Influence of BRCA1 Germline Mutations in the Somatic Mutational Burden of Triple-Negative Breast Cancer

Elisa Napolitano Ferreira*,1, Rafael Canfield Brianese*,1, Renan Valeris Bueno de Almeida*, Rodrigo Duarte Drummond*, Jorge Estefano de Souza†,§, Israel Tojal da Silva†,§, Sandro José de Souza†,§ and Dirce Maria Carraro*,¶

*CIPE, International Research Center, A.C. Camargo Cancer Center, São Paulo, SP, Brazil; †Instituto Metrópole Digital, Federal University of Rio Grande do Norte, Natal, RN, Brazil; ‡Bioinformatics Multidisciplinary Environment (BioME), Federal University of Rio Grande do Norte, Natal, RN, Brazil; §Brain Institute, Federal University of Rio Grande do Norte, Natal, RN, Brazil; ¶National Institute of Science and Technology in Oncogenomics and Therapeutic Innovation, São Paulo, SP, Brazil

Abstract

The majority of the hereditary triple-negative breast cancers (TNBCs) are associated with BRCA1 germline mutations. Nevertheless, the understanding of the role of BRCA1 deficiency in the TNBC tumorigenesis is poor. In this sense, we performed whole-exome sequencing of triplet samples (leucocyte, tumor, and normal-adjacent breast tissue) for 10 cases of early-onset TNBC, including 5 hereditary (with BRCA1 germline pathogenic mutation) and 5 sporadic (with no BRCA1 or BRCA2 germline pathogenic mutations), for assessing the somatic mutation repertoire. Protein-affecting somatic mutations were identified for both mammary tissues, and Ingenuity Pathway Analysis was used to investigate gene interactions. BRCA1 and RAD51C somatic promoter methylation in tumor samples was also investigated by bisulfite sequencing. Sporadic tumors had higher proportion of driver mutations (≥25% allele frequency) than BRCA1 hereditary tumors, whereas no difference was detected in the normal breast samples. Distinct gene networks were obtained from the driver genes in each group. The Cancer Genome Atlas data analysis of TNBC classified as hereditary and sporadic reinforced our findings. The data presented here indicate that in the absence of BRCA1 germline mutations, a higher number of driver mutations are required for tumor development and that different defective processes are operating in the tumorigenesis of hereditary and sporadic TNBC in young women.

Translational Oncology (2019) 12, 1453–1460

Background

Triple-negative breast cancer (TNBC) encompasses a subgroup of breast tumors that are negative for estrogen and progesterone receptors expression and negative for overexpression/amplification of the human epidermal growth factor receptor 2 (HER2). TNBC accounts for about 15% of all breast cancer cases and presents poor outcome due to its aggressive behavior and lack of targeted therapy [1]. Paradoxically, TNBCs have higher response rates to neoadjuvant chemotherapy when compared to other subtypes of breast cancer, and patients who achieve complete pathological response present long-term good prognosis. However, a considerable part of the patients has residual disease and, therefore, poor survival rates [2,3]. This heterogeneous clinical behavior is reflected at the molecular level, and based on gene expression and mutation analysis, researchers have indicated the existence of different tumors subclasses within the TNBC subgroup [4,5]. However, the translation of all these molecular information into clinical practice remains limited.

Address all correspondence to: Dirce Maria Carraro, Genomics and Molecular Biology Group, International Research Center /CIPE, A. C. Camargo Cancer Center. Rua Taguá 440, 1º andar. CEP: 01508-010 São Paulo, SP, Brazil. E-mail: dirce.carraro@accamargo.org.br

†Both authors contributed equally to this manuscript

Received 22 May 2019; Revised 17 July 2019; Accepted 18 July 2019

© 2019 The Authors. Published by Elsevier Inc. on behalf of Neoplasia Press, Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1936-5233/19
https://doi.org/10.1016/j.tranon.2019.07.016
We and others have reported that *BRCA1* is the most frequently mutated gene in women with hereditary breast cancer [6–8], and the majority (60%-80%) of *BRCA1* carriers who develop breast cancer have TNBC [6,9,10]. Also, about 10% of all TNBC patients are *BRCA1* mutation carriers, and this frequency nearly doubles in cases diagnosed before 40 years old [11–13]. Thus, early onset TNBC comprises a significant proportion of hereditary cancer, mainly by germline mutations in *BRCA1*. Besides, somatic hypermethylation of the *BRCA1* promoter region has been detected in about 20% of the TNBC cases [13]. Hence, *BRCA1* inactivation, considering both germline mutations and gene promoter hypermethylation, is present in nearly 30% of all TNBCs and is accentuated (~60%) in tumors diagnosed in younger patients (<40 year of age), as demonstrated in a recent study from our group [13]. These evidences suggest that deficiency of *BRCA1* gene probably triggers the development of TNBC in hereditary and also in, at least a part of, the sporadic tumor group, especially in the early-onset patients.

*BRCA1* gene encodes a multifunctional protein that holds a key function in the maintenance of genomic stability [14]. *BRCA1* protein is essential for DNA double-strand breaks repair through homologous recombination (HR), a high-fidelity repair process that uses the sister chromatid as a template for DNA repair [15]. Thus, *BRCA1* loss of function could predispose cells to errors in DNA replication leading to accumulation of somatic mutations that would lead to tumor development. Yet, how this deficiency modulates the mutational landscape and confers proliferative advantages is poorly understood. Moreover, recently it was proposed that somatic hypermethylation of RAD51C gene promoter leads to a mutation signature similar to what is observed in *BRCA1*-deficient tumors, suggesting a wider impairment of HR pathway in breast cancers. Henceforth, the investigation of the molecular mechanisms that underlies TN tumorigenesis in patients with proficient *BRCA1* could reveal affected biological pathways and potentially suggest therapeutic targets for this subtype of tumor.

Here, for a better understanding of the role of *BRCA1* deficiency in the tumor mutation burden associated to TN tumorigenesis, we examined the somatic mutation repertoire and promoter methylation (*BRCA1* and RAD51C) of tumor and normal adjacent mammary tissue of two groups of TNBC: hereditary *BRCA1*-impaired and sporadic *BRCA1*/2-proficient. We investigated variant allele frequency as an indication of driver events in the tumorigenesis and also enrichment of mutation signatures to provide a set of putative biological process affected by somatic variants occurring in allele frequencies ≥5% in the tumor or normal regions with a minimum coverage of 10× for leukocyte sample from each patient. Somatic mutations were selected using the criterion of minimum variant frequency of 5% in the tumor or normal samples. We selected the variants that occur within the coding sequence and that affect protein sequence—missense, nonsense, splice site alterations, and indels—and that are not reported in dbSNP version 138 [19] or reported with a minor allele frequency less than 1%. Somatic alterations were compared to dbNSFP version 2.4 [20], Catalogue of Somatic Mutations in Cancer v69 [21], 1000genomes [22], NHLBI GO Exome Sequencing Project version ESP6500SI-V2 (http://evs.gs.washington.edu/EVS/), and HapMap [23].

**Methods**

**Samples**

We selected patients from the A. C. Camargo Cancer Center Tumor Biobank that were diagnosed with TNBC at young age (≤40) and that presented available frozen tissue from tumor, paired normal sample, and leukocyte. These patients have been screened for *BRCA1* and *BRCA2* germline mutations in previous studies of our group [6,13] and were classified as sporadic *BRCA1*/2 wild-type (sporadic) or hereditary *BRCA1*-mutated (*BRCA1* hereditary).

**DNA Extraction**

Solid tissue and blood samples were collected following the technical and ethical procedures of A. C. Camargo Tumor Bank [16]. DNA was extracted using QIASymphony DNA Mini kit (QIAGEN, Hilden, Germany), following standard procedures.

**Whole-Exome Sequencing**

Whole-exome sequencing of the tumor, normal, and leukocyte DNA samples was performed using the TargetSeq Exome Enrichment Kit (Life Technologies) or the Nextera Rapid Capture Exome (Illumina) followed by paired-end sequencing at Solid 5500xl System (Life Technologies) or NextSeq 500 (Illumina), respectively. Sequencing reads from Solid 5500xl System were mapped to the reference genome (GRCh37/hg19) with LifeScope Genomic Analysis Software v2.5.1. Sequencing reads from NextSeq 500 were mapped to the reference genome (GRCh37/hg19) with TMAP 4.2.18. Genetic variants (SNVs and indels) were identified following the GATK protocol vs3.2-2-gec30cee [17] and annotated with SnapEff version 3.5d (build 2014/03/05) [18]. Variants were identified in regions with a minimum coverage of 20× for tissue samples (tumor and normal) and a minimum coverage of 10× for leukocyte sample from each patient. Somatic mutations were selected using the criterion of minimum variant frequency of 5% in the tumor or normal samples. We selected the variants that occur within the coding sequence and that affect protein sequence—missense, nonsense, splice site alterations, and indels—and that are not reported in dbSNP version 138 [19] or reported with a minor allele frequency less than 1%. Somatic alterations were compared to dbNSFP version 2.4 [20], Catalogue of Somatic Mutations in Cancer v69 [21], 1000genomes [22], NHLBI GO Exome Sequencing Project version ESP6500SI-V2 (http://evs.gs.washington.edu/EVS/), and HapMap [23].

**Bisulfite Sequencing**

*BRCA1* and RAD51C gene promoter methylation was investigated by bisulfite next-generation sequencing in the Ion Proton platform. Tumor and adjacent normal tissue DNA samples were bisulfite converted using the EZ DNA Methylation Gold kit (Zymo Research). The promoter region of both genes (chr17:41277324–41277487 for *BRCA1* and chr17:56769768–56770061 for RAD51C; see Supplemental Table S6) were PCR-amplified using the Multiplex PCR Plus kit (Qiagen). The amplified products were used for library preparation with the Ion Plus Fragment kit (Thermo Fisher). Samples were considered hypermethylated upon reaching ≥16.1% mean methylation level, the same cutoff determined by maximally selected rank statistics approach by our group [13].

**Validation in The Cancer Genome Atlas (TCGA) Data**

We evaluated germline and somatic mutation data from WES of 155 TNBC cases diagnosed at any age from TCGA and classified them as *BRCA1* hereditary (10 cases) and sporadic (125 cases) as previously mentioned (see section “Samples”). Somatic variants that were identified by Mutect and Muse variant calling software were selected using the same criteria stated before (see section “Whole-Exome Sequencing”).

**Ingenuity Pathway Analysis (IPA)**

We evaluated the functional analysis of the driver genes—genes affected by somatic variants occurring in allele frequencies ≥25%—using the core analysis of IPA software (Qiagen, Hilden, Germany). We only considered the pathways with a score ≥20.

**Signatures of Mutational Processes**

We investigated the patterns in base substitutions and context for the identification of signatures of mutational processes characterized in previous studies [24–28] in the somatic mutations identified in our
Only the signatures curated by the Catalogue of Somatic Mutations in Cancer and observed in breast cancer were considered. The significance of the contribution of each signature on the mutational load of each sample was checked by a likelihood ratio test: the maximum likelihood for the sample was computed with all the signatures present in the model and with one signature excluded, and the significance of the ratio of those likelihoods was adopted as the significance of the excluded signature contribution. P values were corrected by the Benjamini-Hochberg method.

**Statistical Analysis**
Fisher’s exact test and ANOVA with Tukey’s post hoc test were used for comparing the variables with 5% level of significance in the GraphPad Prism 5.04 software.

**Results and Discussion**
First, we investigated the mutational landscape of the two different groups, sporadic and hereditary, of TNBC by evaluating the somatically acquired alterations of both tumors and paired normal breast tissue from five patients harboring BRCA1 germline mutation (BRCA1 Hereditary) and five patients that are BRCA1/2 wild-type (Sporadic) using whole-exome sequencing of triplet samples (blood, tumor, and normal breast tissue) (Supplementary Tables S1-3).

As expected, both groups of tumors exhibited higher number of somatic alterations than the corresponding normal mammary tissue: 34.1 and 18.7 somatic mutations on average per tumor and normal sample, respectively. However, only the comparison between BRCA1 hereditary tumors versus the paired normal breast tissues reached statistical significance (Figure 1, A; P value = 0.0287, ANOVA, post hoc Tukey’s test). Additionally, although a trend towards higher number of somatic mutations in BRCA1 hereditary tumors was noted, no significant difference was detected by comparing the number of somatic mutations from both BRCA1 hereditary and sporadic groups, either between tumors or between normal mammary tissues.

Next, to have insights on the somatic evolutionary process of both TNBC groups, we assessed the potential driver genes that might be underlying the tumorigenic process of these two types of TNBC. Assuming that driver mutations, the founding events of carcinogenesis, tend to occur at higher frequencies in tumor cells, we classified the somatic mutations in two groups: high-frequency mutation (frequency ≥ 25%) and low-frequency mutation (frequency < 25%), and named them as drivers and passengers mutations, respectively.
We observed a higher frequency of driver mutations in tumors compared to normal mammary tissues (Figure 1, B) even when different cutoff of allele frequencies was tested for discriminating drivers from passengers variants (Supplementary Figure 1). Additionally, sporadic TNBC showed significantly higher number of driver mutations than hereditary TNBC (BRCA1 hereditary) (P value = <0.0001, Fisher’s test), whereas no difference was detected in the normal samples in the two groups (Figure 1, B). Moreover, we investigated the proportion of variants according to different allele frequencies rates between the groups and observed a progressive increase in the proportion of variants at higher allele frequency in the sporadic group, which was exclusively observed in tumor tissue (Figure 2). These findings suggest that, in the absence of BRCA1 germline mutations, a higher number of driver mutations have to accumulate for acquiring the malignant phenotype.

In order to validate these findings in a larger cohort of patients unselected for age at diagnosis, we investigated WES data from TCGA. We classified the samples as BRCA1 hereditary and sporadic and investigated the somatic mutation burden in each group using a similar approach as before. BRCA1 hereditary tumors exhibited higher number of somatic mutations (Figure 3, A; P = .0384, Mann-Whitney test). Nevertheless, the proportion of driver alterations was higher in the sporadic group (Figure 3, B; P < .0001, Fisher’s test), supporting our previous hypothesis. Interestingly, in spite of the fact

Figure 2. Proportion of somatic variants according to different allele frequency levels between BRCA1 Hereditary and Sporadic in tumor tissue. Passenger Mutation (≥ 5% and < 25%); Driver mutation (≥ 25%).

Figure 3. A. Somatic acquired alterations of tumors (TNBC) from patients harboring BRCA1 germline mutations (BRCA1 Hereditary) and in patients that are BRCA1/2 wild-type (Sporadic) obtained by whole-exome sequencing data from The Cancer Genome Atlas (TCGA) (*) p-value=0.0385; Mann-Whitney Test. B. Distribution of driver (blue) and passenger (green) alterations in each group of tumors from BRCA1 Hereditary and Sporadic TNBC patients from TCGA (***; p-value < .0001; Fisher’s Test.)
that, in our group of TNBC samples, no significant difference in the absolute number of mutations was detected comparing BRCA1 hereditary and sporadic TNBC, data from TCGA showed higher number of mutation in BRCA1 hereditary TNBC. Indeed, we observed a trend towards higher number of mutations in BRCA1 hereditary tumors, and this could indicate that perhaps this difference did not reach statistical significance in our cohort probably due to the low number of samples. This information is also in consonance with the previous data published by Wen and Leong [29] reporting that BRCA1-deficient breast tumors, irrespective of subtype, display more neoantigen formation and immunogenic phenotype than the BRCA2-mutated or BRCA1/2-proficient breast tumors, suggesting a role for immunotherapy in BRCA1-mutated breast cancers. In fact, the potential for immunotherapy in TNBC is currently being tested [30], but the benefits, especially for early stages and in BRCA1-deficient tumors, are still to be determined.

Additionally, we assessed the BRCA1 and RAD51C somatic promoter methylation in all tumors and in most paired normal adjacent tissue samples (Supplementary Table S7). All BRCA1 hereditary tumors exhibited negative BRCA1 promoter methylation, reinforcing the evidence that BRCA1 germline mutations and promoter hypermethylation are mutually exclusive events, as we previously reported [13]. Moreover, most tumors (hereditary and sporadic) showed higher BRCA1 promoter hypermethylation in comparison with their normal breast tissue counterparts, which were all classified as negative for gene promoter hypermethylation. In spite

Figure 4. Mutational Signatures in BRCA1 Hereditary and Sporadic TNBC.
of the fact that all sporadic tumors showed gene promoter hypermethylation, two of them had >80% mean methylation level but with no clear association between high level of BRCA1 promoter hypermethylation and enrichment for signature 3. Also, the levels of RAD51C promoter methylation were very low across all samples. In a recent study, Polak et al. showed that RAD51C promoter hypermethylation, although associated with signature 3 (associated with HR deficiency), is a rare event in breast cancer. Although limited by sample size, our study agrees with these data as no tumor was identified as having high levels of RAD51C promoter hypermethylation.

We also investigated the somatic mutations for the identification of signatures indicative of mutation processes active during tumor development and progression (Figure 4). Signature 3 — which has been associated with failure in double-strand break repair by HR, as mentioned — was more observed in the BRCA1 hereditary group (4/5) than in sporadic tumors (2/5), when considering a \( P \) value \( \geq 0.1 \). Also, hypermethylation of both BRCA1 and RAD51C promoter was not associated with mutation signatures. Moreover, signatures 1 and 26, believed to be associated with spontaneous deamination of 5-methylcytosine (a common feature in all cancer types) and with defects in DNA mismatch repair, respectively, were equally observed in both groups.

Finally, to uncover the affected biological pathways that outline the tumorigenic process of both TNBC groups of young women, IPA was performed. The list of driver genes affected by mutations in the BRCA1 hereditary group depicted genes highly interconnected with TP53 pathway (Figure 5, A). This network suggested a dysregulation of mechanisms involved with cell death and survival, and embryonic and organismal development (Supplementary Tables 4-5), which are processes associated with cell cycle disturbance. The driver genes in the sporadic group are interconnected with the NfkB complex and histone H3, in addition to TP53, and are related to organismal injury and abnormalities, respiratory disease, and cellular compromise (Figure 5, B), which are mechanisms associated with metabolism disturbance.

Overall, in this work, we propose a portrait of somatic mutation arising in TNBC samples under the strong influence of germline mutations in BRCA1 gene. Although no statistically significant difference was observed in the absolute number of mutation per sample between the BRCA1 hereditary and sporadic groups in our cohort, in the TCGA data analysis, we were able to show that sporadic tumors exhibit, proportionally, more driver mutations than BRCA1 hereditary tumors, at least considering those that occur in the beginning of the tumorigenic process. Our criterion for defining mutations as “drivers” (allele frequency \( \geq 25\% \)) was in terms of value, undoubtedly arbitrary. However, in a serial analysis using distinct cutoffs encompassing values from 15% to 30%, the same pattern of driver and passenger proportion, i.e., significantly higher number of driver than passenger mutations detected in sporadic tumors, was perceived and statistically supported (Supplementary Figure 1). These findings reinforce our suggestion that TNBC tumors that arise in sporadic patients need to acquire and accumulate a higher number of driver mutations for the tumorigenic process to take place, as opposed to BRCA1 hereditary patients, which already have an important inherited driver event. Hence, this finding reinforces the strong effect of BRCA1 loss of function in TNBC tumorigenesis and suggests that different biological processes are active in hereditary, prompted by BRCA1 mutations, and sporadic TNBC diagnosed in young women.

The IPA pointed out the involvement of TP53 in TNBC regardless of the mutation status of BRCA1, where both our networks showed TP53 as a central node. TP53 is highly recognized as frequently mutated in solid tumors, and especially in TNBC, it has
been proposed as a potential therapeutic target [31] Mutations in TP53 can result in dysregulation of important cellular processes such as cell cycle and apoptosis, contributing to the tumorigenic processes. In addition, driver genes of sporadic TNBC showed involvement of biological processes of gene expression regulation by affecting promoter regulation via NFκB complex and epigenetics modulation via histone H3 complex. Recently, activation of the canonical NFκB pathway has been positively correlated to chemotherapy resistance and poor prognosis of TNBC patients [32].

Although limited by the sample size, our study was supported in an independent dataset and was able to demonstrate the existence of differences at the mutational level between tumor from patients with hereditary BRCA1 hereditary TNBC and those with sporadic TNBC. Larger analysis exploring the driving events of cancer under the influence of loss of BRCA1 function may lead to a better understanding on the emergence of TNBC and contribute to the identification of clinically useful biomarkers in these two groups of TNBC.

**Conclusions**

The lower number of driver mutations detected in hereditary BRCA1-related TNBC might reflect a shorter multistep process for tumorigenesis than that which occurs in sporadic TNBC. However, no difference in this process could be observed in normal adjacent mammary tissue of both TNBC groups: hereditary and sporadic. Deciphering the tumorigenic process can be significantly enhanced by analysis of both tumors and normal tissues.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.tranon.2019.07.016.

**Acknowledgements**

The authors are grateful to all patients for donating their biological samples for scientific research. We thank the A. C. Camargo Institutional Biobank for providing the DNA samples for the study and Dr. Luiz Paulo Kowalski (funding from Fundação de Amparo a Pesquisa do Estado de São Paulo—FAPESP process number 2014/072497) for providing the IPA software. We also thank FAPESP: 2014/509443-1; CNPQ: 465682/2014-6 and CAPES: 88887.136405/2017-00.

**Ethics Approval and Consent to Participate**

All procedures performed involving human participants were in accordance with the ethical standards of the A. C. Camargo Cancer Research Ethics Committee (number 1746/13) and with the 1964 Helsinki declaration and its later amendments. Written informed consent was obtained from all patients of the study, who signed the informed consent allowing the use of their biological material, donated for our Biobank, for scientific projects and for data publication. The A. C. Camargo Cancer Center Biobank has approval of the National Ethical Committee under number B-001.

**Consent for Publication**

Not applicable.

**Availability of Data and Material**

Not applicable.

**Competing Interests**

The authors declare that they have no conflict of interest.

**Funding**

This study was supported by the Brazilian agencies Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, 2013/23277-8), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, 23038.004629/2014-19), and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, 305464/2013-2).

**Authors' Contributions**

D. M. C. and S. J. S. were responsible for providing the funding and together with E. N. F. conceived and designed the study; E. N. F. and R. C. B. selected samples, collected data and biological specimens, and performed experiments; R. V. B. A., R. D. C. D., and J. E. S. performed the bioinformatics analysis. E. N. F., R. C. B., and D. M. C. performed data analysis and interpretation and drafted the first version of the manuscript. All authors have critically reviewed and approved the final version of the manuscript.

**References**


