real-time PCR measuring nonsense TP53 transcript at 120h.

2.1.3 NMD Inhibition in HCC1500 Cells: Looking for Point Mutations

Optimized siRNA-mediated NMD inhibition was performed in the HCC1500 breast cancer cells and siRNAs were shown to be effective in reducing UPF1 and UPF2 gene expression up to 6 days (Figure 14).

After NMD inhibition, the search for putative premature terminating mRNAs was achieved through microarray-based gene expression profiling. For this purpose, we chose to analyze the HCC1500 NMD-silenced cells at the 120 hrs, when the control TP53 PTC-harboring transcript showed the maximum level of increase (Figure 13B).

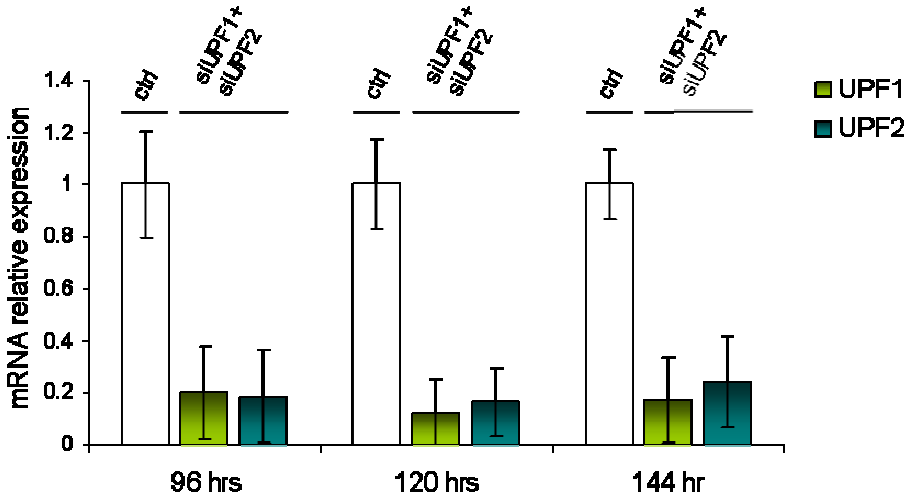

Figure 14. UPF1 and UPF2 siRNA mediated knockdown in HCC1500 cells. Real-time PCR confirmed the relevant degree of simultaneous UPF1 and UPF2 silencing in the HCC1500 cells.

Both real-time PCR and microarray analysis confirmed the good degree of UPF1 and UPF2 genes silencing in all three replicates used to hybridize microarrays, and highlighting the high reproducibility of the experimental setting. Microarray data analysis led to identification of 720 genes, which were significantly up-regulated after UPF1 and UPF2 silencing. This data set included genes involved in ubiquitination, apoptosis, proliferation, DNA damage, amino acids metabolism, as well as genes already reported downregulated by NMD (Mendell *et al*., 2004; Wittmann *et al*., 2006).

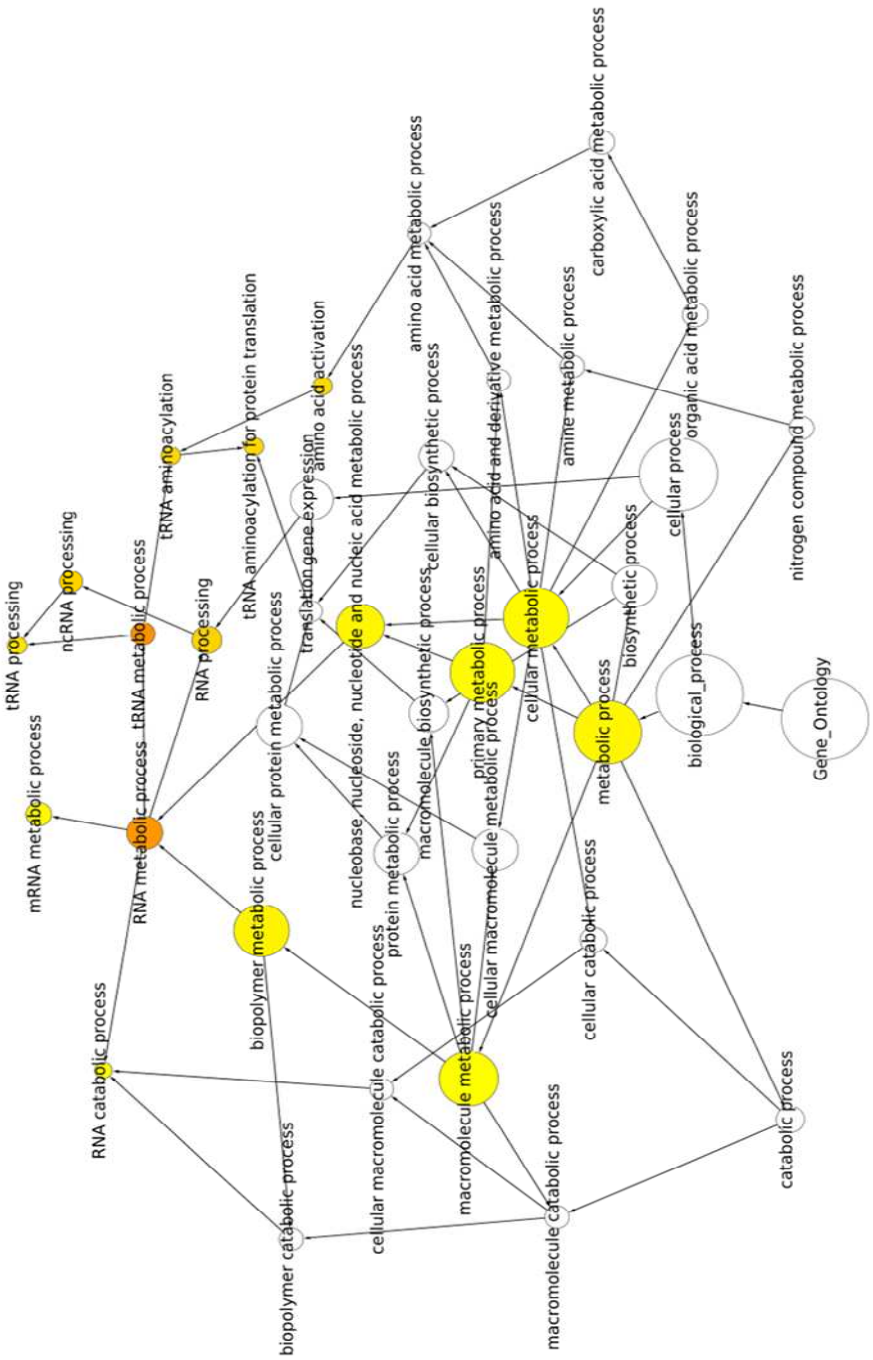

Figure 15. Gene Ontology enrichment of genes significantly up-regulated in HCC1500 cells NMD inhibited. The graph shows the Gene Ontology (GO) biological process terms characterizing genes up-regulated after NMD inhibition (nodes) and their complex network of relationship. Circles dimension is proportional to the number of genes; the color intensity is proportional to statistical significance of terms enrichment (orange means lower p-value). The analysis showed a statistically significant enrichment in genes involved in the RNA and tRNA metabolic processes.

A statistical clustering analysis based on gene ontology classification (Figure 15) interestingly showed a significant enrichment in genes involved in RNA metabolic processes (p-value 4.7 10^{-6}) among up-regulated genes. These data could suggest a possible crosstalk between the mRNA decay pathway, or its core factors UPF1 and/or UPF2, and other pathways related, at least to some extent, to gene expression regulation.

Among the genes up-regulated in NMD-silenced cells, we took into consideration those located in genomic regions with LOH, to focus our attention on bi-allelic inactivation events (i.e. truncating mutation in one allele and LOH of the other one), which could underline putative tumor suppressor genes. This approach led to identification of 286 genes. To further reduce the number of candidate genes, we prioritized those located within genomic regions associated with breast cancer susceptibility (126 genes), whereas genes already reported up-regulated by UPF1 or UPF2 inhibition in previous studies (229 genes) were excluded. In addition, due to the correlation between breast cancer susceptibility genes identified to date and DNA repair pathways, we selected also genes involved in the maintenance of genome integrity and those encoding proteins interacting with BRCA1 or BRCA2. This rationale-based multiple selection strategy led to the identification of 29 genes (Table 2). Some of these genes have intriguing functions that could well support a role in breast cancer susceptibility. For instance, RAP80/UIMC1 is essential for the recruitment of BRCA1 at DNA damage sites and, in recent studies, it has been evaluated for its possible role as breast cancer susceptibility gene (Akbari *et al*., 2009; Osorio *et al*., 2009). ZSWIM7, MUS81, and ERCC8 directly participate in DNA repair pathways. CDKN2AIP, GADD45G, PINX1, RINT1, and TBRG1 genes are involved in pathways that control cell growth and genome stability, acting as tumor suppressor genes.

Table 2. Candidate bi-allelic inactivated genes in non-BRCA1/BRCA2 hereditary breast cancer.

Gene Name	Expression Fold Change	Chromosome Position	Inclusion Criteria	Gene Function
APLF	1.5	2p14	LOH, DNA repair	single-strand and double-strand DNA break repair
APTX	1.8	9p13.3	DNA repair	single-strand DNA and double-strand DNA break repair, base excision repair
C10orf82	2.4	10q25.3	LOH, Br Ca susceptibility	ND
C11orf61	1.6	11q24.2	LOH, Br Ca susceptibility	ND
CCDC25	1.5	8p21.1	LOH, Br Ca susceptibility	ND
CCDC84	2.2	11q23.3	LOH, Br Ca susceptibility	ND
CCL5	3.9	17q12	LOH, Br Ca susceptibility	chemoattractant for monocytes, memory T-helper cells and eosinophils
CDKN2AIP	3.0	4q35.1	LOH, p14ARF, MDM2 & p53-interacting protein	activates TP53/p53 by CDKN2A-dependent and independent pathways
COG3	1.6	13q14.12	LOH, Br Ca susceptibility	involved in ER-Golgi transport
DNAH7	1.5	2q32.3	LOH, Br Ca susceptibility	dynein heavy chain
ERCC8	1.6	5q12.1	LOH & amplification, DNA repair	involved in transcription-coupled repair and nucleotide excision repair (NER) processes.
GABRG1	1.9	4p12	LOH, Br Ca susceptibility	ligand-gated ionic channel that mediates GABA (gamma-aminobutyric acid)-induced inhibition of neurotransmission
GADD45G	3.1	9q22.2	Br Ca susceptibility , DNA repair	involved in the regulation of growth and apoptosis. Mediates activation of stress-responsive MTK1/MEKK4 kinase
GTF2H1	1.7	11p15.1	LOH, DNA repair	component of the core-TFIIH basal transcription factor involved in nucleotide excision repair (NER) process
HOOK2	1.6	19p13.13	LOH & amplification, Br Ca susceptibility	involved in microtubule nucleation or anchoring at centrosomes
MUS81	1.5	11q13.1	Br Ca susceptibility , DNA repair	DNA structure-specific endonuclease that targets 3'-flap structures, replication forks and nicked Holliday junctions
NARG1L	2.3	13q14.11	LOH, Br Ca susceptibility	may belong to a complex displaying N-terminal acetyltransferase activity

Gene Name	Expression Fold Change	Chromosome Position	Inclusion Criteria	Gene Function
PCDH17	1.6	13q21.1	LOH, Br Ca susceptibility	cell-adhesion protein
PCTK1	1.4	Xp11.3	LOH, BRCA1-interacting protein	serine/threonine-protein kinase
PHF11	1.7	13q14.3	LOH, Br Ca susceptibility	ND
PINX1	1.6	8p23.1	LOH, Br Ca susceptibility	inhibits telomerase activity
PNLIPRP3	3.3	10q25.3	LOH, Br Ca susceptibility	pancreatic lipase-related protein
RINT1	2.0	7q22.2	cell cycle checkpoint, RAD50-interacting protein	involved in regulation of membrane traffic between the Golgi and the ER. It is essential for telomere length control and may play a role in G2/M cell cycle checkpoint control
DPCD	1.5	10q24.32	LOH, Br Ca susceptibility	may play a role in the formation or function of ciliated cells
SIDT2	1.4	11q23.3	LOH, Br Ca susceptibility	ND
TBRG1	1.9	11q24.2	LOH, Br Ca susceptibility	acts as a growth inhibitor and it is involved in maintaining chromosomal stability
TGDS	1.6	13q32.1	LOH, Br Ca susceptibility	dTDP-glucose 4,6-dehydratase
RAP80/UIMC1	2.1	5q35.2	LOH & amplification, BRCA1-interacting protein, DNA repair	ubiquitin-binding protein that specifically recognizes and binds Lys-63-linked polyubiquitin chains
ZSWIM7	1.5	17p12	LOH, DNA repair	involved in homologous recombination DNA repair

Table 2. (continued) For each gene expression fold change after NMD inhibition, chromosomal localization, criteria that allow prioritarization (LOH: loss of heterozygosity, Br Ca susceptibility: loci reported in literature as possibly harbouring breast cancer susceptibility genes), and known encoded protein function are reported. In

Despite this rational-based gene selection, most of the variants identified were missense changes and no frameshift nor nonsense mutations were found in HCC1500 cells (Table 3). However, the sequencing of candidate genes highlighted three sequence variants that could impair protein function. DNAH7 gene, encoding a protein that belongs to the dynein heavy chain family, has two missense substitutions, not yet reported in the literature, within exon 38 and exon 43.

Gene	SNP rs# / Mutation Type	Exon	Nucleotide Change	Amino Acid Change
APLF	rs11902811	3	A→G	p.Ile100Val
	rs35002937	6	T→A	p.Ser224Thr
COG3	rs2985959	12	A→G	p.Glu433Glu
	rs3014960	14	A→G	p.Gln497Gln
	rs3014902	23	C→T	p.Ser825Leu
	rs56244086	23	A→G	p.Ser825Ser
DNAH7	rs17838596	10	A→G	p.Ile315Val
	rs6719500	18	A→G	p.Lys825Glu
	rs60214909	35	C→T	p.Ser1856Ser
	MS	38	A→T	p.Asp2045Val
	MS	43	A→G	p.Glu2622Gly
	rs16841018	45	A→G	p.Ile2809Val
GABRG1	rs976156	3	A→G	p.Thr88Thr
HOOK2	S	11	C→A	p.Arg366Arg
	rs35337531	15	G→A	p.Leu462Leu
MUS81	rs34381357	3	C→T (het)	p.Ser115Phe
PHF11	rs2031532	3	A→G	p.Leu57Leu
PINX1	rs1078543	7	A→T	p.Ser254Cys
PNLIPRP3	rs10736251	10	G→A	p.Val381Ile
	rs1897519	10	A→G	p.Arg382Gly
RAP80/UIMC1	rs1700490	3	C→T	p.Ile76Ile
	rs365132	9	C→A	p.Thr448Thr
	rs10475633	12	G→A	p.Gly596Glu

Table 3. Sequence variations identified in candidate genes. The analysis was performed by aligning HCC1500 gene sequences to the reference sequences deposited at NCBI database (SeqScape v2.5 software, Applied Biosystems). S: synonymous; MS: missense; het: heterozygous.

Both sequence variants were predicted to be damaging by *in silico* analysis; in particular, the functional relevance of the Asp2045 residue is supported by its high evolutionary conservation. Indeed, the p.Asp2045Val could impair the AAA3 ATP-binding domain, one of the six AAA domains which form the motor region of the protein, whereas p.Glu2622Gly could alter microtubules-binding ability of the DNAH7 stalk domain.

The synonymous nucleotide substitution within exon 11 of the HOOK2 gene (Table 3) failed to demonstrate any predictable effect on pre-mRNA splicing by in silico analyses

In silico evaluation of the effect of other sequence variants (see Material and Method section) showed that PNPLIPRP3 gene, encoding for a lipase, has an intronic sequence variant (IVS10+15 A>C) that might affect pre-mRNA splicing (Table 4). This intronic substitution was predicted to create a cryptic donor splice site that, in turn, can give rise to a downstream PTC potentially recognized by the NMD pathway. However, this *in silico* prediction has yet to be experimentally confirmed.

Gene	Nucleotide Variant Description
APLF	IVS9 +8 G>T
C10orf82	IVS4 +13 T>C
CCDC25	IVS7 -9 delT
PHF11	IVS8 +24 G>T
PINX1	IVS2 +19 G>A
	IVS4 -26 A>G
	IVS5 -13 A>G
PNLIPRP3	IVS4 +24 delT
	IVS10 +15 A>C
SIDT2	IVS22 -22 C>T
TBRG1	IVS7 +10 A>G
RAP80/UIMC1	IVS12 -17 G>A

Table 4. Intronic variants of candidate genes. Intonic variants located within 30 bp from exon boundaries were *in silico* evaluated for their effect on pre-mRNA splicing. In yellow, the nucleotide variant predicted to alter splicing and possibly creating a premature terminating mRNA. Nucleotide variant description for PHF11 gene refers to transcript variant 1.

2.2 Genomic Profiling of HCC1500 cells

The genomic profile of HCC1500 was analyzed for chromosomal copy number alterations (CNA) that could underline genes causally implicated in breast cancer. This analysis was performed by means of the Affimetrix platform SNP array 6.0, which allows a high resolution in copy number detection.

Overall, HCC1500 showed a high level of chromosomal instability, with a large number of amplifications and deletions (Figure 16), some of which corresponding to well known regions of copy number alteration in breast cancer. Indeed, genomic instability is a feature of breast cancer, and hereditary BRCA1/2-associated cancer in particular; breast cancer derived cell lines have similar, even if more complex, patterns of genomic alterations (Mackay *et al*., 2009).

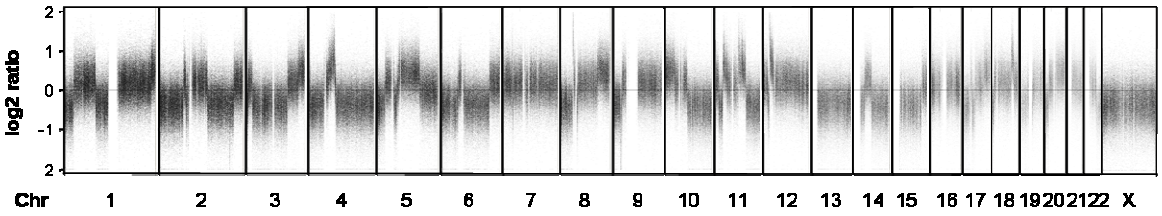

Figure 16. Genome-wide copy number profile of HCC1500 cell line. The values of $\log_2$ ratio derive from tumor hybridization intensity divided by normal reference samples from the HapMap project of ~2 million markers distributed throughout the genome. Log2 ratio 0 corresponds to the dyploid status.

Considering the large number of chromosomal alterations identified, we decided to focus on CNAs that either narrow regions of well known aberrations or represent focal alterations (i.e. alterations involving a small number of genes), thus aiming at the identification of a small number of specific candidate genes.

2.2.1 Genomic Amplifications

In HCC1500 cells relevant amplifications (i.e. > 3 copies) were detected at 1q43-qter, 2q14.3, 3q26.31-qter, 4q13.1-q21.1, 7p15, 8p12, 8q23.1-q24.13, 10p12.1-12.31, 10p15.3, 11p14.3-p14.1, 11q13.2-q14.1, 12p13-p11.21, 17q25.2-25.3, 18q21.32-q22.1, 20q13.33. Some of the amplified chromosomal regions, such as those located at 3q26.31-qter, 4q13.1-q21.1, 7p15.1, 8p12,

resulted from the high level amplification of one parental allele, whereas the other one is lost, thus suggesting the co-presence of oncogenes and tumor suppressor genes at the same chromosomal locations.

8p11-12, 8q24, 11q13, and 20q13 amplicons occur in a significant fraction of breast cancers, and each region harbors several genes that have been implicated in breast cancer development (Teschendorff and Caldas, 2009).

At 8p11-p12 cytoband, which is amplified in 10-15% of human breast cancers, there may be as many as four sub-regions of amplification, raising the possibility that there are multiple cancer-relevant genes in this region (Cooke *et al*., 2008). In HCC1500 cells we succeeded in defining the minimal amplified region to 370 kb encompassing only eight amplified genes (18 copies). Among these, RAB11FIP1, BRF2, PROSC, and ZNF703 genes were previously reported as candidate oncogenes, whereas the hypothesized driver gene FGFR1 is not included. Moreover, at locus 11q13.2-q13.3 HCC1500 cells have a minimal amplified region (9 copies) of 642 kb, including the supposed driver oncogene CCND1.

Another focal amplification was detected at chromosome 2q14.3, where a 580 kb DNA stretch was amplified (4 copies) involving only two genes: MKI67IP and TSN (Figure 17). MKI67IP gene encodes a nucleolar protein that interacts with the forkhead-associated (FHA) domain of the Ki-67 antigen, a proliferation-related protein. Interestingly, FHA domain is a phosphopeptide-binding domain present in a variety of nuclear cellular proteins involved in DNA repair, cell cycle arrest, or pre-mRNA processing. Interestingly, also the breast cancer susceptibility genes NBN and CHK2 encode FHA domain-containing proteins (Takagi *et al*., 2001). TSN is a DNA-binding protein which specifically recognizes conserved target sequences at the breakpoint junctions of chromosomal translocations and has been suggested to function also in DNA damage response (Abeysinghe *et al*., 2003).

Furthermore, at chromosome 7p15.1 a minimal amplified region of 624 kb was detected (5 copies). Interestingly, this CNA has its boundaries within two coding genes, CREB5 and CHN2/beta chimerin (see below), and includes the amplification of only two genes: CPVL, which encodes a carboxypeptidase first characterized in human macrophages, and the hypothetical protein-coding KIAA0644 gene.

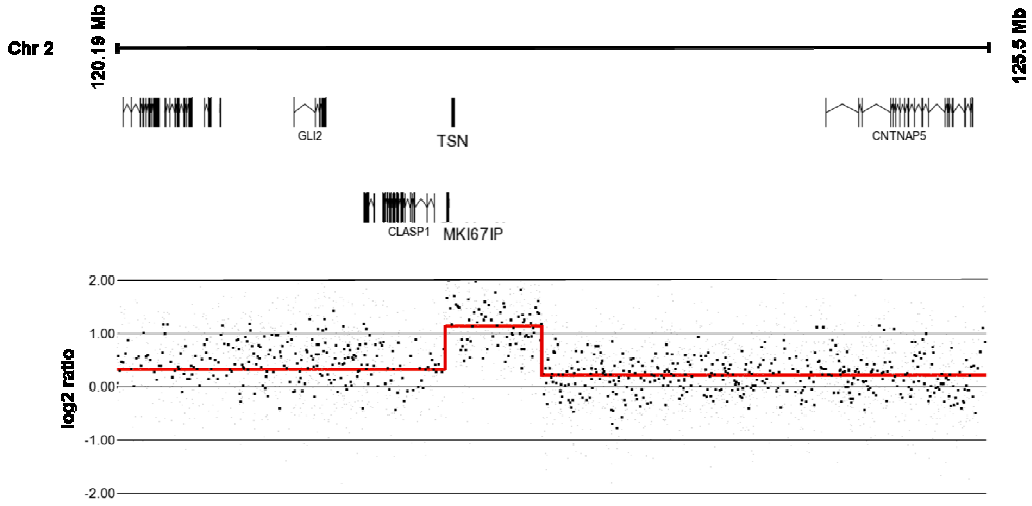

Figure 17. Focal amplification at locus 2q14.3. Identification of a small amplified region containing only the MKI67IP and TSN genes. The graph was generated with Partek Genomic Suite software.

2.2.2 Homozygous Deletions

The search for homozygous deletions in cancer genomes has been shown to be instrumental to the identification of tumor suppressor genes.

Homozygous deletions involving coding genes were identified at three distinct loci in the HCC1500 cell line (Table 5).

Chr	Cytoband	Minimal Region (kb)	Genes Involved
2	q33.3	43.56	PARD3B
9	p21.3	3882.3	IFNA14, IFNA5, KLHL9, INFA6, INFA13, INFA2, INFA8, IFNA1, INFE1, MTAP, CDKN2A, CDKN2B, DMRTA1, ELAVL2
22	q13.1	98.3	GTPBP1, UNC84B, DNAL4

Table 5. Homozygous deletions. Genes inactivated by homozygous deletion in HCC1500 cells.

In particular, chromosome 9p21.3 locus was shown to harbor a ~3.9 Mb homozygous deletion that involves KLHL9, MTAP, DMRTA1, ELAVL2, CDKN2B, CDKN2A (the well known tumor suppressor gene), and eight

members of an INF-α genes cluster. Homozygous deletions of MTAP, CDKN2A (p16/INK4A and p14/ARF), and CDKN2B (p15/INK4b) tumor suppressor genes are often reported in several types of human cancers and in tumor cell lines.
Interestingly, at chromosome 2q33.3 the PARD3B gene, which encodes a protein that co-localize at tight junctions in epithelial cells and participates to cell polarity establishment, is completely inactivated in HCC1500 cells by a complete deletion of one allele, while the other is inactivated by an intragenic alteration encompassing 43.6 kb (Figure 18).

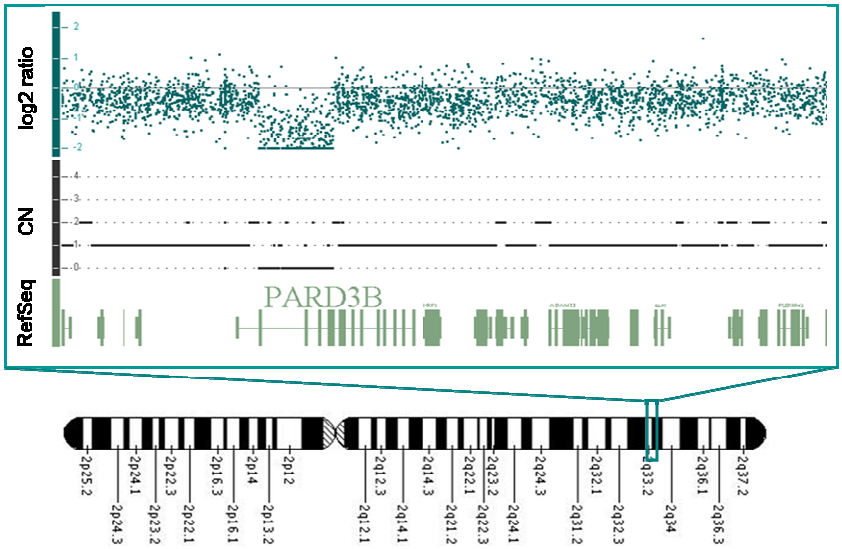

Figure 18. Intragenic homozygous deletion. In HCC1500 cells PARD3B is completely inactivated by loss of one allele and intragenic 43.5 kb deletion encompassing exons 2-10 of the other allele. Log2 ratio: tumor hybridization intensity divided by normal reference samples from HapMap project; CN: copy number state; RefSeq: reference sequences. Graph generated with the Genotyping Console software (Applied Biosystems).

At locus 22q13.1 a 98.3 kb homozygous deletion determined the complete loss of three genes: UNC84B, GTPBP1, and DNAL4. UNC84B is the human ortholog of UNC-84 gene, encoding a protein that acts at the nuclear envelope during nuclear migration and anchoring in *C. elegans*. In mammalian cells it was found to interact with, and thus regulate, Rab5, a ras-related GTPase protein. The other two genes, GTPBP1 and DNAL4, encode a GTP-binding protein and a dynein light chain, respectively.

2.2.3 Copy Number Alterations Driving Intragenic Rearrangements

The high resolution of copy number analysis (0.7 kb on average) provided with the SNP array 6.0 platform (Affymetrix), allowed us to undertake a third strategy for the identification of even more focal alterations involving single genes.

Genes clearly affected by intragenic genomic rearrangements are reported in tables 6 and 7. These genes are involved in a broad range of cell functions, most of which have been reported in the oncogenic pathways usually subverted in breast cancer and including: DNA repair, protein modification, cell shape and adhesion, DNA transcription, and cell signaling.

In order to discriminate genes that might represent neutral target of genomic instability (also referred to as "passenger alterations") from genes whose alteration is selected for during tumor development (i.e. "driver alterations"), this analysis was integrated with sequence data derived from genome-wide or candidate gene scans of breast and other cancers. Interestingly, somatic point mutations were reported in 40% (12 out of 30) of the genes identified (The Catalogue of Somatic Mutations in Cancer - COSMIC - described in Forbes et al., 2008; Wood *et al*., 2007) (Tables 6 and 7).

Several genes among those listed represent appealing candidates for breast cancer related genes. For example, LRP1B was found to be frequently inactivated both via genomic or epigenetic modifications in several types of cancer (lung adenocarcinoma, esophageal squamous cell carcinoma, glioblastoma multiforme, acute lymphoblastic leukemia), thus suggesting a role as tumor suppressor gene (Liu *et al*., 2000; Sonoda *et al*., 2004; Yin *et al*., 2009; Taylor *et al*., 2007). Similarly, PTPRT was found frequently mutated in colorectal and lung cancer (Wang *et al*., 2004 and COSMIC).

Table 6. Rearranged genes with loss of either 5' or 3' partion of their coding sequence

	Chr	Gytoband	Gene	Protein Molecular Function
5'-gene **loss**	Chr 2	q22.1-q22.2	**LRP1B**	Low density lipoprotein (LDL) receptor. It is a putative tumor suppressor gene
		q35	**XRCC5**	Single stranded DNA-dependent ATP-dependent helicase. Involved in double-strand break DNA repair by nonhomologous end joining (NHEJ)
	Chr 3	p24.3	**ZNF385D**	zinc-finge protein
	Chr 6	q12	**LGSN**	May act as a component of the cytoskeleton or as a chaperone for the reorganization of intermediate filament proteins during terminal differentiation in the lens
	Chr 11	p15.3	**GALNTL4**	May catalyze the initial reaction in O-linked oligosaccharide biosynthesis
		p14.3	**GAS2**	Caspase-3 substrate that plays a role in regulating microfilament and cell shape changes during apoptosis. . In HCC1500 cells it has also gain at the 3'-end
	Chr 14	q13.1	**NPAS3**	Belong to the bHLH-PAS family of transcription factors expressed in the brain. There are evidence of chromosomal rearrangement in psychiatric disease.
	Chr 16	q24.1	**COX4I1**	Nuclear-encoded subunit IV of cytochrome c oxidase complex of the mitochondrial respiratory chain enzyme. In HCC1500 cells it has also gain at the 3'-end
	Chr 18	p11.23	**PTPRM**	Encoded a receptor-type protein tyrosine phosphatase. It is involved in cell-cell adhesion and may play a key role in signal transduction and growth control
3'-gene **loss**	Chr 2	p13.3-p13.2	**EXOC6B**	Involved in the docking of exocystic vesicles with fusion sites on the plasma membrane.
		p12	**CTNNA2**	Gene that encode a alpha-N-catenin, a protein that mediates the link between cadherins and cytoskeleton
	Chr 3	q22.3	**STAG1**	Component of cohesin complex, that is required for the cohesion of sister chromatids after DNA replication. It is a target of the transcriptional activity of p53.
	Chr4	p16.1	**EVC**	Encodes a protein containing a leucine zipper and a transmembrane domain. This gene has been implicated in both Ellis-van Creveld syndrome (EvC) and Weyers acrodental dysostosis
	Chr 8	q11.22-q11.23	**PXDNL**	Peroxidase
	Chr 9	p21.2	**MOBKL2B**	May regulate kinases activity
	Chr 17	q11.2	**ZNF207**	zinc-finge protein
	Chr 22	q13.2	**SERHL2**	Serine hydrolase. In HCC1500 cells it has also gain at the 5'-end

Table 7. Rearranged genes with gain of either 5' or 3' partion of their coding sequence

	Chr	Gytoband	Gene	Protein Molecular Function
5'-gene **gain**	Chr 1	p22.3	**CLCA4**	Belongs to the calcium sensitive chloride conductance protein family. In HCC1500 cells it has also loss at 3'-end
	Chr 4	p12	**CORIN**	Serine protease involved in the processing pro-atrial natriuretic peptide, a cardiac hormone. In HCC1500 cells it has also loss at 3'-end
	Chr 7	p15.1	**CHN2/beta chimerin**	It is an intracellular receptor for the second messenger diacylglycerol and the phorbol esters, and as GTPase-activating protein negatively regulates p21rac. In HCC1500 cell it has also loss at the 3'-end
	Chr 12	p11.21	**FGD4**	Involved in the regulation of the actin cytoskeleton Activates CDC42, a member of the Ras-like family of Rho- and Rac proteins
		P12.3	**PLCZ1**	Phospholipase Ca^{2+}-dependent acting in the patway of phosphatidylinositol signaling
	Chr 18	q21.31-q21.32	**ALPK2**	Protein with kinase activity
	Chr 20	q12-q13.11	**PTPRT**	Receptor-type protein tyrosine phosphatase.
3'-gene **gain**	Chr 3	q26.31	**NAALADL2**	In a cell line from a patient with Cornelia de Lange syndrome, this gene was found to be disrupted by a t(3;17) translocation
	Chr 5	q23.2	**ALDH7A1**	Aldehyde dehydrogenase
	Chr 7	p15.1	**CREB5**	CRE (cAMP response element)-biding protein that activates transcription in a CRE-dependent way
	Chr 8	q23.1	**OXR1**	Mitochondrial protein involved in the cell protection against oxidative damage
	Chr 12	p12.3	**PIK3C2G**	Member of the PI-3-kinase. The C2 domain in C-terminal region acts as calcium-dependent phospholipid binding motifs that mediate translocation of proteins to membranes.
			PLEKHA5	Protein with a PH (Pleckstrin Homology)-domain that binds phosphatidylinositol trisphosphate

CHN2/chimerin, which is an effector of the second messenger diacylglycerol (DAG), stimulates p21rac GTPase activity leading to its inactivation (Yang *et al*., 2005). CHN2 transcript levels are significantly down-regulated in human breast cancer cell lines as well as in breast tumors. An insufficient expression level of CHN2 is expected to lead to higher p21rac activity, which acts on cell growth and proliferation associated with the RAS protein. An impaired control on the activity of p21rac could therefore play a role in the progression from low-grade to high-grade tumors (Dieckmann *et al*., 1991). Notably, p21rac is reported to up-regulate the expression of cyclin D1, which we found amplified in HCC1500 cells (Yang *et al*., 2005).

GAS2 encodes a protein that possibly enhances p53 stabilization by its inhibitory effect on m-calpain protease. Impaired expression of GAS2 gene could negatively affect p53 protein accumulation, therefore contributing to a lower cells susceptibility to p53-mediated apoptosis (Benetti *et al*., 2001). Notably, also STAG1, rearranged in HCC1500 cells, is involved in p53-mediated apoptotic process (Anazawa *et al*., 2004).

Of particular interest is the alteration that affects the CTNN2A gene, encoding a member of catenin protein family. This genomic rearrangement causing the 3' gene loss consists of a deletion of 3.78 Mb that does not involve any other gene.

As shown in figure 19, some of the CNAs analyses exemplify the complexity of copy number alterations, as also described for breast cancer amplicons at 8p12 and 11q13 (Kwek *et al*., 2009). In particular, in a 900 kb region at chromosome 12p12.3, HCC1500 cells showed distinct contiguous amplicons, whose boundaries fall within the coding sequence of three reference sequences. In particular, the rearranged PIK3C2G gene (Figure 19), due to the well known involvement of PI3K pathways in breast cancer (Leary *et al*., 2008), could represent another gene whose alteration contribute to the impairment this cancer related signaling cascade.

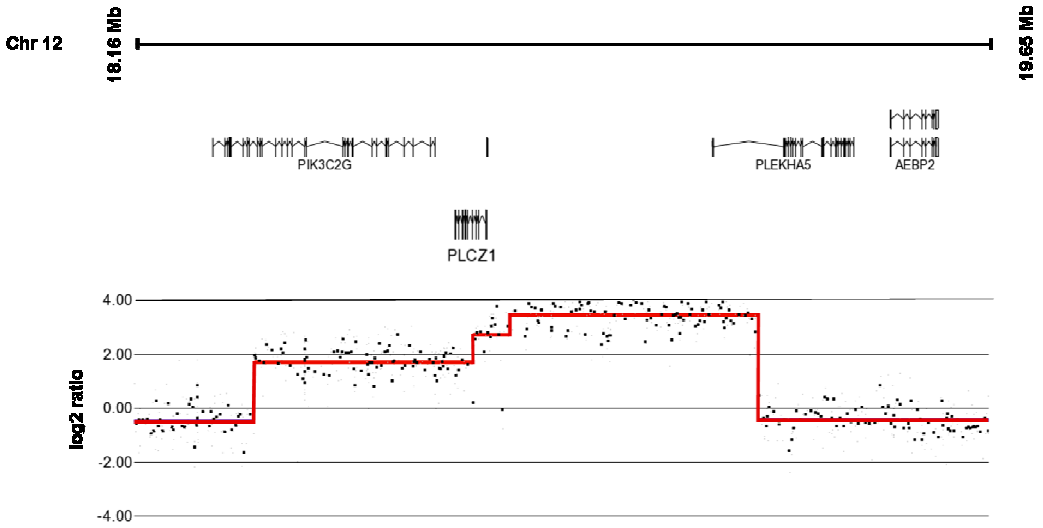

Figure 19. Focal high level DNA amplification at 12p12.3 locus. Within a 900 kb region a complex amplification determines the rearrangement of PIK3C2G, PLCZ1, and PLEKHA5 genes. The graph was generated with Partek Genomic Suite software.

2.2.4 Future directions: investigation of candidate gene relevance in non-BRCA1/BRCA2 hereditary breast cancer

Due to all the evidences described above (GINI approach, focal amplifications, homozygous deletion, and intragenic rearrangements), mutational characterization of HCC1500 cells provided some intriguing candidate genes that deserve further investigation (Table 8).

Combining all the evidences obtained with the genomic analyses described so far a number of different genes have been identified that might represent reasonable candidates with a direct role in breast cancer pathogenesis. Though many of the alterations affecting these genes might prove functionally related to the tumorigenic process, we sought a reclassification that could prioritize those genes that more likely deserve further investigation. Table 8 reports a subset of 12 genes that were selected based on their previously reported involvement in cancer and/or relevance of the biochemical pathway and /or peculiarity of the molecular alterations (i.e. specifically pointing to single genes).

Gene Name	Identification Strategy	Alteration in HCC1500 Cells	Selection Criteria: experimental evidences and protein function
DNAH7	GINI	MS, LOH	bi-allelic inactivation
PNLIPRP3	GINI	SM, LOH	putative PTC
TSN	CNA	focal amplification	putative oncogenetic function
MKI67IP	CNA	focal amplification	putative oncogenetic function
PARD3B	HD	complete inactivation	putative tumor suppressor
LRP1B	CNA	intagenic rearrangement	frequently inactivated in several tumor types
XRCC5	CNA	intagenic rearrangement	involved in DNA repair process (NHEJ)
GAS2	CNA	intagenic rearrangement	involved in the arrest of cell growth and in p53-mediated apoptosis
STAG1	CNA	intagenic rearrangement	involved in p53-mediated apoptosis
CHN2/ beta chimerin	CNA	intagenic rearrangement	low expression in breast cancer; regulates the activity of p21rac
PIK3C2G	CNA	intagenic rearrangement	belong to PI3K signaling pathway
PTPRT	CNA	intagenic rearrangement	frequently inactivated in some cancers

Table 8. Relevant genes selected for further evaluations. With the different analytical approaches (GINI and CNA analysis) in HCC1500 cells were identified several genes with specific alterations (MS: missense mutation; SM: splicing mutation) whose relevance in breast cancer susceptibility and progression could be better investigate.

Discussion

In this project, we characterized a breast cancer cell line by investigating, by means of different analytical approaches, both sequence point mutations and DNA copy number alterations (CNAs). The HCC1500 cell line was established from a patient who probably harbored a germline predisposing alteration, based on a number of features including early disease onset, a first-degree relative affected by breast cancer, and a family history for colon cancer. Involvement of the high-penetrance genes BRCA1, BRCA2, and TP53 was specifically ruled out by mutational screening.

Using the GINI strategy, we aimed at identifying putative tumor suppressor genes involved in non-BRCA1/BRCA2 hereditary breast cancer predisposition or progression and inactivated by nonsense mutations and deletions. Although we efficiently inhibited the NMD pathway, by specifically knocking-down UPF1 and UPF2 genes, and reduced false positives results by excluding those genes having features known to be associated with NMD targets, this approach was not as effective as expected, with the exception of the lipase coding gene PNLIPRP3, which had an intronic substitution predicted to alter pre-mRNA splicing and leading to a PTC. Among the candidate genes sequenced, we also identified two potentially deleterious missense mutations affecting DNAH7 and a synonymous amino acid change within HOOK2.

It is now well known that NMD core factors are involved in a broad range of distinct cellular processes, among which genome integrity and telomere maintenance, as well as in the regulation of many classes of physiologic transcripts (Isken and Maquat, 2008; Mendell *et al*., 2004). Thus, although our NMD inhibition approach was very specific, many *wild type* transcripts, either direct or indirect targets of the pathway, were stabilized and probably generated a noisy background weakening the detection power of the GINI approach.

Interestingly, in addition to the previously known physiologic transcripts regulated by the NMD pathway, NMD-inhibited HCC1500 cells showed a large number of mRNAs encoding proteins involved in the transcriptional and post-transcriptional regulation of gene expression. Indeed, the Gene Ontology enrichment analysis, performed on genes up-regulated following NMD

inhibition, evidenced a statistically significant presence of transcripts assigned to the "RNA metabolic process" (p-value $4.7 \cdot 10^{-6}$) and, more specifically to genes involved in "tRNA metabolic process" (p-value $2.9\ 10^{-6}$), and "RNA processing" (p-value $9 \cdot 10^{-4}$) (Figure 15, page 75). These data are consistent with previous findings suggesting a relationship between the NMD pathway and cellular homeostatic mechanisms (Mendell *et al*., 2004), and suggest a more complex and wide relationship between the NMD pathway itself, or UPF1 and UPF2 factors, and protein synthesis regulation.

With the advent of massive parallel sequencing technologies, which provide a comprehensive analysis of entire cancer genomes, a more detailed portrait of genomic alterations in breast cancer is now emerging (Teschendorff and Caldas, 2009). The genome-wide analysis of somatic gene alterations highlighted the presence of only a few genes frequently mutated, mainly PIK3CA and TP53, in contrast with a larger number of genes mutated at low frequency (Wood *et al*., 2007). These recent findings not only evidenced a relatively low number of genes affected by point mutations, but also suggested that nonsense mutations, in particular, are much less common than expected (Wood *et al*., 2007; http://www.sanger.ac.uk/genetics/CGP/cosmic/). This new light shed on the mutational spectrum of breast cancers likely contributes to explain the apparently low efficiency of the GINI strategy in the identification of truncated transcripts. Moreover, the GINI approach was recently used to analyze six breast cancer cell lines (Muggerud *et al*., 2009). This study found only one nonsense mutated gene in one of the cell line analyzed, thus confirming a low frequency of nonsense mutations.

In previous studies, the GINI approach was successfully applied to microsatellite instable cell lines that had a phenotype that highly increased the probability of frameshift-derived PTC (Huusko *et al*., 2004a; Ionov *et al*., 2004; Ivanov *et al*., 2007; El-Bchiri *et al*., 2005). Conversely, in a very recent work, Buffart *et al*. (2009) failed to identify nonsense mutated genes in two gastric cancer cell lines that were not characterized by microsatellite instability.

The second approach used to identify relevant DNA alterations in HCC1500 cell line is the CNA analysis. This analysis was performed through genome-wide genotyping of ~2 million markers (SNP Array 6.0, Affymetrix), allowing us a

better analytical resolution compared to conventional cytogenetic techniques and providing the means to precisely map the boundaries of CNAs and to pinpoint relevant genes. To date, few cytogenetic studies addressed the issue of chromosomal imbalances in non-BRCA1/2 hereditary breast tumors (Gronwald *et al*., 2005; Mangia *et al*., 2008). In line with data on CNA frequently found in non-BRCA1/BRCA2 breast tumors, HCC1500 cells show loss at 8p23.1-p23.3, and gain at 8q23.1-q24 and 19q12-qter, as well as 20q, which is significantly associated with non-BRCA1/BRCA2 ductal carcinomas (like the tumor of origin of the HCC1500 cell line) (Gronwald *et al*., 2005).

Supported by the high analytical resolution, we investigated three types of genomic alterations that could pinpoint candidate cancer genes: small amplified regions, homozygous deletions, and copy number alterations causing intragenic rearrangements. The picture of breast cancer genomes that emerges from large-scale studies is one of remarkable complexity, a feature characterizing cell lines as well as primary tumors, in which genome copy number changes and epigenomic modifications seem to deregulate the largest number of genes (Korkola ang Gray, 2010).

We succeeded in identifying small altered loci, as the focal amplification at 2q14.3 locus affecting only two genes, MKI67IP and TSN, both of which having biological functions possibly compatible with an oncogenic activity. TSN protein binds to a variety of single-stranded DNA sequences and consensus TSN binding sites were found to be significantly over-represented near translocation and deletion breakpoints, and it is thus suspected of mediating chromosomal translocations in lymphoid malignancies and in solid tumors (Abeysinghe *et al*., 2003). Its amplification could thus be involved in conferring a proliferative advantage to the genomic unstable HCC1500 cells.

Among the altered genes identified, some evidences, related to their biological function, suggest the existence of multiple alterations acting together in the activation of common oncogenes (CHN2/beta chimerin rearrangenent and cyclin D1 amplification) and likely confirming proliferative advantages, as well as resistance to apoptosis (rearrangement at STAG1 and GAS2 loci).

Among the rearranged genes, a very high proportion (40%) was already reported as affected by point mutations in several cancer types, thus supporting

a direct role of some of these genes in cancer development rather than a random accumulation of “passenger alterations”.

The evidences achieved so far at either gene and genomic level provide an intriguing starting point for future investigations. Together, GINI approach ad CNA analysis unveiled a subset of genes whose involvement in breast cancers initiation and progression can be further investigated. Genes highlighted by the searching of point mutations (PNLIPRP3, DNAH7) will be evaluate experimentally, in the attempt to clearly establish the detrimental effect of the sequence changes identified. Moreover, it will be interesting to evaluate whether any of the expressed genes rearranged in HCC1500 cells give rise to chimeric transcripts with possible oncogenic functions.

In conclusion, using a combined approach for the search of point mutations (GINI analysis) and copy number alterations we provide evidences suggesting that the major driving mutational mechanism in the HCC1500 cell line likely relies on a genomic instability that leads to multiple rearrangements and/or copy number imbalances. These data are also in line with a recently emerging scenario in which chromosomal aberrations and fusion transcripts would play previously unexpected relevant roles in breast cancer tumorigenesis (Teschendorff and Caldas, 2009).

By this approach we were also able to select a reasonable number of genes whose involvement in breast cancers is suggested by the specific mutational events identified as well as their biological function or previously demonstrated involvement into breast cancer pathways.

Bibliography

Abeysinghe SS, Chuzhanova N, Krawczak M, Ball EV, Cooper DN (2003). Translocation and gross deletion breakpoints in human inherited disease and cancer I: Nucleotide composition and recombination-associated motifs. Hum Mutat. 22(3): 229-44.

Agata S, Viel A, Della Puppa L, Cortesi L, Fersini G, Callegaro M, Dalla Palma M, Dolcetti R, Federico M, Venuta S, Miolo G, D'Andrea E, Montagna M (2006). Prevalence of BRCA1 genomic rearrangements in a large cohort of Italian breast and breast/ovarian cancer families without detectable BRCA1 and BRCA2 point mutations. Genes Chromosomes Cancer 45(9): 791-7.

Akbari MR, Ghadirian P, Robidoux A, Foumani M, Sun Y, Royer R, Zandvakili I, Lynch H, Narod SA. (2009). Germline RAP80 mutations and susceptibility to breast cancer. Breast Cancer Res Treat. 113(2): 377-81.

Anazawa Y, Arakawa H, Nakagawa H, Nakamura Y (2004). Identification of STAG1 as a key mediator of a p53-dependent apoptotic pathway. Oncogene 23(46):7621-7..

Antoniou A, Pharoah PD, Narod S, *et al.* (2003). Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case series unselected for family history: a combined analysis of 22 studies. Am J Hum Genet 72(5): 1117-30

Aretini P, D'Andrea E, Pasini B, et al. (2003). Different expressivity of BRCA1 and BRCA2: analysis of 179 Italian pedigrees with identified mutation. Breast Cancer Res Treat. 81(1): 71-9.

Argos M, Kibriya MG, Jasmine F, Olopade OI, Su T, Hibshoosh H, Ahsan H (2008). Genomewide scan for loss of heterozygosity and chromosomal amplification in breast carcinoma using single-nucleotide polymorphism arrays. Cancer Genet Cytogenet 182(2): 69-74

Behm-Ansmant I, Kashima I, Rehwinkel J, Sauliere J, Wittkopp N, Izaurralde E (2007). mRNA quality control: an ancient machinery recognizes and degrades mRNA with nonsense codons. FEBD Lett 581: 2845-2853

Benetti R, Del Sal G, Monte M, Paroni G, Brancolini C, Schneider C (2001). The death substrate Gas2 binds m-calpain and increases susceptibility to p53-dependent apoptosis. EMBO J 20(11): 2702-14.

Bergman A, Karlsson P, Berggren J, Martinsson T, Björck K, Nilsson S, Wahlström J, Wallgren A, Nordling M (2007). Genome-wide linkage scan for breast cancer susceptibility loci in Swedish hereditary non-BRCA1/2 families: suggestive linkage to 10q2332-q253. Genes Chromosomes Cancer 46(3): 302-9

Berliner JL, Fay AM (2007). Risk assessment and genetic counseling for hereditary breast and ovarian cancer: recommendations of the National Society of Genetic Counselors. J Genet Couns 16(3): 241-60

Bloethner S, Mould A, Stark M, Hayward NK (2008). Identification of ARHGEF17, DENND2D, FGFR3, and RB1 mutations in melanoma by inhibition of nonsense-mediated mRNA decay. Genes Chromosomes Cancer 47(12):1076-85

Buffart TE, Tijssen M, El-Bchiri J, Duval A, Wiel MA, Ylstra B, Meijer GA, Carvalho B (2009). NMD inhibition fails to identify tumour suppressor genes in microsatellite stable gastric cancer cell lines. BMC Med Genomics 29;2:39

Chang YF, Imam JS, Wilkinson MF (2007). The nonsense-mediated decay RNA survelliance pathway. Annu Rev Biochem 76: 51-74.

Chenevix-Trench G, Milne RL, Antoniou AC, Couch FJ, Easton DF, Goldgar DE; CIMBA (2007). An international initiative to identify genetic modifiers of cancer risk in BRCA1 and BRCA2 mutation carriers: the Consortium of Investigators of Modifiers of BRCA1 and BRCA2 (CIMBA). Breast Cancer Res 9(2): 104.

Conti E, Izaurralde E (2005). Nonsense-mediated mRNA decay: molecular insights and mechanistic variations across species. Curr Opin Cell Biol 17: 316-325.

Cooke SL, Pole JC, Chin SF, Ellis IO, Caldas C, Edwards PA (2008). High-resolution array CGH clarifies events occurring on 8p in carcinogenesis. BMC Cancer 8: 288.

Cox A, Dunning AM, Garcia-Closas M, et al (2007). A common coding variant in CASP8 is associated with breast cancer risk. Nat Genet 39: 352-358.

De Nicolo A, Parisini E, Zhong Q, Dalla Palma M, Stoeckert KA, Domchek SM, Nathanson KL, Caligo MA, Vidal M, Cusick ME, Garber JE (2009). Multimodal assessment of protein functional deficiency supports pathogenicity of BRCA1 p.V1688del. Cancer Res. 69(17): 7030-7.

Diekmann D, Brill S, Garrett MD, Totty N, Hsuan J, Monfries C, Hall C, Lim L, Hall A (1991). Bcr encodes a GTPase-activating protein for p21rac. Nature. 351(6325): 400-2.

Dumitrescu RG, Cotarla I (2005). Understanding breast cancer risk - where do we stand in 2005? J Cell Mol Med 9(1): 208-221.

Easton DF, Pooley KA, Dunning AM, et al (2007). Genome-wide association study identifies novel breast cancer susceptibilty loci. Nature 447: 1087-93.

El-Bchiri J, Buhard O, Penard-Lacronique V, Thomas G, Hamelin R, Duval A (2005). Differential nonsense mediated decay of mutated mRNAs in mismatch repair deficient colorectal cancers. Hum Mol Genet 14(16):2435-42.

El-Bchiri J, Guilloux A, Dartigues P, Loire E, Mercier D, Buhard O, Sobhani I, de la Grange P, Auboeuf D, Praz F, Fléjou JF, Duval A (2008). Nonsense-mediated mRNA decay impacts MSI-driven carcinogenesis and anti-tumor immunity in colorectal cancers. PLoS One 3(7): e2583.

Federico M, Maiorana A, Mangone L, Turchetti D, Canossi B, Cortesi L, Romagnoli R, Silingardi V (1999). Identification of families with hereditary breast and ovarian cancer for clinical and mammographic surveillance: the Modena Study Group proposal. Breast Cancer Res Treat 55(3):213-21

Ferrari F, Bortoluzzi S, Coppe A, Sirota A, Safran M, Shmoish M, Ferrari S, Lancet D, Danieli GA, Bicciato S (2007). Novel definition files for human GeneChips based on GeneAnnot. BMC Bioinformatics 8: 446.

Finch A, Beiner M, Lubinski J, *et al*. (2006). Salpingo-oophorectomy and the risk of ovarian, fallopian tube, and peritoneal cancers in women with a BRCA1 or BRCA2 Mutation. JAMA. 2006 Jul 12;296(2):185-92.

Forbes SA, Bhamra G, Bamford S, Dawson E, Kok C, Clements J, Menzies A, Teague JW, Futreal PA, Stratton MR (2008). The Catalogue of Somatic Mutations in Cancer (COSMIC). Curr Protoc Hum Genet Chapter 10:Unit 10.11.

Goldgar DE, Easton DF, Deffenbaugh AM, Monteiro AN, Tavtigian SV, Couch FJ; (2004). Integrated evaluation of DNA sequence variants of unknown clinical significance: application to BRCA1 and BRCA2. Am J Hum Genet 75(4):535-44.

Gonzalez-Neira A, Rosa-Rosa JM, Osorio A, Gonzalez E, Southey M, Sinilnikova O, Lynch H, Oldenburg RA, van Asperen CJ, Hoogerbrugge N, Pita G, Devilee P, Goldgar D, Benitez J (2007). Genomewide high-density SNP linkage analysis of non-BRCA1/2 breast cancer families identifies various candidate regions and has greater power than microsatellite studies. BMC Genomics 8: 299.

Gronwald J, Jauch A, Cybulski C, Schoell B, Böhm-Steuer B, Lener M, Grabowska E, Górski B, Jakubowska A, Domagała W, Chosia M, Scott RJ, Lubiński J (2005). Comparison of genomic abnormalities between BRCAX and sporadic breast cancers studied by comparative genomic hybridization. Int J Cancer 114(2): 230-6.

Gudmundsdottir K, Ashworth A (2006). The roles of BRCA1 and BRCA2 and associated proteins in the manteinance of genomic stability. Oncogene 25: 5864-5874.

Huusko P, Ponciano-Jackson D, Wolf M, *et al.* (2004a). Nonsense-mediated decay microarray analysis identifies mutations of EPHB2 in human prostate cancer. Nat Genet 36(9): 979-83.

Huusko P, Juo SH, Gillanders E, *et al.* (2004b). Genome-wide scanning for linkage in Finnish breast cancer families. Eur J Hum Genet 12(2): 98-104.

Ionov Y, Nowak N, Perucho M, Markowitz A, Cowell JK (2004). Manipulation of nonsense mediated decay identifies gene mutations in colon cancer cells with microsatellite instability. Oncogene 23 (3): 639-645.

Isken O, Maquat LE (2007). Quality control of eukaryotic mRNA: safeguarding cells from abnormal mRNA function. Genes Dev 21: 1833-3856.

Isken O, Maquat LE (2008). The multiple lives of NMD factors: balancing the roles in gene and genome regulation. Nat Rev Genet 9: 699-712.

Ivanov I, Lo KC, Hawthorn L, Cowell JK, Ionov Y (2007). Identifying candidate colon cancer tumor suppressor genes using inhibition of nonsense-mediated mRNA decay in colon cancer cells. Oncogene 26 (20): 2873-84.

Kainu T, Juo SH, Desper R *et al.* (2000). Somatic deletions in hereditary breast cancers implicate 13q21 as a putative novel breast cancer susceptibility locus. Proc Natl Acad Sci U S A. 97(17): 9603-8.

Kallioniemi A, Kallioniemi OP, Sudar D, Rutovitz D, Gray JW, Waldman F, Pinkel D (1992). Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. Science 258(5083): 818-21.

Korkola J, Gray JW (2010). Breast cancer genomes-form and function. Curr Opin Genet Dev. 2010 Epub ahead of print.

Kunnev D, Ivanov I, Ionov Y (2009). Par-3 partitioning defective 3 homolog (C elegans) and androgen-induced prostate proliferative shutoff associated protein genes are mutationally inactivated in prostate cancer cells. BMC Cancer 9: 318.

Kwek SS, Roy R, Zhou H, Climent J, Martinez-Climent JA, Fridlyand J, Albertson DG (2009). Co-amplified genes at 8p12 and 11q13 in breast tumors cooperate with two major pathways in oncogenesis. Oncogene 28(17):1892-903.

Le Hir H, Izaurralde E, Maquat LE, Moore MJ (2000). The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junction. EMBO J 19: 6860-6869.

Lareau LF, Inada M, Green RE, Wengrod JC, Brenner SE (2007). Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements. Nature 446 (7138): 926-9.

Leary RJ, Lin JC, Cummins J, *et al*. (2008). Integrated analysis of homozygous deletions, focal amplifications, and sequence alterations in breast and colorectal cancers. Proc Natl Acad Sci U S A. 105(42):16224-9.

Lejeune F, Maquat LE (2005). Mechanistic links between nonsense-mediated mRNA decay and pre-mRNA splicing in mammalian cells. Curr Opin Cell Biol 17: 309-315.

Liu CX, Musco S, Lisitsina NM, Yaklichkin SY, Lisitsyn NA. (2000). Genomic organization of a new candidate tumor suppressor gene, LRP1B. Genomics 69(2): 271-4.

Lykke-Andersen J, Shu MD, Steitz JA (2000). Human Upf proteins target an mRNA for nonsense-mediated decay when bound downstream of a termination codon. Cell 103: 1121-1131.

Mackay A, Tamber N, Fenwick K, Iravani M, Grigoriadis A, Dexter T, Lord CJ, Reis-Filho JS, Ashworth A. (2009). A high-resolution integrated analysis of genetic and expression profiles of breast cancer cell lines. Breast Cancer Res Treat. 118(3): 481-98.

Maguire P, Holmberg K, Kost-Alimova M, Imreh S, Skoog L, Lindblom A (2005). CGH analysis of familial non-BRCA1/BRCA2 breast tumors and mutation screening of a candidate locus on chromosome 17q11.2-12. Int J Mol Med 16(1): 135-41.

Malacrida S, Agata S, Callegaro M, Casella C, Barana D, Scaini MC, Manoukian S, Oliani C, Radice P, Barile M, Menin C, D'Andrea E, Montagna M (2008). BRCA1 p.Val1688del is a deleterious mutation that recurs in breast and ovarian cancer families from Northeast Italy. J Clin Oncol. 26(1): 26-31.

Mangia A, Chiarappa P, Tommasi S, Chiriatti A, Petroni S, Schittulli F, Paradiso A (2008). Genetic heterogeneity by comparative genomic hybridization in BRCAx breast cancers. Cancer Genet Cytogenet. 182(2): 75-83.

Maquat LE (2004). Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics. Nat Rev Mol Cell Biol 5: 89-99.

McGlincy N, Smith CWJ (2008). Alternative splicing resulting in nonsense-mediated mRNA decay: what is the meaning of nonsense? Trends BiochemSci 33(8): 385-93.

Mendell JT, Sharifi NA, Meyers JL, Martinez-Murillo F, Dietz HC (2004). Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. Nature Genet 36: 1073-1078.

Miller BJ, Wang D, Krahe R, Wright FA (2003). Pooled analysis of loss of heterozygosity in breast cancer: a genome scan provides comparative

evidence for multiple tumor suppressors and identifies novel candidate regions. Am J Hum Genet 73(4): 748-67.

Muggerud AA, Edgren H, Wolf M, Kleivi K, Dejeux E, Tost J, Sørlie T, Kallioniemi O (2009). Data integration from two microarray platforms identifies bi-allelic genetic inactivation of RIC8A in a breast cancer cell line. BMC Med Genomics: 2: 26.

Mühlemann O, Eberle AB, Stalder L, Orozco RZ (2008). Recognition and elimination of nonsense mRNA. Bioch Biophys Acta 1779(9): 538-49.

Mullan PB, Quinn JE, Harkin DP (2006). The role of BRCA1 in transcriptional regulation and cell cycle control. Oncogene 25: 5854-5863.

Nagy E, Marquat LE (1998). A role for termination-codon position within intron-containig genes: when nonsense affects RNA abbundance. Trend Biochem Sci 23: 198-199.

Noesie EN, Dietz HC (2001). A strategy for disease gene identification through nonsense-mediated mRNA decay inhibition. Nat Biotechnol 19: 434-439.

Oldenburg RA, Meijers-Heijboer H, Cornelisse CJ, Devilee P (2007). Genetic susceptibility for breast cancer: how many more genes to be found? Crit Rev Oncol Hematol 63(2): 125-49.

Oldenburg RA, Kroeze-Jansema KH, Houwing-Duistermaat JJ, *et al.* (2008). Genome-wide linkage scan in Dutch hereditary non-BRCA1/2 breast cancer families identifies 9q21-22 as a putative breast cancer susceptibility locus. Genes Chromosomes Cancer 47(11): 947-56.

Osorio A, Milne RL, Honrado E, Barroso A, Diez O, Salazar R, de la Hoya M, Vega A, Benítez J (2007). Classification of missense variants of unknown significance in BRCA1 based on clinical and tumor information. Hum Mutat 28(5): 477-85.

Osorio A, Barroso A, García MJ, Martínez-Delgado B, Urioste M, Benítez J. (2009) Evaluation of the BRCA1 interacting genes RAP80 and CCDC98 in familial breast cancer susceptibility. Breast Cancer Res Treat. 113(2): 371-6.

Palacios J, Robles-Frías MJ, Castilla MA, López-García MA, Benítez J (2008). The molecular pathology of hereditary breast cancer. Pathobiology 75(2): 85-94.

Pinyol M, Bea S, Plà L, Ribrag V, Bosq J, Rosenwald A, Campo E, Jares P (2007). Inactivation of RB1 in mantle-cell lymphoma detected by nonsense-mediated mRNA decay pathway inhibition and microarray analysis. Blood 109(12): 5422-9.

Prucka SK, McIlvried DE, Korf BR (2008). Cancer risk assessment and the genetic counseling process: using hereditary breast and ovarian cancer as an example. Med Princ Pract 17(3): 173-89.

Ray ME, Yang ZQ, Albertson D, Kleer CG, Washburn JG, Macoska JA, Ethier SP (2004). Genomic and expression analysis of the 8p11-12 amplicon in human breast cancer cell lines. Cancer Res 64(1): 40-7.

Ripperger T, Gadzicki D, Meindl A, Schlegelberger B (2009). Breast cancer susceptibility: current knowledge and implications for genetic counselling. Eur J Hum Genet 17(6): 722-31.

Rosa-Rosa JM, Pita G, González-Neira A, Milne RL, Fernandez V, Ruivenkamp C, van Asperen CJ, Devilee P, Benitez J (2009a). A 7 Mb region within 11q13 may contain a high penetrance gene for breast cancer. Breast Cancer Res Treat 118(1): 151-9.

Rosa-Rosa JM, Pita G, Urioste M, Llort G, Brunet J, Lázaro C, Blanco I, Ramón y Cajal T, Díez O, de la Hoya M, Caldés T, Tejada MI, González-Neira A, Benítez J (2009b). Genome-wide linkage scan reveals three putative breast-cancer-susceptibility loci. Am J Hum Genet. 84(2): 115-22.

Rossi MR, Hawthorn L, Platt J, Burkhardt T, Cowell JK, Ionov Y (2005). Identification of inactivating mutations in the JAK1, SYNJ2, and CLPTM1 genes in prostate cancer cells using inhibition of nonsense-mediated decay and microarray analysis. Cancer Genet Cytogenet 161(2): 97-103.

Seitz S, Rohde K, Bender E, Nothnagel A, Kölble K, Schlag PM, Scherneck S. (1997). Strong indication for a breast cancer susceptibility gene on chromosome 8p12-p22: linkage analysis in German breast cancer families. Oncogene. 14(6): 741-3.

Sonoda I, Imoto I, Inoue J, Shibata T, Shimada Y, Chin K, Imamura M, Amagasa T, Gray JW, Hirohashi S, Inazawa J. (2004). Frequent silencing of low density lipoprotein receptor-related protein 1B (LRP1B) expression by genetic and epigenetic mechanisms in esophageal squamous cell carcinoma. Cancer Res 64(11): 3741-7.

Stacey SN, Manolescu A, Sulem P, et al (2007). Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer. Nat Genet 39: 865-9.

Stalder L, Mühlemann O (2008). The meaning of nonsense. Trends Cell Biol 18 (7): 315-21.

Stratton MR, Rahman N (2008). The emerging landscape of breast cancer susceptibility. Nat Genet 40(1): 17-22.

Taylor KH, Pena-Hernandez KE, Davis JW, Arthur GL, Duff DJ, Shi H, Rahmatpanah FB, Sjahputera O, Caldwell CW (2007). Large-scale CpG

methylation analysis identifies novel candidate genes and reveals methylation hotspots in acute lymphoblastic leukemia. Cancer Res 67(6): 2617-25.

Takagi M, Sueishi M, Saiwaki T, Kametaka A, Yoneda Y (2001). A novel nucleolar protein, NIFK, interacts with the forkhead associated domain of Ki-67 antigen in mitosis. Biol Chem 276(27):25386-91.

Tavtigian SV, Deffenbaugh AM, Yin L, Judkins T, Scholl T, Samollow PB, de Silva D, Zharkikh A, Thomas A (2006). Comprehensive statistical study of 452 BRCA1 missense substitutions with classification of eight recurrent substitutions as neutral. J Med Genet 43: 295–305.

Tavtigian SV, Greenblatt MS, Goldgar DE, Boffetta P (2008). Assessing pathogenicity: overview of results from the IARC Unclassified Genetic Variants Working Group. Hum Mutat 29:1261–4.

Thompson D, Szabo CI, Mangion J *et al.* (2002). Evaluation of linkage of breast cancer to the putative BRCA3 locus on chromosome 13q21 in 128 multiple case families from the Breast Cancer Linkage Consortium. Proc Natl Acad Sci U S A. 99(2): 827-31.

Thorslund T, West SC (2007). BRCA2: a universal recmbinase regulator. Oncogene 26: 7720-30.

Teschendorff AE, Caldas C. (2009). The breast cancer somatic 'muta-ome': tackling the complexity. Breast Cancer Res 11(2): 301.

Tuna M, Knuutila S, Mills GB (2009). Uniparental disomy in cancer. Trends Mol Med 15(3): 120-8.

Turnbull C, Rahman N (2008). Genetic predisposition to breast cancer: past, present, and future. Annu Rev Genomics Hum Genet 9: 321-45.

Vallon-Christersson J, Cayanan C, Haraldsson K, Loman N, Bergthorsson JT, Brøndum-Nielsen K, Gerdes AM, Møller P, Kristoffersson U, Olsson H, Borg A, Monteiro AN (2001). Functional analysis of BRCA1 C-terminal missense mutations identified in breast and ovarian cancer families. Hum Mol Genet. 10(4): 353-60.

van Beers EH, van Welsem T, Wessels LF, Li Y, Oldenburg RA, Devilee P, Cornelisse CJ, Verhoef S, Hogervorst FB, van't Veer LJ, Nederlof PM (2005). Comparative genomic hybridization profiles in human BRCA1 and BRCA2 breast tumors highlight differential sets of genomic aberrations. Cancer Res 65(3): 822-7.

Venkitaraman AR (2009). Linking the cellular functions of BRCA genes to cancer pathogenesis and treatment. Annu Rev Pathol Mech Dis 4: 461-87.

Wang Z, Shen D, Parsons DW, Bardelli A, Sager J, Szabo S, Ptak J, Silliman N, Peters BA, van der Heijden MS, Parmigiani G, Yan H, Wang TL, Riggins G, Powell SM, Willson JK, Markowitz S, Kinzler KW, Vogelstein B, Velculescu VE (2004). Mutational analysis of the tyrosine phosphatome in colorectal cancers. Science 304(5674): 1164-6.

Wiechec E, Overgaard J, Hansen LL (2008). A fragile site within the HPC1 region at 1q253 affecting RGS16, RGSL1, and RGSL2 in human breast carcinomas. Genes Chromosomes Cancer 47(9): 766-80.

Wittmann J, Hol EM, Jäck HM (2006). hUPF2 silencing identifies physiologic substrates of mammalian nonsense-mediated mRNA decay. Mol Cell Biol 26(4): 1272-87.

Wood LD, Parsons DW, Jones S, *et al*. (2007). The genomic landscapes of human breast and colorectal cancers. Science 318(5853): 1108-13.

Yang C, Liu Y, Leskow FC, Weaver VM, Kazanietz MG (2005). Rac-GAP-dependent inhibition of breast cancer cell proliferation by {beta}2-chimerin. J Biol Chem. 280(26): 24363-70.

Yin D, Ogawa S, Kawamata N, Tunici P, Finocchiaro G, Eoli M, Ruckert C, Huynh T, Liu G, Kato M, Sanada M, Jauch A, Dugas M, Black KL, Koeffler HP (2009). High-resolution genomic copy number profiling of glioblastoma multiforme by single nucleotide polymorphism DNA microarray. Mol Cancer Res 7(5): 665-77.

Zheng W, Long J, Gao YT, Li C, Zheng Y, Xiang YB, Wen W, Levy S, Deming SL, Haines JL, Gu K, Fair AM, Cai Q, Lu W, Shu XO (2009). Genome-wide association study identifies a new breast cancer susceptibility locus at 6q251. Nat Genet 41(3): 324-8.

Publications

Osorio A, Milne RL, Pita G, Peterlongo P, ..., **Casella C**, *et al*. (2009). Evaluation of a candidate breast cancer associated SNP in ERCC4 as a risk modifier in BRCA1 and BRCA2 mutation carriers Results from the Consortium of Investigators of Modifiers of BRCA1/BRCA2 (CIMBA). Br J Cancer; 101(12): 2048-54.

Scaini MC, Rossi E, de Siqueira Torres PL, Zullato D, Callegaro M, **Casella C**, Quaggio M, Agata S, Malacrida S, Chiarion-Sileni V, Vecchiato A, Alaibac M, Montagna M, Mann GJ, Menin C, D'Andrea E (2009). Functional impairment of p16(INK4A) due to CDKN2A pGly23Asp missense mutation. Mutat Res; 671(1-2): 26-32.

Antoniou AC, Sinilnikova OM, McGuffog L, Healey S, ... , **Casella C**, *et al*. (2009). Common variants in LSP1, 2q35 and 8q24 and breast cancer risk for BRCA1 and BRCA2 mutation carriers. Hum Mol Genet; 18(22): 4442-56.

Malacrida S, Agata S, Callegaro M, **Casella C**, Barana D, Scaini MC, Manoukian S, Oliani C, Radice P, Barile M, Menin C, D'Andrea E, Montagna M (2008). The BRCA1 p Val1688del is a deleterious mutation that recurs in breast/ovarian cancer families from North-East Italy. J Clin Oncol; 26 (1): 26-31.

Acknowledgements / Ringraziamenti

Innanzitutto ringrazio la Prof.ssa Paola Zanovello per l'ottimo lavoro svolto nella coordinazione della scuola di dottorato e per la sua gentile disponibilità.

Un grande grazie, davvero di cuore, ad Emma e Marco, per il supporto, i consigli e le critiche; spero di riuscire a mettere a frutto almeno parte delle tante cose che siete riusciti a trasmettermi in questi anni.

Un grandissimo grazie a tutte le mie compagne di viaggio! Un particolare Grazie a Mery, Monia e Federica, con le quali ho condiviso i bei momenti (ed anche quelli un po' meno belli) della vita di lab...fino a tarda sera! Se non ci foste state voi! Devo solo dirvi: Grazie! E poi un grandissimo Grazie a Daniela, la vicina di bancone che tutti vorrebbero! Sempre pronta a dare una mano. E ancora un grande grazie a Simona e Monica, per i consigli e l'aiuto che avete saputo darmi (oltre che per le buonissime torte & affini...). Grazie anche a Silvia, Silvia ed Eugenia: per aver dato un nuovo tocco con le vostre personalità il laboratorio!

Un grazie particolare alla mia famiglia. Grazie mamma: mi hai sempre sostenuto ed aiutata, ed ogni volta lo apprezzo sempre di più. Lo so che hai sempre avuto fiducia in me e che ora sarai contenta di questo traguardo, anche se mi porterà un po' lontana ..... ma solo per poco tempo!

Il ringraziamento finale è per Fabio, che più di tutti ha vissuto con me questo dottorato.