Analytica Chimica Acta 1046 (2019) 154–162

Contents lists available at ScienceDirect

Analytica Chimica Acta

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

A multi-gene panel beyond *BRCA1/BRCA2* to identify new breast cancer-predisposing mutations by a picodroplet PCR followed by a next-generation sequencing strategy: a pilot study

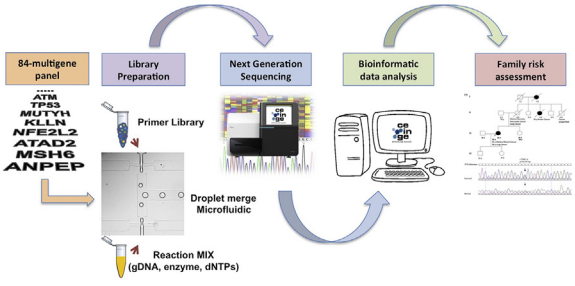
Marcella Nunziato [a, b, 1], Maria Valeria Esposito [a, b, 1], Flavio Starnone [a, b], Maria Angela Diroma [a], Alessandra Calabrese [a, c], Valentina Del Monaco [a], Pasqualina Buono [a, d], Giuseppe Frasci [c], Gerardo Botti [c], Massimiliano D'Aiuto [c], Francesco Salvatore [a, b, **], Valeria D'Argenio [a, b, *]

[a] CEINGE-Biotecnologie Avanzate, via Gaetano Salvatore 486, 80145, Naples, Italy
[b] Department of Molecular Medicine and Medical Biotechnologies, University of Naples Federico II, via Sergio Pansini 5, 80131, Naples, Italy
[c] Istituto Nazionale Tumori–IRCCS Fondazione Pascale, via Mariano Semmola 52, 80131, Naples, Italy
[d] Department of Movement Sciences and Wellness (DiSMEB), University of Naples Parthenope, via Medina 40, 80133 Naples, Italy

HIGHLIGHTS

- A picodroplet-PCR and NGS-based analysis of an 84 genes panel is proposed.
- Twenty-two *BRCA*-negative patients with familial breast cancer have been analyzed.
- Two pathogenic mutations and 8 potentially pathogenic variants were identified.
- Extended gene profiling better defines the predisposition risk for breast cancer.
- The diagnostic sensitivity of familial breast cancer testing is improved.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 16 July 2018
Accepted 14 September 2018
Available online 18 September 2018

Keywords:
Hereditary breast and ovarian cancer
Gene panel testing
Cancer risk
Next-generation sequencing

ABSTRACT

By analyzing multiple gene panels, next-generation sequencing is more effective than conventional procedures in identifying disease-related mutations that are useful for clinical decision-making. Here, we aimed to test the efficacy of an 84 genes customized-panel in *BRCA1* and *BRCA2* mutation-negative patients. Twenty-four patients were enrolled in this study. DNA libraries were prepared using a picodroplet PCR-based approach and sequenced with the MiSeq System. Highly putative pathogenic mutations were identified in genes other than the commonly tested *BRCA1/2*: 2 pathogenic mutations one in *TP53* and one in *MUTYH*; 2 missense variants in *MSH6* and *ATM*, respectively; 2 frameshift variants in *KLLN*, and *ATAD2*, respectively; an intronic variant in *ANPEP*, and 3 not functionally known variants (a frameshift variant in *ATM* a nonsense variant in *ATM* and a missense variant in *NFE2L2*). Our results show that this molecular screening will increase diagnostic sensitivity leading to a better risk assessment in

* Corresponding author. CEINGE-Biotecnologie Avanzate, and Dipartimento di Medicina Molecolare e Biotecnologie Mediche – Ed. 19, Via Sergio Pansini 5, 80131, Naples, Italy.,
** Corresponding author. CEINGE-Biotecnologie Avanzate, and Dipartimento di Medicina Molecolare e Biotecnologie Mediche – Ed. 19, Via Sergio Pansini 5, 80131, Naples, Italy.
E-mail addresses: salvator@unina.it (F. Salvatore), dargenio@ceinge.unina.it (V. D'Argenio).
[1] Equal contributors.

https://doi.org/10.1016/j.aca.2018.09.032

Germinal predisposing mutations
Picodroplet PCR

breast cancer patients and their families. This strategy could also reveal genes that have a higher penetrance for breast and ovarian cancers by matching gene mutation with familial and clinical data, thereby increasing information about hereditary breast and ovarian cancer genetics and improving cancer prevention measures or therapeutic approaches.

1. Introduction

Recent advances in next-generation sequencing (NGS) technology enabled the use of multiple gene panel tests to identify disease-related mutations [1]. These strategies, by allowing the simultaneous analysis of many genes of interest in several patients, can improve molecular diagnosis because they increase diagnostic accuracy and sensitivity, besides reducing analytical time and costs. Consequently, NGS-based gene panel analyses are now frequently used for the molecular diagnosis of several diseases [2–5]. Hereditary breast and ovarian cancer (HBOC) accounts for about 15% of all breast cancers (BCs), and is thus one of the most common hereditary cancer syndromes [6]. It is associated with an increased lifetime risk of multiple cancers, and often occurs at an earlier age with respect to the general population. Therefore, it is important to identify at-risk subjects, i.e., carriers of cancer-predisposing germline mutations, in order to offer patients and their families appropriate cancer surveillance screening and prophylactic therapeutic or surgical options [6–9].

It is generally agreed that HBOC is mainly associated to germline mutations in the tumor suppressor genes *BRCA1* and *BRCA2* [6,10]. Women with heterozygous mutations in the *BRCA1* or *BRCA2* genes have a lifetime risk of developing breast cancer of up to 80% and 60% respectively, and a cumulative risk of developing ovarian cancer of up to 60% [6,7]. Therefore, *BRCA* gene testing is currently offered worldwide to identify at-risk women, namely, those with a personal and/or familial history that fulfills the National Comprehensive Cancer Network guidelines for HBOC testing (NCCN) [11]. However, only a small proportion (25%) of such at-risk subjects is found to carry a deleterious mutation of the *BRCA* genes, which suggests that other genes may be involved in the development of HBOC [12,13].

In this scenario, NGS-based gene panels that simultaneously analyze diverse breast cancer-related genes are recently being used for inherited cancer risk identification and definition [14–17]. However, their use in the clinical setting is debatable because the number of causative mutations in genes other than the *BRCA* genes has been reported to be lower than expected [18,19]. Conversely, the number of variants with difficult-to-interpret clinical value is increasing, which suggests that caution be exercised regarding the application of these tests in routine molecular diagnostics [18,19].

Here, we report the analysis of a customized 84-gene NGS-based panel designed to analyze the definition of inherited risk of breast cancer. By increasing the diagnostic sensitivity of this risk by molecular testing, our strategy supports the use of a multistep diagnostic approach to improve the clinical management of patients and their families and ameliorate cancer prevention and surveillance programs.

2. Material and methods

2.1. Selection of patients and preparation of DNA samples

Twenty-four patients (21 women with BC, 1 man with BC, and 2 unaffected women with a positive family history) attending the Senology Department of the “Istituto Nazionale dei Tumori - Fondazione G. Pascale” of Naples, Italy, and who did not have a *BRCA1/2* mutation at conventional screening except 2 (see below) were enrolled in this study (Table 1). Pre-test genetic counseling was carried out to evaluate each patient's familial cancer history considering not only breast and ovarian cancers but also other types of cancer. Although all 24 patients had at least one breast cancer case in their family, they represent a heterogeneous group differing in terms of onset time (early versus late), breast cancer features, and/or a variety of cancers of a different nature, some with a high mortality rate even in young patients. Therefore, we postulated that a multiple gene panel could be used to investigate the predisposition risk of other inherited cancer genes beyond the two classical breast cancer genes conventionally used. Clinical data, including personal and family history, were collected and a three-generation family pedigree was constructed for each subject showing the number and the type of cancers present in each family (Table 1 and Figs. 1–3). Only 2 women with breast cancer were found to carry a mutation in the *BRCA2* gene and served also as positive controls to verify our analytical procedure [20]. All patients gave their written informed consent to the study that was carried out according to the tenets of the Helsinki Declaration and approved by the Istituto Nazionale Tumori - Fondazione G. Pascale Ethics Committee (protocol number 3 of 03/25/2009).

Genomic DNA was isolated from peripheral blood using the MasterPure™ Complete DNA Purification Kit (Epicentre Madison, WI, USA) according to the manufacturer's instructions. DNA quantity was evaluated using the NanoDrop 2000c Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA), while DNA quality was assessed by 0.8% agarose gel electrophoresis.

2.2. Gene panel selection

We designed a panel of 84 genes constituted by *BRCA1* and *BRCA2*, and other genes reported to be associated with breast and/or ovarian cancer [12–19]. For each gene we included the coding regions, 100 bp in each of the intronic boundaries, the promoter and the 3’ UTR for a total target size of about 1 Mb.

2.3. Custom-enriched library preparation and next generation sequencing

All samples were analyzed according to the flowchart reported in Fig. S1 in Supplementary Material and Methods. As an alternative to other methodologies for libraries preparation used in previous experimentations, including shotgun libraries, amplicons or probes-based approaches [1,10,20,21], we used herein a different strategy. In fact, the targeted enrichment of our customized panel was carried out using the RainDance ThunderStorm System (RainDance Technologies, Billerica, MA, USA), which is a microfluidic device that generates millions of unique picodroplet PCR reactions for high throughput multiplex amplification [1]. The miniaturized system that also includes a liquid handler reduces both hands-on time and operator-dependent variability. Moreover, when working with GC-rich regions, repetitive regions or pseudogenes, the separation of individual amplicons into individual droplet micro-reactors

Table 1
Clinical features of patients enrolled in the study. All patients belonged to families in which at least one case of breast cancer had been reported.

ID	Sex	Age	Under/Over 40 at onset	Proband Disease Status	BRCA genes Mutation Status	Menarche Age	Pregnancies	Other Risk Factors	Familial Cancer (case number)
P1	F	46	Over	Breast cancer	WT	N.R.	N.R.	N.R.	Breast (1)
P2	F	41	Under	Breast cancer	WT	N.R.	1	Smoke	Breast (1), ovarian (1)
P3	F	49	Over	Breast cancer	WT	14	1	No	Breast (3), intestine (1), lymphoma (1), lung (2)
P4	F	44	–	Healthy	WT	11	2 + 1 Abortion	No	Throat (1), breast (1), prostate (2), bone (4), lung (1)
P5	F	71	Over	Breast cancer	WT	12	1 + 1 Abortion	Smoke	Brain (1), breast (1)
P6	M	66	Over	Breast cancer	WT	–	–	No	Breast (1)
P7	F	55	Over	LABC	WT	11	4	Smoke	Breast (3)
P8	F	49	Under	Breast cancer	WT	12	1 + 2 Abortion	No	Breast (6), prostate (1)
P9	F	38	Under	Breast cancer	WT	11	3	No	Uterus (1), breast (2), bone (1), pancreas (1), throat (1), lung (1)
P10	F	37	Under	Breast cancer	WT	14	0	Oral Contraceptives, Smoke	Melanoma (1), intestine (1), leukemia (1), bladder (1), breast (4)
P11	F	72	Over	Breast cancer	WT	12	2	No	Intestine (2), lung (1), liver (1), breast (3)
P12	F	57	–	Healthy	WT	15	2	Oral Contraceptives, Smoke	Colon (1), breast (3), lung (5), liver (2), stomach (4), bladder (1), pancreas (1)
P13	F	43	Under	Breast cancer	WT	12	1	No	Breast (3), uterus (1), lung (1), prostate (1)
P14	F	51	Under	Breast cancer	BRCA2	13	3	Oral Contraceptives, Smoke	Breast (3), uterus (1)
P15	F	43	Under	Breast cancer	BRCA2	12	3 + 1 Abortion	Oral Contraceptives, Smoke	Breast (3), uterus (1)
P16	F	52	Over	Breast cancer	WT	11	0	No	Throat (1), breast (1)
P17	F	64	Over	Breast cancer	WT	13	2	Smoke	Breast (2), brain (1)
P18	F	64	Over	Breast cancer	WT	13	2	No	Uterus (1), prostate (1), breast (1), kidney (1)
P19	F	39	Under	Breast cancer	WT	13	2	Oral Contraceptives, Smoke	Liver (1), breast (2)
P20	F	54	Over	Breast cancer	WT	N.R.	0	Smoke, Ovarian Stimulation	Throat (1), breast (1)
P21	F	54	Over	Breast cancer	WT	12	2 + 1 Abortion	No	Lung (2), breast (2), kidney (2), uterus (1)
P22	F	80	Over	Breast cancer	WT	14	5 + 1 Abortion	Smoke	Breast (4)
P23	F	42	Under	Breast cancer	WT	13	0	Ovarian Stimulation, Oral Contraceptives	Breast (3), pancreas (2), lung (1), brain (1)
P24	F	48	Over	Breast cancer	WT	11	2 + 2 Abortion	Oral Contraceptives, Smoke	Breast (2), prostate (2), leukemia (1), lung (2)

P, Patient; F, Female; M, Male; WT, Wild Type; N.R., not reported; LABC, Local Advanced Breast Cancer.

facilitates high quality coverage similar to that obtained with less complex regions, thereby ensuring great coverage uniformity, which is a prerequisite for any clinical application of NGS. All these features are prerequisites for the routine analysis of multiple genes in a high number of patients simultaneously. Consequently, given these improvements, we used droplet PCR together with an NGS system as a more reliable method for the enrichment of NGS libraries, and validated this approach in a pilot study group.

Eight μg of genomic DNA/sample were fragmented (average length = 3000 bp) using the M220 Focused-ultrasonicator Covaris (Covaris, Inc., Woburn, MA, USA). Next, 1.5 μg of sheared DNA/sample were used to generate the ThunderStorm custom emulsion for the amplification of the target genes (amplicon length = 300–500 bp). After amplification, the emulsions were broken, and the enriched libraries were further processed to amplify the targets and attach personal indices adaptors. Enriched libraries were quality-assessed on a DNA 1000 LabChip (Agilent Technologies, Santa Clara, CA, USA). Twelve libraries were pooled for each sequencing run. Sequencing reactions were performed with the Illumina MiSeq System (PE 300x2) (Illumina Inc., San Diego, CA, USA). Causative or potentially interesting variants were confirmed by Sanger sequencing.

2.4. Bioinformatics

Bioinformatic analyses consisted of 6 steps as shown in Fig. S2 in Supplementary Material and Methods. Raw fastq files were filtered after quality checking with the NGS QC Toolkit v2.3.3 to remove primers and trim low quality ends of reads (quality score, QS < 20) [22]. Reads were aligned against the hg38 genome assembly (GenBank:GCA_000001405.15) using the BWA software v0.7.12, mem algorithm [23]. Post-alignment processing was performed using: i) SAMtools v1.2 for sorting, indexing and merging bam files; (ii) Picard tools v1.129 (http://picard.sourceforge.net) to add read groups; (iii) GATK v3.3 for insertions/deletions (indels) realignment and base QS recalibration [24,25]. Indels were realigned around known indels of the 1000 Genomes phase 1, after conversion of genome coordinates from hg19 (GenBank: GCA_000001405.1) to hg38 assembly, using CrossMap software v0.1.6 [26,27]. dbSNP variants (release 142, http://www.ncbi.nlm.nih.gov/SNP/) were used as reference for base QS recalibration. Variant calling and filtering were performed using the GATK HaplotypeCaller v3.3 in GVCF mode and SAMtools mpileup (see online supplementary Methods) [24,25]. Lastly, variant annotation and prioritization were performed using SnpEff v.4.1b and dbNSFP v3.0 beta1 tools (see Supplementary Material and Methods) [28–30].

3. Results

3.1. NGS data analysis

We obtained an average of high quality reads of 77%

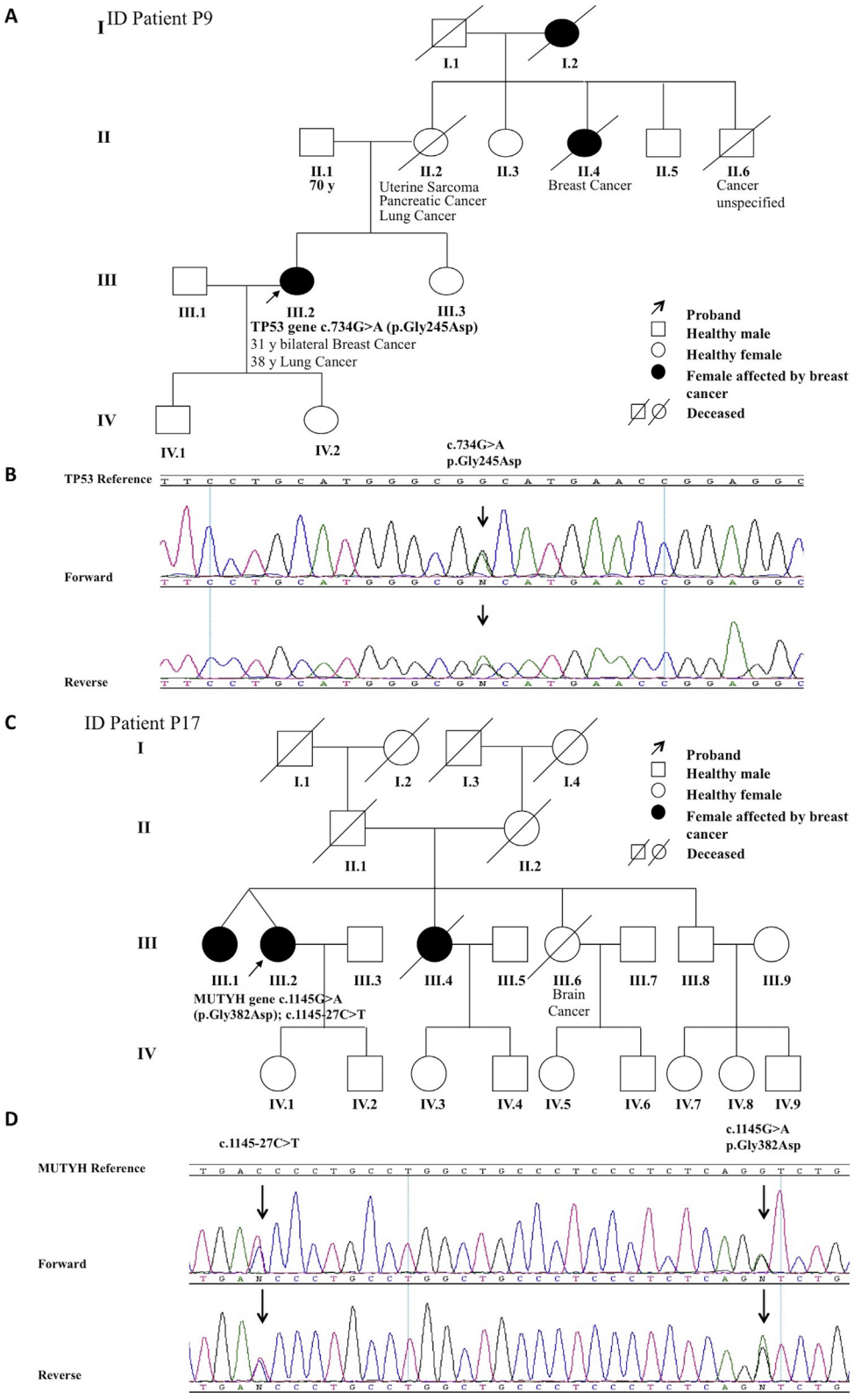


Fig. 1. Pedigree of P9 and P17. (A) P9, now 39 years old, was diagnosed with bilateral breast cancer at the age of 31 years, and with a primitive lung cancer at the age of 38. The patient's mother died from cancer, and during her life she was affected by uterine sarcoma, pancreatic cancer and lung cancer. One aunt died from breast cancer at the age of 30 years and an uncle died from unspecified cancer (both in the maternal arm). **(B)** P9 was negative for *BRCA* mutations but we identified a known pathogenic variant (c.734G > A, p.Gly245Asp [(rs121912656]) in the *TP53* gene, which was confirmed by Sanger sequencing. **(C)** P17 was diagnosed with breast cancer she was over the age of 40 years. Two sisters died from breast and brain cancer, respectively. The patient's twin sister also developed breast cancer. **(D)** In this patient we found the pathogenic variant c.1145G > A, p.Gly382Asp (rs36053993) in the *MUTYH* gene, as shown in the Sanger electropherogram.

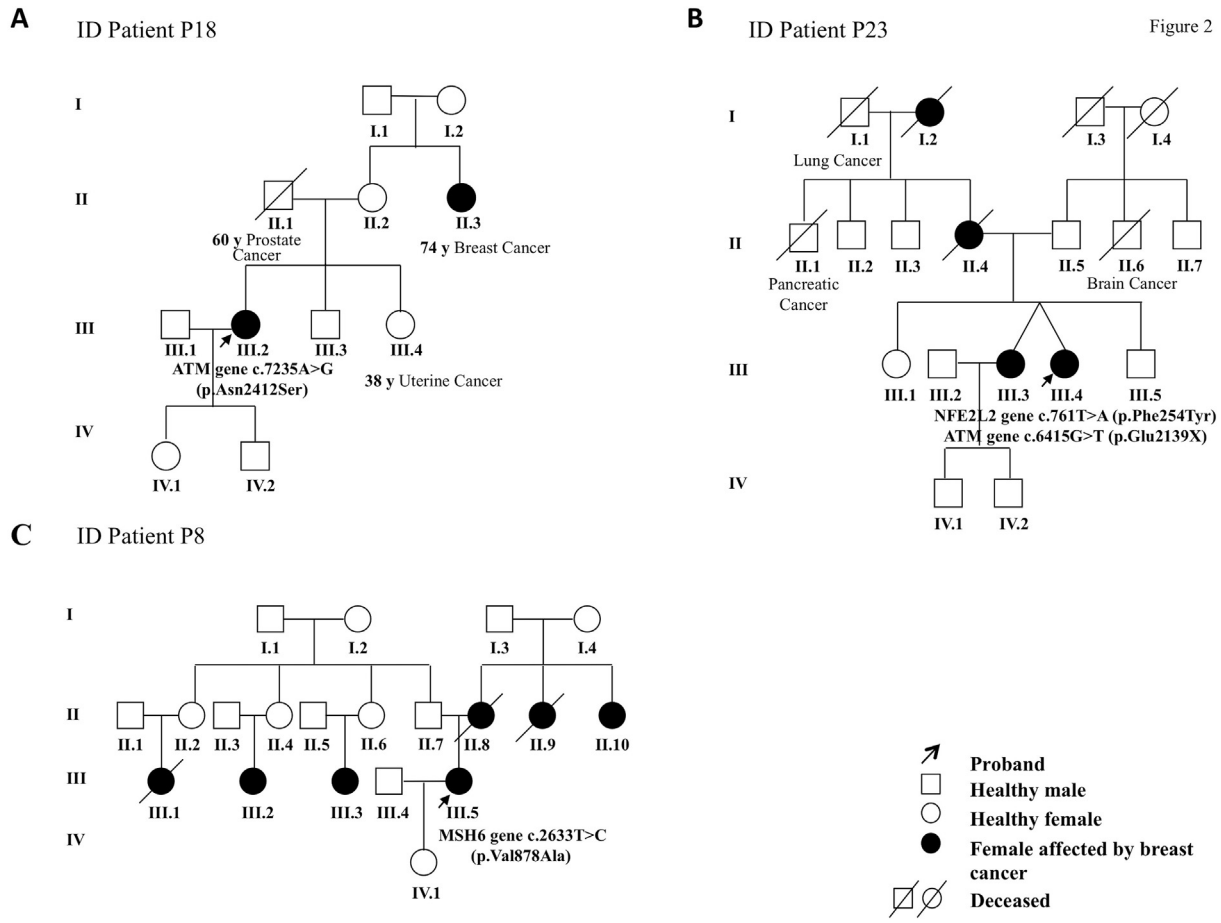


Figure 2

Fig. 2. Pedigrees of patients carrying at least one prioritized variant. Pedigrees of P18, P23 and P8 (A–C**).** The gene and their variants are reported on the top of each panel. The arrows indicate the proband from each family.

(minimum = 71%, maximum = 81%) (Table S1 in Supplementary Results). The alignment showed a very high percentage of mapped reads in each sample and a few unmapped reads (average 5.0%). The average mapping quality ranged from 47.8 to 55.2, which confirms high alignment efficiency. Uniquely mapped reads represented the largest fraction of the total amount of reads (from 91% to 96%). Fixing a threshold of minimum base quality score = 30, more than 90% of selected regions were almost entirely covered (>80% of their total exon length) and only 6 regions out of 2316 exons were not covered at all. The mean sample coverage ranged from 24.00 to 286.76 (Fig. S3A and Table S1 in Supplementary Results). To call only highly reliable variants, we considered a depth of coverage $\geq$10 acceptable. The percentage of covered target sites, meant as single bases, was variable within samples at a coverage depth $\geq$10 and $\geq$ 100 (Fig. S3B in Supplementary Results). One sample (patient 4) that had a very low coverage depth was not considered in the subsequent analysis.

Overall, in all the 24 samples we found 2324 single nucleotide variants and 752 indels after variant filtration with a minimum variant depth of 10 (Table S3 in Supplementary Results). The distribution of single nucleotide variants and indels within genomic regions was homogeneous among variants detected by the 3 methods (SAMtools, GATK VQSR and GATK HF) (Fig. S4A and B in Supplementary Results). About 50% of single nucleotide variants were missense mutations and only 2 STOP gain mutations were observed (GATK and VQSR), one of which was confirmed by both SAMtools and GATK HF (Fig. S4C in Supplementary Results). Eleven frameshifts/indels were identified by GATK HF, but only 4 were detected also by another method, consequently the remaining mutations were not considered reliable (Fig. S4D in Supplementary Results). See also supplementary information for further details.

3.2. Reliability of variant calling

The reliability of the targeted enrichment method described herein was assessed by evaluating its capability to detect gene sequence variants in the *BRCA1* and *BRCA2* genes. In fact, we had previously analyzed the 24 patients enrolled in this study using an NGS-based protocol to assess their *BRCA* mutational status [20]. In particular, in the *BRCA1* gene, the Thunderstorm System coupled with the above-mentioned bioinformatic strategy, identified 86 variants detected also by the first screening analysis (88%), 1 variant was identified only with the JSI Sequence Pilot software, while 11 were detected only by Thunderstorm but not confirmed by Sanger sequencing. Similarly, we found 110 variants in the *BRCA2* gene that were detected also by the first screening analysis (87.4%), 5 variants detected only with the JSI Sequence Pilot software, and 11 detected only by Thunderstorm but not confirmed by Sanger sequencing. Therefore, based on these data, the sensitivity of our procedure was 96.6% and its specificity >99.9%.

3.3. Mutation identification in high risk BC patients

The variant annotation and prioritization strategy revealed a small number of variants of potential interest as cancer-predisposing mutations in our patients. In fact, we analyzed 11

Figure 3

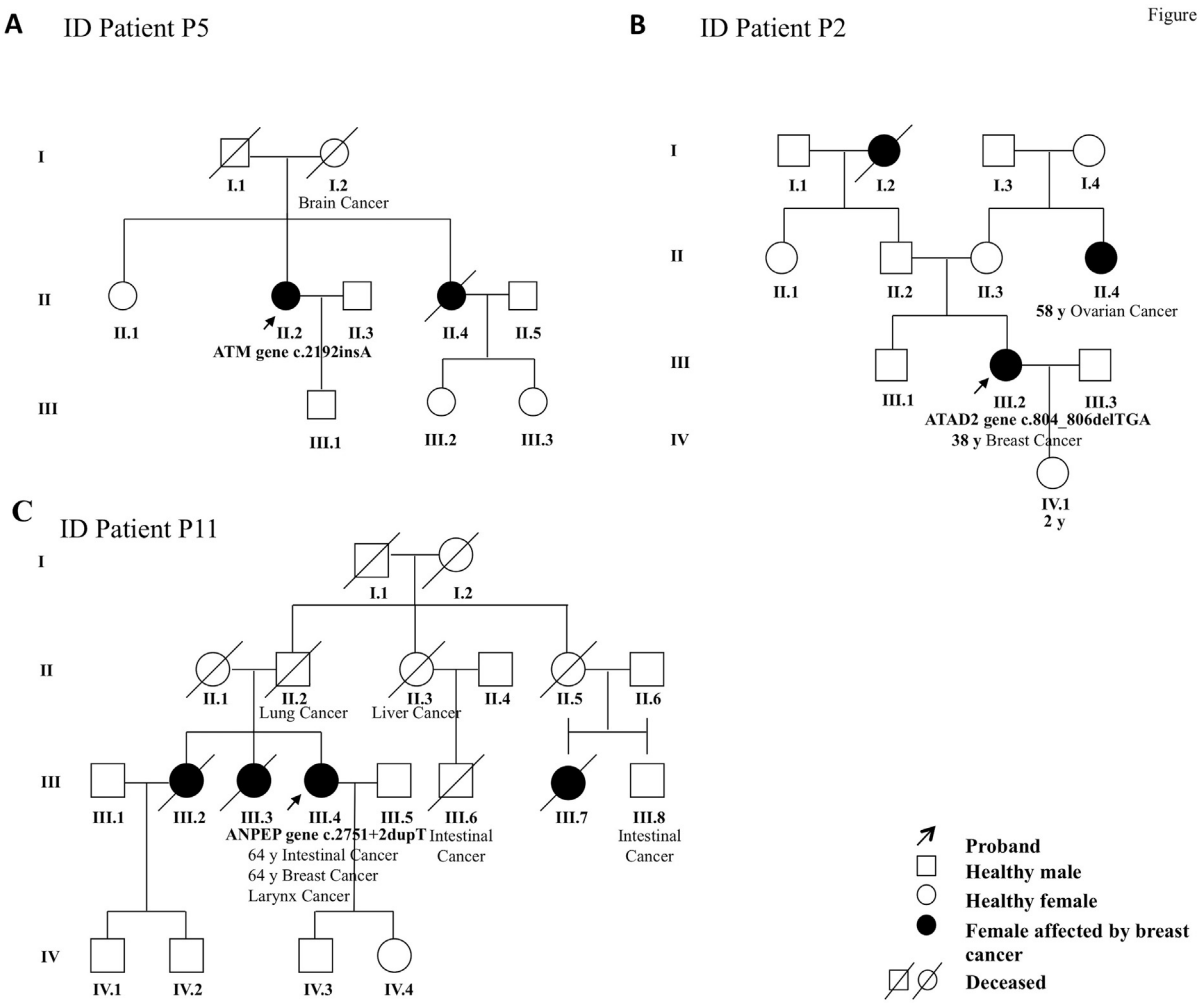


Fig. 3. Pedigrees of patients carrying at least one prioritized variant. Pedigrees of P5, P2 and P11 (A–C). The gene and their variants are reported on the top of each panel. The arrows indicate the proband from each family.

variants prioritized through the bioinformatic pipeline by Sanger sequencing and found that 3 of them are classified as pathogenic in the ClinVar Database (https://www.ncbi.nlm.nih.gov/clinvar/): 1 is the *BRCA2* mutation already identified in the first level analysis in 2 patients (Positive Controls in this study); the remaining 2 occurr one in *TP53* and one in *MUTYH*. The remaining 8 prioritized variants consist of 1 intronic variant, 3 frameshifts, 1 nonsense and 3 missense variants. All prioritized variants were confirmed by Sanger sequencing and are reported in Table 2.

In Patient 9 (P9) we identified a known pathogenic variant (rs121912656) in the *TP53* gene, namely c.734G > A (p.Gly245Asp) that maps on exon 7/11. *TP53* encodes the tumor suppressor gene p53. This variant alters p53 DNA binding site, characterized by 8 conserved residues. P9 was found to be wild-type in the first level analysis of *BRCA1/2*, but the many different types of cancers within her family (breast, uterus, throat, bone, pancreas and lung) and the young age of BC onset strongly suggested a hereditary component. During follow-up, Patient 9 was diagnosed with a new primitive lung cancer (Fig. 1A and B). Therefore, this patient's personal and familial history is compatible with a pathogenic mutation in the *TP53* gene, given the absence of other major risk factors (i.e. smoke, oral contraceptives, ovarian stimulation).

In P17 we found a pathogenic variant in the *MUTYH* gene (rs36053993). *MUTYH* encodes an adenine DNA glycosylase involved in oxidative DNA damage repair. The variant is located on cDNA position c.1145G > A (p.Gly382Asp) in exon 13/16. At protein level the variant mapped on Nudix hydrolase domain. P17 was wild-type at *BRCA1/2* screening analysis. The family history was positive principally for breast and brain cancers (Fig. 1C and D).

Patient 18 was found to carry 2 prioritized variants in the *ATM* and *KLLN* genes. The positions on cDNA are c.7235A > G (p.Asn2412Ser) and c.339_340delAG (p.Tyr113fs), respectively. The *ATM* variant is on exon 49/63 and despite the ClinVar annotation of uncertain significance (rs786203311), it has been associated to hereditary cancer-predisposing syndrome. At protein level, this variant results in an alteration of the FAT (FRAP, ATM and TRRAP) domain. The *KLLN* gene encodes Killin, a 20 kDa nuclear protein that is necessary for p53-induced apoptosis. The *KLLN* variant is a frameshift variant (rs749052307). Although P18 was diagnosed with BC over the age of 40 years old, she was wild-type for *BRCA1/2* at a screening analysis. She was enrolled in this study due to a cluster of tumors in the family (kidney, breast, prostate and uterus). During follow-up, this patient had 2 new primitive cancers: colon cancer and triple negative breast cancer (Fig. 2A).

Patient 23 has 2 previously unreported prioritized variants in *ATM* and *NFE2L2*. In *ATM* there is a stop variant in cDNA position c.6415G > T (p.Glu2139X - rs1339238483) in exon 44/63. At protein level the stop is in the FAT domain. In the *NFE2L2* gene we validated c.761T > A (p.Phe254Tyr) in exon 5/5. P23 resulted wild-type for *BRCA1/2* at screening analysis. This patient was enrolled in the

Table 2
Variants prioritized through the bioinformatic analysis and validated by the Sanger method.

Gene	Patient	HGVS cDNA	HGVS Protein	Reference SNP ID number	ClinVar Status	Number of Pathogenicpredictions	Conserved Site	gnomAD Browser Data[a]	ExAC Data[b]	DANN Score[c]
TP53	P9	c.734G > A	p.Gly245Asp	rs121912656	Pathogenic	7/9	N/A	0.000004061 (1/246,236)	1/121,378	0.9977
BRCA2	P14	c.5796_5797delTA	p.His1932Glnfs	rs80359537	Pathogenic	–	N/A	–	1/121,378	–
BRCA2	P15	c.5796_5797delTA	p.His1932Glnfs	rs80359537	Pathogenic	–	N/A	–	1/121,378	–
MUTYH	P17	c.1145G > A	p.Gly382Asp	rs36053993	Pathogenic	8/9	N/A	0.002958 (1/338)	1/357	0.9985
ATM	P18	c.7235A > G	p.Asn2412Ser	rs786203311	UCV	4/9	Highly conserved	–	–	0.9889
KLLN	P18	c.339_340delAG	p.Tyr113fs	rs749052307	NA	–	N/A	0.001778 (1/703)	–	–
ATAD2	P2	c.804_806delTGA	p.Asp269del	rs539981908	NA	–	N/A	0.001712 (1/584)	1/116,788	–
MSH6	P8	c.2633T > C	p.Val878Ala	rs2020912	Benign	2/9	N/A	0.005119 (1/195)	1/189	0.7259
ANPEP	P11	c.2751+2dupT	–	rs751522046	NA	–	N/A	0.0001056 (1/9469)	–	–
NFE2L2	P23	c.761T > A	p.Phe254Tyr	Unreported	–	3/9	Highly conserved	–	–	0.9886
ATM	P23	c.6415G > T	p.Glu2139X	rs1339238483	NA	2/9	Highly conserved	0.000004091 (1/244,414)	–	0.9968
ATM	P5	c.2192dupA	p.Tyr731fs	rs1478081526	NA	–	N/A	0.000004065 (1/246,020)	–	–

HGVS, Human Genome Variation Society (http://www.hgvs.org).
ClinVar, Clinical Variation database (https://www.ncbi.nlm.nih.gov/clinvar/).
NA, not available, N/A, not applicable.
[a] gnomAD Browser Data: The data set provided spans 123,136 exome sequences and 15,496 whole-genome sequences from unrelated individuals sequenced as part of various disease-specific and population genetic studies.
[b] ExAC Data: The data set provided on spans 60,706 unrelated individuals sequenced as part of various disease-specific and population genetic studies.
[c] DANN Score: DANN is a pathogenicity scoring methodology developed at the University of California, Irvine. It is based on deep neural networks. The value range is 0–1, with 1 given to the variants predicted to be the most damaging [40].

present study due to the early onset of breast cancer and to the cluster of tumors in the family (breast, pancreas, lung and brain) (Fig. 2B).

In P8 we prioritized and validated a variant in the *MSH6* gene. Despite 6 cases of BC and 1 of prostate cancer in the family, P8 was wild-type at *BRCA1/2* screening (Fig. 2C). The *MSH6* variant is in cDNA position c.2633T > C (p.Val878Ala) in exon 4/10, and is classified "benign" in the ClinVar Database (rs2020912). In P5 we validated a hitherto unknown frameshift variant in the *ATM* gene. The variant is in cDNA position c.2192dupA (p.Tyr731fs - rs1478081526). P5 was wild-type at *BRCA1/2* screening and has a family history of breast and brain cancers (Fig. 3A). In P2 we validated a known variant (rs539981908) in the *ATAD2* gene classified as "NA" (not available) for clinical significance (Fig. 3B). This variant, namely c.804_806delTGA (p.Asp269del), causes the in-frame deletion of one amino acidic residue. Finally, P11 had an intronic variant in the *ANPEP* gene (Fig. 3C). This variant, c.2751+2dupT, affects the canonic donor splice site in intron 20 and has been predicted by Human splicing finder (http://www.umd.be/HSF3/) to affect the splicing through the loss of the wild-type donor site and the potential activation of a new cryptic site.

4. Discussion

The molecular analysis reported herein confirmed the pathogenic mutations in *BRCA2*, used as positive controls in 2 patients, and the absence of *BRCA1* and *BRCA2* pathogenic mutations in the other 21 patients. The reliability of this new strategy was confirmed by comparing all the polymorphisms and variants of unknown significance (VUS) identified in *BRCA1* and *BRCA2* with those previously detected in these genes. Using the bioinformatics pipeline followed by Sanger sequencing validation, we prioritized candidate cancer-predisposing variants in the other analyzed genes. The results showed that the two *BRCA2*-mutated patients had no mutations in the other genes, while the other 21 patients had 10 prioritized variants, of which 2 pathogenic mutations, 6 VUS or variants with no functionally known significance, 1 benign and 1 novel variant (Table 2). These variants are distributed differently among the patients and among the genes. In fact, 8/21 patients carry at least one prioritized variant, and just 8/82 genes, not including *BRCA1* and *BRCA2,* present prioritized variants. In particular, as shown in Table 2, we found three different variants in *ATM*, and one variant in each of the following genes *TP53*, *MUTHY*, *KLLN*, *ATAD2*, *MSH6*, *ANPEP* and *NFE2L2*.

We detected 2 pathogenetic mutations in two *BRCA*-negative patients (Patient 9 and Patient 17). One mutation (c.734G > A [p.Gly245Asp]) occurred in the *TP53* gene. This finding supports NCCN guidelines [11] that recommend *TP53* testing in *BRCA*-negative individuals with BC diagnosed before the age of 35 years. Germline mutations in *TP53* have been associated to the rare cancer-predisposing syndrome, Li-Fraumeni [31], which features a wide spectrum of cancers, usually with early onset, also during childhood [32]. The personal and familial history of Patient 9 is compatible with Li-Fraumen-related phenotypes. Thus, our multigene panel testing not only revealed a germline cancer-predisposing mutation in the patient, but also a risk of cancer in the members of this family, who can consequently now be offered appropriate cancer surveillance. The other pathogenic mutation identified in Patient 17 was c.1145G > A (p.Gly382Asp) in the *MUTYH* gene. *MUTYH* encodes a protein involved in DNA repair, whose mutations are usually associated to colorectal cancers, and, as shown by multiple gene panel testing for cancer risk assessment, also to breast cancer [33]. In detail, it has been recently estimated that *MUTYH* monoallelic mutation carriers have an 11% cumulative risk of BC, which confirms that this gene should be included in hereditary cancer risk assessment [34].

Among the other prioritized variants, some could increase the risk of cancer. In fact, we found a nonsense variant in *ATM* (c.6415G > T, p.Glu2139X), an intronic variant in the canonical splicing site of *ANPEP* (c.2751+2dupT), and 2 frameshift variants,

one in *KLLN* (c.339_340delAG, p.Tyr113fs) and one in *ATM* (c.2192dupA, p.Tyr731fs). Of the other three, two are likely pathogenetic since one is a deletion and one is a missense mutation in the *NFE2L2* gene. Notably, the overall frequency of germline mutations in DNA damage repair genes other than the *BRCA* genes in our population screened in a third level center was 28.5%, and *ATM* was the most frequently affected gene (33.3%). Although NGS sensitivity supersedes that of Sanger sequencing, we used Sanger as a second independent method to verify the specific variants of potential interest that we identified by NGS, and found they were all confirmed based on the extent of coverage and the quality of the reads (Q-score $\geq$ 30).

NGS-based screening of a large panel of genes increases the number of novel variants and variants of unknown clinical significance; however, in a research setting, it's important to increase the number of biological variants that may successively be tested *in vitro* and in vivo to assess the clinical importance. Notably, the greater the routine application of gene panel testing, the greater the epidemiological data on the role of specific genes and/or mutations, which in turn will prompt functional studies to clarify their role as tumor-predisposing factors. Therefore, although not regulated by validated guidelines, the presence of a germline pathogenic (or possibly pathogenic) variant in a patient with a personal and/or family history alerts clinicians to the need for a patient monitoring program, also in view of treatment strategies.

The method we used is only one of the possible strategies that can be used to enrich multiple target genes that will be simultaneously sequenced by NGS. The aim of our study was to test this approach since it is based on a picodroplet PCR strategy that ensures a high specificity of PCR reactions and an enhanced capability of multiplexing (up to 20,000 different amplicons/sample), since the amplicons are physically separated within the droplets (that act as individual micro-reactors). As mentioned above, operator time and operator-dependent variability is greatly reduced thanks to the miniaturized microfluidic device and the liquid handler incorporated in the system. Another advantage of this system is that, when working with GC-rich regions, repetitive regions or pseudogenes the separation of individual amplicons into individual droplet micro-reactors facilitates high coverage because it reduces the complexity of these regions thereby ensuring great coverage uniformity [35], which is essential when NGS is applied in the clinical setting. Therefore, the tested droplet PCR combined with NGS is an alternative and reliable method with which to enrich NGS libraries.

Although limited to a small group of patients, our study shows the efficacy of NGS-based approaches in hereditary cancer risk assessment in our patients, based on clinical features and family history, and indicates the need to introduce an enlarged gene panel test in *BRCA*-negative HBOC wild type patients [12–19]. Because molecular tests for the analysis of large genomic regions of interest are becoming more widely used thanks also to more powerful and more sensitive NGS platforms, there is a need to standardize these procedures for clinical applications. The critical point is, of course, the large amounts of data generated and the difficulties related to the interpretation of identified variants. The interpretation of variants of unknown significance is critical for the management of patients and is already a challenge for HBOC counseling, especially in the case of variants that are open to controversial interpretations [18,19,36,37]. Obviously, this issue will be amplified, and will receive much more critical evaluation, consequent to the widespread use of gene panels that include genes whose role in BC onset is not well established, as in the cases reported herein. Functional *in vitro* evaluations combined with molecular analysis and genotype/phenotype correlations, within the same family could shed light on the pathogenetic effects of novel variants [18,19,38,39].

5. Conclusions

The identification of at-risk subjects in affected families is crucial for cancer surveillance and prevention programs. Given the above considerations, multi-gene testing is advisable and should be included in a multistep diagnostic flowchart for the evaluation of HBOC risk, after first-level evaluation of *BRCA* mutation status. The widespread application of this strategy in diagnostic settings will increase our ability to understand a larger fraction of familial cancer cases which at the moment is only 25% based on the two main genes *BRCA1/2*, and thus globally improve the care of cancer patients.

In conclusion, here we describe an affordable methodological strategy with which to identify BC mutations in genes other than the *BRCA* genes. Our strategy revealed a BC-predisposing mutation in patients who were negative at first level *BRCA* testing, thereby improving the diagnostic sensitivity of this molecular test. Identification of predisposing mutations will aid surgical and therapeutic decision-making, particularly in view of the novel targeted drugs that are being developed.

Funding

The study was supported by grants PON03PE_00060_2 and PON03PE_00060_7 (Campania - Bioscience) from the Italian Ministry of University and Research (to F.S.).

Declarations of interests

None.

Ethics approval

All patients gave their written informed consent to the study that was carried out according to the tenets of the Helsinki Declaration and approved by the Istituto Nazionale Tumori - Fondazione G. Pascale Ethics Committee (protocol number 3 of 03/25/2009).

Acknowledgements

The authors thank Jean Ann Gilder (Scientific Communication srl., Naples, Italy) for editing the text, Vittorio Lucignano (CEINGE—Biotecnologie Avanzate) for technical assistance and Antonella Telese (CEINGE—Biotecnologie Avanzate) for her support in the setting of some experimental procedures.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.aca.2018.09.032.

References

[1] V. Precone, V. Del Monaco, M.V. Esposito, F.D.E. De Palma, A. Ruocco, F. Salvatore, V. D'Argenio, Cracking the code of human diseases using next-generation sequencing: applications, challenges, and perspectives, BioMed Res. Int. 2015 (2015), https://doi.org/10.1155/2015/161648.

[2] I. Bernardis, L. Chiesi, E. Tenedini, L. Artuso, A. Percesepe, V. Artusi, M.L. Simone, R. Manfredini, M. Camparini, C. Rinaldi, A. Ciardella, C. Graziano, N. Balducci, A. Tranchina, G.M. Cavallini, A. Pietrangelo, V. Marigo, E. Tagliafico, Unravelling the complexity of inherited retinal dystrophies molecular testing: added value of targeted next-generation sequencing, BioMed Res. Int. 2016 (2016), https://doi.org/10.1155/2016/6341870.

[3] M.D. Mruthyunjaya, A. Chapla, A.H. Shyamasunder, D. Varghese, M. Varshney, J. Paul, M. Inbakumari, F. Christina, R.T. Varghese, K.A. Kuruvilla, T.V. Paul, R. Jose, A. Regi, J. Lionel, L. Jeyaseelan, J. Mathew, N. Thomas, Comprehensive maturity onset diabetes of the young (MODY) gene screening in pregnant women with diabetes in India, PLoS One 12 (2017) 1–15, https://doi.org/

10.1371/journal.pone.0168656.
[4] I. Bock, K. Németh, K. Pentelényi, P. Balicza, A. Balázs, M.J. Molnár, V. Román, J. Nagy, G. Lévay, J. Kobolák, A. Dinnyés, Targeted next generation sequencing of a panel of autism-related genes identifies an EHMT1 mutation in a Kleefstra syndrome patient with autism and normal intellectual performance, Gene 595 (2016) 131–141, https://doi.org/10.1016/j.gene.2016.09.027.
[5] V. D'Argenio, G. Frisso, V. Precone, A. Boccia, A. Fienga, G. Pacileo, G. Limongelli, G. Paolella, R. Calabro, F. Salvatore, DNA sequence capture and next-generation sequencing for the molecular diagnosis of genetic cardiomyopathies, J. Mol. Diagn. 16 (2014) 32–44, https://doi.org/10.1016/j.jmoldx.2013.07.008.
[6] A. Paul, S. Paul, The breast cancer susceptibility genes (BRCA) in breast and ovarian cancers, Front Biosci. (Landmark Ed). 19 (2014) 605–618, https://doi.org/10.1021/ac901991x.
[7] C. Engel, C. Fischer, Breast cancer risks and risk prediction models, Breast Care 10 (2015) 7–12, https://doi.org/10.1159/000376600.
[8] L.C. Hartmann, N.M. Lindor, The role of risk-reducing surgery in hereditary breast and ovarian cancer, N. Engl. J. Med. 374 (2016) 454–468, https://doi.org/10.1056/NEJMra1503523.
[9] Q. Wang, Cancer predisposition genes: molecular mechanisms and clinical impact on personalized cancer care: examples of Lynch and HBOC syndromes, Acta Pharmacol. Sin. 37 (2016) 143–149, https://doi.org/10.1038/aps.2015.89.
[10] M. Nunziato, F. Starnone, B. Lombardo, M. Pensabene, C. Condello, F. Verdesca, C. Carlomagno, S. De Placido, L. Pastore, F. Salvatore, V. D'Argenio, Fast detection of a BRCA2 large genomic duplication by next generation sequencing as a single procedure: a case report, Int. J. Mol. Sci. 18 (2017) 2487, https://doi.org/10.3390/ijms18112487.
[11] M.B. Daly, R. Pilarski, M. Berry, S.S. Buys, M. Farmer, S. Friedman, J.E. Garber, N.D. Kauff, S. Khan, C. Klein, CE NCCN Guidelines ® Insights genetic/Familial High-Risk Assessment : breast and Ovarian, Version 2. 2017 Featured Updates to the NCCN Guidelines, 15, 2017, pp. 989–997, https://doi.org/10.6004/jnccn.2017.0132.
[12] H. Kobayashi, S. Ohno, Y. Sasaki, M. Matsuura, Hereditary breast and ovarian cancer susceptibility genes (Review), Oncol. Rep. 30 (2013) 1019–1029, https://doi.org/10.3892/or.2013.2541.
[13] S.A. Narod, Breast cancer in young women, Nat. Rev. Clin. Oncol. 9 (2012) 460–470, https://doi.org/10.1038/nrclinonc.2012.102.
[14] S.S. Buys, J.F. Sandbach, A. Gammon, G. Patel, J. Kidd, K.L. Brown, L. Sharma, J. Saam, J. Lancaster, M.B. Daly, A study of over 35,000 women with breast cancer tested with a 25-gene panel of hereditary cancer genes, Cancer 123 (2017) 1721–1730, https://doi.org/10.1002/cncr.30498.
[15] E. Castellanos, B. Gel, I. Rosas, E. Tornero, S. Santín, R. Pluvinet, J. Velasco, L. Sumoy, J. Del Valle, M. Perucho, I. Blanco, M. Navarro, J. Brunet, M. Pineda, L. Feliubadaló, G. Capellá, C. Lázaro, E. Serra, A comprehensive custom panel design for routine hereditary cancer testing: preserving control, improving diagnostics and revealing a complex variation landscape, Sci. Rep. 7 (2017) 1–12, https://doi.org/10.1038/srep39348.
[16] T. Judkins, B. Leclair, K. Bowles, N. Gutin, J. Trost, J. McCulloch, S. Bhatnagar, A. Murray, J. Craft, B. Wardell, M. Bastian, J. Mitchell, J. Chen, T. Tran, D. Williams, J. Potter, S. Jammulapati, M. Perry, B. Morris, B. Roa, K. Timms, Development and analytical validation of a 25-gene next generation sequencing panel that includes the BRCA1 and BRCA2 genes to assess hereditary cancer risk, BMC Canc. 15 (2015) 215, https://doi.org/10.1186/s12885-015-1224-y.
[17] L. Castéra, S. Krieger, A. Rousselin, A. Legros, J.J. Baumann, O. Bruet, B. Brault, R. Fouillet, N. Goardon, O. Letac, S. Baert-Desurmont, J. Tinat, O. Bera, C. Dugast, P. Berthet, F. Polycarpe, V. Layet, A. Hardouin, T. Frébourg, D. Vaur, Next-generation sequencing for the diagnosis of hereditary breast and ovarian cancer using genomic capture targeting multiple candidate genes, Eur. J. Hum. Genet. 22 (2014) 1305–1313, https://doi.org/10.1038/ejhg.2014.16.
[18] E.F. Cobain, K.J. Milliron, S.D. Merajver, Updates on breast cancer genetics: clinical implications of detecting syndromes of inherited increased susceptibility to breast cancer, Semin. Oncol. 43 (2016) 528–535, https://doi.org/10.1053/j.seminoncol.2016.10.001.
[19] F.C. Nielsen, T. Van Overeem Hansen, C.S. Sørensen, Hereditary breast and ovarian cancer: new genes in confined pathways, Nat. Rev. Canc. 16 (2016) 599–612, https://doi.org/10.1038/nrc.2016.72.
[20] V. D'Argenio, M.V. Esposito, A. Telese, V. Precone, F. Starnone, M. Nunziato, P. Cantiello, M. Iorio, E. Evangelista, M. D'Aiuto, A. Calabrese, G. Frisso, G. D'Aiuto, F. Salvatore, The molecular analysis of BRCA1 and BRCA2: next-generation sequencing supersedes conventional approaches, Clin. Chim. Acta 446 (2015) 221–225, https://doi.org/10.1016/j.cca.2015.03.045.
[21] V. D'Argenio, E. Notomista, M. Petrillo, P. Cantiello, V. Cafaro, V. Izzo, B. Naso, L. Cozzuto, L. Durante, L. Troncone, G. Paolella, F. Salvatore, A. Di Donato, Complete sequencing of Novosphingobium sp. PP1Y reveals a biotechnologically meaningful metabolic pattern, BMC Genom. 15 (2014) 384, https://doi.org/10.1186/1471-2164-15-384.
[22] R.K. Patel, M. Jain, NGS QC toolkit: a toolkit for quality control of next generation sequencing data, PLoS One 7 (2012), https://doi.org/10.1371/journal.pone.0030619.
[23] H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform, Bioinformatics 25 (2009) 1754–1760, https://doi.org/10.1093/bioinformatics/btp324.
[24] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, The sequence alignment/map format and SAMtools, Bioinformatics 25 (2009) 2078–2079, https://doi.org/10.1093/bioinformatics/btp352.
[25] M. a. DePristo, E. Banks, R.E. Poplin, K.V. Garimella, J.R. Maguire, C. Hartl, a. a. Philippakis, G. del Angel, M. a Rivas, M. Hanna, A. McKenna, T.J. Fennell, a. M. Kernytsky, a. Y. Sivachenko, K. Cibulskis, S.B. Gabriel, D. Altshuler, M.J. Daly, A framework for variation discovery and genotyping using next- generation DNA sequencing data, Nat. Genet. 43 (2011) 491–498, https://doi.org/10.1038/ng.806.A.
[26] G.R. Abecasis, A. Auton, L.D. Brooks, M.A. DePristo, R.M. Durbin, R.E. Handsaker, H.M. Kang, G.T. Marth, G.A. McVean, 1000 Genomes Project Consortium, an integrated map of genetic variation from 1, 092 human genomes, Nature 491 (2013) 56–65, https://doi.org/10.1038/nature11632.An.
[27] H. Zhao, Z. Sun, J. Wang, H. Huang, J.P. Kocher, L. Wang, CrossMap: a versatile tool for coordinate conversion between genome assemblies, Bioinformatics 30 (2014) 1006–1007, https://doi.org/10.1093/bioinformatics/btt730.
[28] P. Cingolani, A. Platts, L.L. Wang, M. Coon, T. Nguyen, L. Wang, S.J. Land, X. Lu, D.M. Ruden, A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff, Fly (Austin) 6 (2012) 80–92, https://doi.org/10.4161/fly.19695.
[29] X. Liu, X. Jian, E. Boerwinkle, dbNSFP: a lightweight database of human nonsynonymous SNPs and their functional predictions, Hum. Mutat. 32 (2011) 894–899, https://doi.org/10.1002/humu.21517.
[30] X. Liu, C. Wu, C. Li, E. Boerwinkle, dbNSFP v3.0: a one-stop database of functional predictions and annotations for human nonsynonymous and splice-site SNVs, Hum. Mutat. 37 (2016) 235–241, https://doi.org/10.1002/humu.22932.
[31] J. Kamihara, H.Q. Rana, J.E. Garber, Germline TP53 mutations and the changing landscape of Li-Fraumeni syndrome, Hum. Mutat. 35 (2014) 654–662, https://doi.org/10.1002/humu.22559.
[32] J.M. Valdez, K.E. Nichols, C. Kesserwan, Li-Fraumeni syndrome: a paradigm for the understanding of hereditary cancer predisposition, Br. J. Haematol. 176 (2017) 539–552, https://doi.org/10.1111/bjh.14461.
[33] P.-H. Lin, W.-H. Kuo, A.-C. Huang, Y.-S. Lu, C.-H. Lin, S.-H. Kuo, M.-Y. Wang, C.-Y. Liu, F.T.-F. Cheng, M.-H. Yeh, H.-Y. Li, Y.-H. Yang, Y.-H. Hsu, S.-C. Fan, L.-Y. Li, S.-L. Yu, K.-J. Chang, P.-L. Chen, Y.-H. Ni, C.-S. Huang, Multiple gene sequencing for risk assessment in patients with early-onset or familial breast cancer, Oncotarget 7 (2016) 8310–8320, https://doi.org/10.18632/oncotarget.7027.
[34] A.K. Win, J.C. Reece, J.G. Dowty, D.D. Buchanan, M. Clendenning, C. Rosty, M.C. Southey, J.P. Young, S.P. Cleary, H. Kim, M. Cotterchio, F.A. Macrae, K.M. Tucker, J.A. Baron, T. Burnett, L. Le Marchand, G. Casey, R.W. Haile, P.A. Newcomb, S.N. Thibodeau, J.L. Hopper, S. Gallinger, I.M. Winship, N.M. Lindor, M.A. Jenkins, Risk of extracolonic cancers for people with biallelic and monoallelic mutations in MUTYH, Int. J. Canc. 139 (2016) 1557–1563, https://doi.org/10.1002/ijc.30197.
[35] A. Gnirke, A. Melnikov, J. Maguire, P. Rogov, E.M. LeProust, W. Brockman, T. Fennell, G. Giannoukos, S. Fisher, C. Russ, S. Gabriel, D.B. Jaffe, E.S. Lander, C. Nusbaum, Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing, Nat. Biotechnol. (2009), https://doi.org/10.1038/nbt.1523.
[36] E.B. Eccles, G. Mitchell, A.N.A. Monteiro, R. Schmutzler, F.J. Couch, A.B. Spurdle, E.B. Gómez-García, R. Driessen, N.M. Lindor, M.J. Blok, P. Moller, M. de la Hoya, T. Pal, S. Domchek, K. Nathanson, C. Van Asperen, O. Diez, K. Rheim, D. Stoppa-Lyonnet, M. Parsons, D. Goldgar, BRCA1 and BRCA2 genetic testing-pitfalls and recommendations for managing variants of uncertain clinical significance, Ann. Oncol. 26 (2015) 2057–2065, https://doi.org/10.1093/annonc/mdv278.
[37] V. D'Argenio, M.V. Esposito, J.A. Gilder, G. Frisso, F. Salvatore, Should a BRCA2 stop codon human variant, usually considered a polymorphism, be classified as a predisposing mutation? Cancer 120 (2014) 1594–1595, https://doi.org/10.1002/cncr.28605.
[38] M.V. Esposito, M. Nunziato, F. Starnone, A. Telese, A. Calabrese, G. D'Aiuto, P. Pucci, M. D'Aiuto, F. Baralle, V. D'Argenio, F. Salvatore, A novel pathogenic BRCA1 splicing variant produces partial intron retention in the mature messenger RNA, Int. J. Mol. Sci. 17 (2016), https://doi.org/10.3390/ijms17122145.
[39] M. Rodríguez-Balada, B. Roig, L. Martorell, M. Melé, M. Salvat, E. Vilella, J. Borràs, J. Gumà, In silico, in vitro and case-control analyses as an effective combination for analyzing BRCA1 and BRCA2 unclassified variants in a population-based sample, Canc. Genet. 209 (2016) 487–492, https://doi.org/10.1016/j.cancergen.2016.09.003.
[40] D. Quang, Y. Chen, X. Xie, DANN: a deep learning approach for annotating the pathogenicity of genetic variants, Bioinformatics (2015), https://doi.org/10.1093/bioinformatics/btu703.