Frequency of the TP53 R337H variant in sporadic breast cancer and its impact on genomic instability

Carolina Mathias, Stéfanne Bortoletto, Ariana Centa, Heloisa Komechen, Rubens S. Lima, Aline S. Fonseca, Ana Paula Sebastião, Cicero A. Urban, Emerson W. S. Soares, Carolina Prando, Ronald C. Figueiredo, Iglenir J. Cavalli, Luciane R. Cavalli, & Enilze M. F. S. Ribeiro

The R337H is a TP53 germline pathogenic variant that has been associated with several types of cancers, including breast cancer. Our main objective was to determine the frequency of the R337H variant in sporadic breast cancer patients from Paraná state, South Brazil, its association with prognosis and its impact in genomic instability. The genotyping of 805 breast cancer tissues revealed a genotypic and allelic frequency of the R337H variant of 2.36% and 1.18%, respectively. In these R337H+ cases a lower mean age at diagnosis was observed when compared to the R337H− cases. Array-CGH analysis showed that R337H+ patients presented a higher number of copy number alterations (CNAs), compared to the R337H−. These CNAs affected genes and miRNAs that regulate critical cancer signaling pathways; a number of these genes were associated with survival after querying the KMplot database. Furthermore, homozygous (R337H+/R337H+) fibroblasts presented increased levels of copy number variants when compared to heterozygous or R337H− cells. In conclusion, the R337H variant may contribute to 2.36% of the breast cancer cases without family cancer history in Paraná. Among other mechanisms, R337H increases the level of genomic instability, as evidenced by a higher number of CNAs in the R337H+ cases compared to the R337H−.

The R337H is a germline pathogenic variant at codon 337 of exon 10 of the TP53 gene, which occurs outside the DNA binding region in the dimerization region of the p53 protein. The exchange of arginine by a histidine (R337H CGC → CAC) characterizes the variant, which affects the structure (tetramer formation) of the protein, leading to an unstable molecule with the eventual loss of function.

The R337H was first reported in the southern region of Brazil as a germline variant in pediatric patients with adrenocortical tumors (ACTs), where it was identified in 95% of the cases. It was later characterized as a low penetrance TP53 variant. Large population studies evaluating the presence of TP53 R337H in the state of Paraná in southern Brazil revealed its presence in 0.27% and 0.30% of 214,087 tested newborns. These clustered populations and others reported with the TP53 R337H variant can be attributed to a founder effect.

Families carrying the germline TP53 R337H variant present a higher but variable incidence of cancer, while breast, stomach, and brain cancers are the most common in the higher age groups.

A study conducted with tumor samples from the Brazilian states of São Paulo and Rio Grande do Sul reported an 8.6% (70/815) frequency of the TP53 R337H variant among women with breast cancer without family history.
family history. These and other authors have found breast cancer cases and other tumors positive for R337H associated with Li–Fraumeni syndrome (LFS) or Li–Fraumeni-like syndrome, hereditary and non-hereditary breast cancer, and/or asymptomatic high risk patients. Despite the accumulating and compelling evidence of the association between R337H and breast cancer, few studies have investigated the cellular phenotype and/or the genomic “consequences” of the R337H variant in breast cancer cells.

Genomic instability is a hallmark of cancer that occurs in virtually all types of cancer. Genomic instability is essential for tumor progression. The instability can be evident as several DNA and/or chromosome alterations, including gene amplifications, chromosome rearrangements, and changes in DNA copy number.

The main objective of this study was to determine the frequency of the TP53 R337H variant in patients diagnosed with sporadic breast cancer from the state of Paraná and its association with clinical and histopathological parameters to assess its potential prognostic value in this population. To our knowledge, no studies have assessed the direct impact of the R337H variant on the genomic instability of these tumors. We also aimed to determine the patterns of the genome-wide copy number alterations (CNAs) in these cases and their corresponding effects on signaling and functional cellular pathways. In addition, to verify whether R337H+/R337H+ and Wt/R337H+ cells contribute equally to genomic instability, we generated normal cell cultures of fibroblasts from individuals with homozygous and heterozygous TP53 R337H variants exposed to a DNA damage agent and evaluated the patterns of their copy number variations (CNVs).

**Results**

**TP53 R337H variant frequency.** Nineteen patients were identified by real-time PCR as heterozygous carriers of TP53 R337H among 805 women with sporadic breast cancer. All 19 positive patients had their genotypes confirmed by Sanger sequencing. The genotypic and allelic frequencies were 2.36% and 1.18%, respectively.

**TP53 R337H status and clinical-histopathological parameters, and survival.** The breast cancer patients were subdivided into two groups according to the R337H variant status. One group of patients harbored the R337H variant (R337H+; n = 19). The other group were non-carriers (R337H−; n = 50). The non-carriers were selected based on the criteria described in the Materials and Methods. The R337H+ group had a significantly lower mean age at diagnosis compared to the R337H− group (47.88 ± 11.56 and 58.52 ± 15.18 years, respectively; Student’s t-test, t = 2.97, P < 0.05). In the living patients in the R337H+ group, a significantly lower mean age at diagnosis was observed compared to the R337H− group (47.90 ± 9.92 and 57.94 ± 15.9 years, respectively; Student’s t-test, t = 2.22, P < 0.05). No other clinical-histopathological parameters were significantly associated with the R337H mutation status (Table 1). The analysis (multiple logistic regression) considering the R337H status and the clinical variables (age at diagnosis, tumor size, lymph node metastasis, ER, PR and HER2 receptor status), also did not show any significance. Comparison of the survival curves of 14/19 R337H+ patients and 44/50 R337H− patients indicated no significant differences between the two groups of patients (Kaplan–Meier test, P > 0.05).

**Analysis of copy number alterations (CNAs).** To determine the patterns of CNAs that could be influenced by the TP53 R337H variant, we performed genome-wide array-comparative genomic hybridization (CGH) analysis using an oligonucleotide array-CGH platform (Agilent Technologies, Inc.). This analysis was conducted in nine R337H+ of the 19 breast cancer patients and in nine R337H− patients.

In the R337H+ cases, 467 CNAs were observed, with an average of 51.89 ± 33.29 alterations per case. The gains of copy number were more frequent, accounting for 57.6% (269/467) of all alterations. The main cytobands with CNAs (> 30% of cases) observed in this group were 8q11.1–q24.3 (89% of cases), 8p12–p11.21 (78%), and 1q21.1–q44 (67%), followed by 1p36.33–p36.32, 6p25.3–p21.1, 8q24.3, 11p15.5, 11q13.4–q25, 13q11–q34, 14q21.1–q32.33, 14q32.33, 17p13.3–p11.2, 19p13.3, 19p13.11, and 20q11.21 (44%). Among these gains of copy number were more frequent, accounting for 82.75% (128/204) of all alterations. The main cytobands with CNAs (> 30% of the cases) observed in each group of patients are presented in Table 2.

In the R337H− group, 204 CNAs were observed, with an average of 22.67 ± 16.78 alterations per case. The gains of copy number were more frequent, accounting for 82.75% (128/204) of all alterations. The main cytobands with CNAs observed in this group were 1q43.32.33 (89% of cases), 8p11.22, 8q11.1–q24.3 and 22q11.22 (78%), 7p22.3 (67%), 1p36.33–p36.32 (56%), and 8q24.3, 10q26.3, 16p13.3–p11.1, and 17p13.3–p11.2. In these cytobands (except 17p, which was observed with loss of copy number), high levels of gains (log 2 > 2.0) were observed in 8p11.22 and 22q11.22.

Finally, the comparison of the total number of CNAs in both groups, which was measured by the comparison of the total “number of calls” in the Cytogenomics aberration interval base reports, revealed a significantly higher number of CNAs in the R337H+ group of patients (t = 2.35; P < 0.05). The findings indicated that genomic instability was more frequent in patients with the R337H variant. The main cytobands with CNAs (> 30% of the cases) observed in each group of patients are presented in Table 2.

**Functional enrichment pathways.** To determine the function of the microRNAs (miRNAs) mapped in these cytobands (Table S1) that could be affected by the presence of CNAs, pathway enrichment analysis was performed using DIANA-miRPath v.3.0P. MiRNAs corresponding to target genes and the main signaling pathways involved were identified. Due to the large number of cytobands affected, especially in the R337H+ group of patients, only cytobands affected more than 50% of the cases were considered.
In the R337H+ group of patients, 76 significant Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were observed (P < 0.05). Among the top 15 pathways observed (based on P-value), those affected by the largest number of miRNAs were the “Pathways in cancer” and “Proteoglycans in cancer,” with 82 miRNAs and 79 miRNAs, respectively. In this group, the p53 signaling pathway was among the significant pathways involved, being affected by 55 miRNAs (Table 3).

In the R337H− group of patients, 26 significant pathways were observed (P < 0.05). Among the top 15 pathways observed (based on P-value), the ones affected by the largest number of miRNAs (10 miRNAs) were “Proteoglycans in Cancer, Hippo, Ras, Pluripotency stem cells regulating and Thyroid hormone signaling pathways, Pathways in cancer, Axon guidance and Focal adhesion” (Table 3).

Finally, we identified the miRNA targets predicted to be regulated by these miRNAs, using Tarbase v.7.0, miRNA target gene (miTG) scores > 0.7 indicated microT-CDS interactions. Only genes correlated with miRNAs that presented strong evidence in the validation methods cited in the Materials and Methods section were considered. This analysis revealed 256 and 180 miRNA targets for the R337H+ and R337H− groups, respectively. Upon integration with the genes that were identified in the affected cytobands by array-CGH (3079 in the R337H+ group and 365 in the R337H− group), we observed that there were 43 genes in common in the R337H+ group and seven genes in the R337H− group (Fig. 1, Table S2). Four genes (CCNE2, MTDH, RDH10, and SNAI2) were commonly observed in both the R337H+ and R337H− groups of patients. The genes affected in the R337H+ group were in the chromosome regions. These genes were mainly affected by CNAs in these cases as revealed by the array-CGH analysis, including 1q21–q44, 2q11-37.3, 8q21–q24, 16q23.2, and 17q25.3. Interestingly, in the R337H− group, all the genes identified in the integration analysis mapped at the 8q region, and were commonly affected in these cases in the array-CGH analysis. These results indicated that CNAs could affect genes that are also potentially regulated by miRNAs.

Table 1. Association of the clinical and pathological variables and survival with TP53 R337H.

| Variable                  | TP53 R337H+ | TP53 R337H− | P value       |
|---------------------------|-------------|-------------|---------------|
| Age (yrs)                 | 47.88 ± 11.56 (n = 19) | 58.52 ± 15.18 (n = 50) | t = 2.97; P < 0.05 |
| Tumor size (cm)           | 2.76 ± 1.43 (n = 19) | 3.02 ± 1.62 (n = 50) | t = 0.62; P > 0.50 |
| T1                        | 26.31% (n = 5) | 30% (n = 15) |               |
| T2                        | 63.15% (n = 12) | 64% (n = 32) |               |
| T3                        | 10.5% (n = 2)  | 6% (n = 3) |               |
| Tumor grade               |             |             |               |
| I                         | 5.26% (n = 1)  | 18% (n = 9) |               |
| II                        | 42.10% (n = 8) | 50% (n = 25) |               |
| III                       | 42.10% (n = 8) | 32% (n = 16) |               |
| Tumor stage               |             |             |               |
| I                         | 15.8% (n = 3)  | 15.9% (n = 7) |               |
| II-A                      | 15.8% (n = 3)  | 29.5% (n = 13) |               |
| II-B                      | 26.31% (n = 5) | 36.36% (n = 16) |               |
| II-A                      | 10.52% (n = 2) | 9.1% (n = 4) |               |
| Lymph node metastasis     |             |             |               |
| Positive                  | 42.1% (n = 8)  | 56% (n = 28) | χ² = 0.03; P > 0.10 |
| Negative                  | 36.8% (n = 7)  | 44% (n = 22) |               |
| ER expression             |             |             |               |
| Positive                  | 52.63% (n = 10) | 80% (n = 40) | χ² = 1.15; P > 0.20 |
| Negative                  | 26.3% (n = 5)  | 20% (n = 10) |               |
| PR expression             |             |             |               |
| Positive                  | 47.36% (n = 9) | 84% (n = 42) | χ² = 2.62; P > 0.10 |
| Negative                  | 26.31% (n = 5) | 16% (n = 8) |               |
| HER2 over expression      |             |             |               |
| Positive                  | 21.05% (n = 4) | 28% (n = 14) | χ² = 0.03; P > 0.80 |
| Negative                  | 47.36% (n = 9) | 72% (n = 36) |               |
| Patients alive            |             |             |               |
| Age at diagnosis          | 47.90 ± 9.92 | 57.94 ± 15.82 |               |
| Survival (mos)            | 59.25 ± 46.14 | 69.74 ± 42.47 |               |
| Patients deceased         |             |             |               |
| Age at diagnosis          | 55 ± 9.89 | 56.78 ± 11.21 |               |
| Survival (mos)            | 16 ± 11.31 | 37.22 ± 23.39 |               |

In the R337H+ group of patients, 76 significant Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were observed (P < 0.05). Among the top 15 pathways observed (based on P-value), those affected by the largest number of miRNAs were the “Pathways in cancer” and “Proteoglycans in cancer,” with 82 miRNAs and 79 miRNAs, respectively. In this group, the p53 signaling pathway was among the significant pathways involved, being affected by 55 miRNAs (Table 3).

In the R337H− group of patients, 26 significant pathways were observed (P < 0.05). Among the top 15 pathways observed (based on P-value), the ones affected by the largest number of miRNAs (10 miRNAs) were “Proteoglycans in Cancer, Hippo, Ras, Pluripotency stem cells regulating and Thyroid hormone signaling pathways, Pathways in cancer, Axon guidance and Focal adhesion” (Table 3).

Finally, we identified the miRNA targets predicted to be regulated by these miRNAs, using Tarbase v.7.0, miRNA target gene (miTG) scores > 0.7 indicated microT-CDS interactions. Only genes correlated with miRNAs that presented strong evidence in the validation methods cited in the Materials and Methods section were considered. This analysis revealed 256 and 180 miRNA targets for the R337H+ and R337H− groups, respectively. Upon integration with the genes that were identified in the affected cytobands by array-CGH (3079 in the R337H+ group and 365 in the R337H− group), we observed that there were 43 genes in common in the R337H+ group and seven genes in the R337H− group (Fig. 1, Table S2). Four genes (CCNE2, MTDH, RDH10, and SNAI2) were commonly observed in both the R337H+ and R337H− groups of patients. The genes affected in the R337H+ group were in the chromosome regions. These genes were mainly affected by CNAs in these cases as revealed by the array-CGH analysis, including 1q21–q44, 2q11-37.3, 8q21–q24, 16q23.2, and 17q25.3. Interestingly, in the R337H− group, all the genes identified in the integration analysis mapped at the 8q region, and were commonly affected in these cases in the array-CGH analysis. These results indicated that CNAs could affect genes that are also potentially regulated by miRNAs.

Association of the target miRNA genes with survival using KMPlot database. The genes identified in both groups of patients in the aforementioned integrative analysis were queried in the KM Plot database.
to determine their potential association with survival outcome in breast cancer patients. This analysis was performed by querying the database in all the groups of breast cancer cases available and in breast cancer cases based on the TP53 variants status. The type of TP53 variants in the cases was not disclosed (Table S2).

In the R337H+ group, 72.1% (31/43) of the genes were associated with survival. Seventeen and 14 of the genes were identified in patients with higher and lower survival rates, respectively. Five genes in this group were only observed in cases that presented with TP53 variants in the KMplot database. Overexpression of three of these genes (ECM1, MMP16, and CTHRC) was associated with significantly worse survival. Overexpression of MCL1 and STAT1 was associated with better survival. Four other genes (ITGA6, HOXD10, FASN, and BUP1) were observed only in cases that were negative for TP53 variants in the KMplot database. Higher expression of ITGA6 was significantly associated with better survival. Higher expression of HOXD10, FASN, and BUP1 was associated with worse survival.

In the R337H− group of patients, 85.7% (6/7) genes were associated with survival. Only SNAI2 (also found in the R337H+ group) was not associated with survival. Overexpression of OXR1 was the only gene in this group that was significantly associated with TP53 status, conferring a higher survival rate in patients with no TP53 variants. Interestingly, in the R337H+ group, three genes (IGFBP5, MAF, and SMYD3) that were not associated with survival in the breast cancer cases in general were associated with survival specifically in cases with TP53 variants.

Table 2. Main affected cytobands (> 30% of the cases) and their corresponding patterns of CNAs in both groups of patients analyzed.

| Cytoband | R337H+ |  | R337H− |  |
|----------|--------|---|--------|---|
| 1p36.13  | -      | 0 | Gain   | 3 (33) |
| 1p36.33–p36.32 | High gain | 4 (44) | Gain | 5 (56) |
| 1q21.2–q44 | Gain | 6 (67) | - | 0 |
| 2q11.1–q37.3 | Gain | 3 (33) | - | 0 |
| 4p16.1  | Gain | 3 (33) | Gain | 3 (33) |
| 4q22.2–q28.1 | Gain | 3 (33) | - | 0 |
| 6p22.1–p11.2 | Loss | 3 (33) | - | 0 |
| 6p25.3–p21.1 | Gain | 4 (44) | - | 0 |
| 6q11.1–q27 | Loss | 3 (33) | - | 0 |
| 7p15.3–p12.3 | Gain | 3 (33) | - | 0 |
| 7q12.2–q13.3 | High gain | 7 (78) | - | 0 |
| 8q11.1–q24.3 | Gain | 8 (89) | Gain | 7 (78) |
| 8q24.3 | Gain | 4 (44) | Gain | 4 (44) |
| 9p24.3–p13.1 | Loss | 3 (33) | - | 0 |
| 9q34.11 | High gain | 3 (33) | - | 0 |
| 9q34.2–q34.3 | Gain | 3 (33) | Gain | 3 (33) |
| 10q11.23–q26.3 | Loss | 3 (33) | - | 0 |
| 10q26.3 | Gain | 3 (33) | Gain | 4 (44) |
| 11p15.5 | Gain | 4 (44) | Gain | 3 (33) |
| 11q12.2–q14.1 | Gain | 3 (33) | - | 0 |
| 11q13.3–q13.4 | - | 0 | Gain | 3 (33) |
| 11q13.4–q25 | Loss | 4 (44) | - | 0 |
| 12q12–q24.3 | Gain | 3 (33) | Gain | 3 (33) |
| 13q11–q34 | Loss | 4 (44) | - | 0 |
| 13q34 | High gain | 3 (33) | Gain | 3 (33) |
| 14q21.1–q32.3 | Loss | 4 (44) | - | 0 |
| 14q24.2–q24.3 | Gain | 3 (33) | - | 0 |
| 16p12.3 | Gain | 3 (33) | Gain | 3 (33) |
| 16q11.2–q24.3 | Loss | 3 (33) | - | 0 |
| 17p11.3–p11.2 | Loss | 4 (44) | Loss | 4 (44) |
| 17q11.1–q25.3 | Gain | 3 (33) | - | 0 |
| 17q25.3 | Gain | 3 (33) | Gain | 3 (33) |
| 17p13.3–p13.3 | High gain | 3 (33) | Gain | 3 (33) |
| 18q11.1–q12.3 | Gain | 3 (33) | - | 0 |
| 19p13.11 | High gain | 4 (44) | - | 0 |
| 19p13.3 | Loss | 4 (44) | - | 0 |
| 19p13.3 | Gain | 3 (33) | - | 0 |
| 21q22.3 | High gain | 3 (33) | Gain | 3 (33) |
| 22q11.22 | High gain | 3 (33) | High gain | 7 (78) |
| 22q13.31 | High gain | 3 (33) | - | 0 |
| 22q13.33 | Gain | 3 (33) | - | 0 |
| 22q13.3 | High gain | 3 (33) | Gain | 3 (33) |
Table 3. Top 15 pathways, their corresponding number of genes and miRNAs, observed to be affected by the main CNAs, in the two groups of breast cancer patients analyzed (presented by the number of affected miRNAs). KEGG, Kyoto Encyclopedia of Genes and Genomes; ECM, Extracellular Matrix.

| KEGG pathway                                                                 | #MiRNAs | #Genes | P value     |
|------------------------------------------------------------------------------|---------|--------|-------------|
| Pathways in cancer                                                           | 82      | 318    | 8.80E−04    |
| Proteoglycans in cancer                                                      | 79      | 171    | 1.24E−02    |
| Hippo signaling pathway                                                      | 77      | 125    | 1.57E−06    |
| Focal adhesion                                                               | 76      | 175    | 5.28E−05    |
| Signaling pathways regulating pluripotency of stem cells                    | 76      | 122    | 6.29E−03    |
| Oxytocin signaling pathway                                                   | 72      | 132    | 6.95E−06    |
| Axon guidance                                                                | 72      | 111    | 6.29E−03    |
| TGF-beta signaling pathway                                                   | 67      | 70     | 5.28E−05    |
| ErbB signaling pathway                                                       | 66      | 79     | 2.55E−05    |
| Renal cell carcinoma                                                         | 65      | 62     | 2.47E−04    |
| Pancreatic cancer                                                            | 62      | 59     | 4.95E−06    |
| Glioma                                                                       | 62      | 57     | 1.57E−06    |
| ECM-receptor interaction                                                     | 62      | 70     | 6.29E−03    |
| Colorectal cancer                                                            | 60      | 57     | 4.51E−06    |
| Long-term depression                                                         | 60      | 51     | 2.08E−03    |
| p53 signaling pathway                                                        | 55      | 60     | 2.56E−02    |

**R337H+ breast cancer patient**

| KEGG pathway                                                                 | #MiRNAs | #Genes | P value     |
|------------------------------------------------------------------------------|---------|--------|-------------|
| Proteoglycans in cancer                                                      | 10      | 65     | 1.74E−03    |
| Hippo signaling pathway                                                      | 10      | 44     | 4.30E−03    |
| Ras signaling pathway                                                        | 10      | 69     | 6.19E−03    |
| Signaling pathways regulating pluripotency of stem cells                    | 10      | 40     | 5.23E−02    |
| Pathways in cancer                                                           | 10      | 97     | 8.69E−02    |
| Axon guidance                                                                | 10      | 38     | 1.84E−01    |
| Focal adhesion                                                               | 10      | 59     | 1.84E−01    |
| Thyroid hormone signaling pathway                                            | 10      | 31     | 1.84E−01    |
| ErbB signaling pathway                                                       | 9       | 34     | 6.29E−06    |
| Regulation of actin cytoskeleton                                             | 9       | 64     | 5.23E−02    |
| Amphetamine addiction                                                        | 9       | 21     | 4.30E−03    |
| Glioma                                                                       | 8       | 26     | 5.13E−06    |
| Lysine degradation                                                           | 8       | 16     | 6.29E−06    |
| Renal cell carcinoma                                                         | 7       | 24     | 7.33E−02    |
| Synaptic vesicle cycle                                                       | 6       | 18     | 2.27E−01    |

**R337H− breast cancer patient**

Figure 1. Venn diagram showing the integration of genes located at the cytobands most affected by CNAs and the target genes regulated by the miRNAs mapped at the cytobands in the TP53 R337H+ (left) and R337H− (right) group of patients.
copy number variations (cnV) analysis. To further verify the differences in the level of genomic instability in R337H+ cells, we performed genome-wide CNVs analyses in homozygous (R337H+/R337H+) and heterozygous (Wt/R337H+) fibroblast cell cultures. Specific targets (9q, 9q33–34, 11p, and 11p15) considered to harbor genes potentially involved in adrenocortical (ACT)29,30 and breast cancer31,32 tumorigenesis were evaluated (Table 4). Statistically significantly increased levels of CNVs (all $P < 0.05$) in chromosomes 1, 2, 3, 6, 7, 8, 10, 11, 13, 14, 15, 16, and 22 were found only in homozygous R337H/R337H cells when compared to the other cells. Remarkably, the most consistent losses and gains were identified in 11p, suggesting that this is a susceptible target to CNV, which is also known to be related to IGF-2 overexpression in ACTs.

Discussion

In this study, we report the frequency of the TP53 R337H variant in tumor tissues of patients with no family history of breast cancer, and its association with clinical and histopathologic parameters and survival outcome. We also conducted a comprehensive computational analysis to determine the impact of the TP53 R337H variant on genomic instability, evaluating both the breast tumor tissue of the patients and homozygous (R337H+/R337H+) and heterozygous (Wt/R337H+) fibroblast cell cultures established from a patient with adrenocortical cancer and a variant carrier, respectively.

Table 4. Comparisons* of the CNV means between homozygous (R337H+/R337H+) and heterozygous (Wt/R337H+) human fibroblasts. Only statistically significant differences ($p < 0.05$) are shown. *The comparisons were performed between homozygous ($n = 8$, 4 treated and 4 untreated assays) and heterozygous ($n = 8$, 4 treated and 4 untreated assays). Chr, chromosomes; G, gain; L, loss.

| Chr | CNV > 10 kb | CNV > 50 kb | CNV > 100 kb |
|-----|------------|-------------|--------------|
|     | Hetero | Homo | $P$ value | Hetero | Homo | $P$ value | Hetero | Homo | $P$ value |
| chr1 | G + L | 4.7 | 20.4 | 0.022 | chr1 | L | 0.5 | 1.5 | 0.038 | chr10 | L | 0 | 0.7 | 0.046 |
| chr1 | G | 10.1 | 42.0 | 0.033 | chr1 | G + L | 2.5 | 6.7 | 0.033 | chr14 | G + L | 0.8 | 1.8 | 0.028 |
| chr2 | G | 4.7 | 14.7 | 0.027 | chr2 | G + L | 2.7 | 5.8 | 0.049 | chr15 | L | 0 | 0.7 | 0.046 |
| chr3 | G | 4.2 | 16.8 | 0.043 | chr3 | L | 0.8 | 3.4 | 0.006 | chr15 | G + L | 1.7 | 3.8 | 0.017 |
| chr6 | G | 2.3 | 7.0 | 0.040 | chr6 | G + L | 1.8 | 6.1 | 0.037 | chr15 | G | 0.5 | 2.3 | 0.016 |
| chr7 | G | 1.2 | 25.3 | 0.032 | chr7 | G | 0.3 | 2.0 | 0.028 | chr16 | G + L | 0.5 | 2.4 | 0.016 |
| chr8 | G | 3.0 | 15.7 | 0.041 | chr8 | G | 0.3 | 4.7 | 0.028 |
| chr10 | G | 6.3 | 32.0 | 0.030 | chr10 | G + L | 0.5 | 6.1 | 0.004 |
| chr10 | G + L | 8.0 | 37.8 | 0.022 | chr10 | G + L | 1.8 | 6.4 | 0.038 |
| chr11 | G | 2.7 | 13.3 | 0.042 | chr11 | G + L | 1.8 | 5.7 | 0.031 |
| chr11 | G + L | 4.8 | 24.1 | 0.034 | chr11 | G | 2.1 | 4.7 | 0.023 |
| chr13 | G | 0 | 2.6 | 0.014 | chr13 | L | 0.7 | 3.1 | 0.043 |
| chr14 | G | 2.5 | 6.1 | 0.029 | chr14 | G + L | 2.8 | 7.8 | 0.008 |
| chr15 | G | 3.5 | 8.4 | 0.041 | chr15 | G + L | 1.0 | 4.8 | 0.024 |
| chr16 | G | 2.6 | 18.8 | 0.040 | chr16 | G + L | 1.3 | 6.4 | 0.048 |
| chr22 | G + L | 2.3 | 19.1 | 0.029 | chr22 | L | 0 | 1.7 | 0.045 |

| Specific targets (11 and 11p15) | | | |
| chr11p | G | 1.3 | 6.3 | 0.039 | chr11p | G + L | 0.3 | 2.4 | 0.048 | chr11p | L | 0 | 0.7 | 0.046 |
| chr11p | L | 0.06 | 4.7 | 0.032 | chr11p | G + L | 0 | 1.3 | 0.048 |
| chr11p | G + L | 2.0 | 11.0 | 0.014 |
| chr11p | G | 0.5 | 5.0 | 0.023 |
ranging from 1.57% in Curitiba, the capital of Paraná, to 7.6% in the eastern Paraná state, where agricultural activities are more prominent. It is important to point out, however, that in this study only the TP53 R337H variant was genotyped. Therefore, it is very likely that both the R337H carriers and non-carriers groups, harbor other somatic and/or TP53 germline variants.

The association of clinical parameters in the groups of breast cancer patients we studied based on the R337H variant revealed a lower average age (47.88 ± 11.56) at diagnosis of the R337H+ group compared to the R337H− group (47.88 ± 11.56 vs. 58.52 ± 15.18 years; P < 0.05). Younger age at diagnosis was also observed by Giacomazzi et al.12 and Andrade et al.13 in patients with breast cancer positive for the R337H variant. A previous study reported that 12.1% of the patients diagnosed with breast cancer were diagnosed before 45 years of age compared to 5.1% diagnosed after 55 years of age. The same was observed in the study by Gomes et al.14, in which two cases of breast cancer with the R337H variant had a lower age at diagnosis. In these cases, however, both patients had a family history of breast and/or other types of cancer. Altogether, the present and prior studies provide evidence of an association between the R337H variant and early age of diagnosis, regardless of a family history of cancer.

To determine the impact of the R337H variant on the tumor genome stability of breast cancer patients, we performed a genome-wide evaluation of CNAs. Genomic instability is a hallmark of cancer hallmark that can be evident as the presence of chromosome regions with copy number gains/amplifications and losses/deletions20,21. These alterations can directly affect the expression of genes and miRNAs mapped at these chromosome regions33,34. In particular, miRNAs have been shown to be commonly affected targets for genomic instability35,36, which can significantly modulate tumor progression, through the regulation of critical cancer genes, such as the TP5337.

In our analysis, we observed a significantly higher frequency of CNAs in the R337H+ breast cancer patients than in the R337H− group. In addition, a significantly higher number of cases in the R337H+ group displayed the main affected cytobands compared to the R337H− group. These results showed a higher level of genomic instability in the R337H+ patients and a preferential involvement of the most affected cytobands. Not surprisingly, as demonstrated by KEGG pathway enrichment analysis based on the mapping of genes and miRNAs in these affected cytobands, we observed functional signaling pathways that were potentially affected in both groups of patients. However, a larger number of these pathways were observed in the R337H+ group. Critical cancer-associated pathways among the top 15 most significantly affected, such as Proteoglycans in cancer, Pathways in cancer, ErbB, Hippo, and Ras pathways, were observed in both groups, likely not reflecting the presence of the R337H variant but the tumorigenic process itself.

Interestingly, in the R337H+ group, the TP53 and TGFβ signaling pathways, which were not observed in the R337H− group, were among the pathways mostly affected by the miRNAs present in the cytobands with CNAs. Crosstalk between these pathways, via the Smad signal transduction pathway38, has been reported, and although the mechanisms involved remain to be fully elucidated, additional signaling pathways, such as phosphoinositide 3-kinase/AKT and extracellular signal-regulated kinase39,40 and the involvement of miRNA regulation41 have been suggested.

As cited above, considering that CNAs are one of the mechanisms that affect miRNA expression (and thus miRNA target expression)33,34, we next identified genes that were potentially altered by CNAs and that were also targets of the miRNAs mapped in these main affected regions. In the R337H+ and R337H− groups, we observed 43 and seven genes (four genes were common to both groups), respectively, which could be affected by these two distinct molecular mechanisms. The genes affected in the R337H+ group were located in the cytobands, which were mainly affected by CNAs in these cases, such as 1q21.2, 1q44, 2q13, 2q31.1, 2q32.2, 2q35, 8q21.3, 8q22.3, 8q23.1, 16q23.2, and 17q25.3 (in the R337H− group, only genes mapped at 8q were observed in this integration analysis). Although the impact of such alterations in gene and miRNA expression has to be confirmed in experimental expression assays, the observations support the finding that CNAs can affect genes that are also potentially regulated by miRNAs33,34,41-44. Several of these genes were previously identified as members of the main signaling pathways observed and, interestingly, displayed direct protein interactions with p53 (data not shown—String Network v. 11.0 [https://string-db.org/]). These genes included BUB1, CCNE2, MCL1, MYC, SNAI2, and STAT1. Several miRNAs mapped at the above cytobands have been previously reported as regulators of TP53 gene expression in tissue samples of patients with sporadic breast cancer45. However, little is known about the roles of miRNAs in patients with TP53 variants. A limited number of reports in cancer have shown the phenotypic consequences of variant's TP53 upon miRNA binding46-49, such as gain of function of the R157H and miR-12850, and R273H and miR-27a48. In our study, in the R337H+ group the miR-128 was among the miRNAs previously described in these studies46. However, no miRNA present in the R337H− group was previously associated with TP53 variants. One recent study has shown the association of a polymorphism in miR-605 with the occurrence of multiple primary tumours in R337H carriers that meet the LF criteria45. However, this miRNA, which mapped at the 10q21.1 cytoband, was not among the main regions affected by CNAs in any of the groups of the patients in this study.

The query of the KMplot database of the genes we identified after the integration analysis above could indicate their association with survival of breast cancer patients. The analysis in the R337H+ group revealed significant associations with 72.1% of the genes, five of which (CYP17A1, ECM1, MCL1, MMP16, and STAT1) were associated specifically in cases that presented with TP53 variants in the KMplot database. Four other genes (ITGA6, HOXD10, FASN, and BUP1) were observed only in the breast cancer cases in the database that were negative for TP53 variants. In addition, three genes (IGFBP5, MAF, and SMYD3) that did not associate with survival in breast cancer cases in general were associated with survival specifically in cases with TP53 variants. In the R337H− group of patients, only OX1R was significantly associated with TP53 status. Interestingly, this gene (Human Oxidation Resistance 1), originally identified as a protein that decreases genomic mutations in Escherichia coli51, prevented reactive oxidation species formation and reduced the duration of gamma-ray-induced G2/M
results indicate that the presence of the R337H variant is associated with an increased level of genomic instability TP53 allele. Altogether, these which showed a significant increase in CNVs compared to cells with one wild-type
regulate critical cancer signaling pathways. This instability was also observed in R337H+/R337H+ fibroblast cells, history in Paraná state of Brazil. Among other mechanisms, R337H increases the level of breast cancer genomic from the HNSG Pathology Service.
Biorepository, respectively, and 282 were from paraffin-embedded formalin-fixed (FFPE) tissue blocks acquired
Cytogenetics and Oncogenetics Laboratory Biorepository (collected at HNSG and HC), and the UOPECCAN
tal Nossa Senhora Das Graças (HNSG) and Hospital de Clínicas (HC), both from Curitiba and União Oeste
pean descent, who had been diagnosed with breast cancer were collected during primary surgery at Hospi-
as R337H53–56. Letouzé et al.56 analyzed 25 ACT tumors, 13 of which with the R337H variant. The authors uti-
TP53
Previous studies have shown the additional impact of CNVs on tumors harboring germline TP53 variants, such as R337H55–56. Letouzé et al.56 analyzed 25 ACT tumors, 13 of which with the R337H variant. The authors utilized high-resolution single nucleotide polymorphism analysis to demonstrate that the cases with the wild-type
TP53
displayed distinct genomic profiles, with significantly fewer rearrangements, compared to the cases with the R337H variant. This finding was also observed in patients with Li-Fraumeni, where an increased number of CNVs were observed in patients carrying germline variants in the TP53 gene, such as R337H55.
In conclusion, the TP53 R337H variant may contribute 2.36% of all breast cancer cases without family cancer history in Paraná state of Brazil. Among other mechanisms, R337H increases the level of breast cancer genomic instability, as evidenced by the presence of a higher number of CNAs potentially affecting genes/miRNAs that regulate critical cancer signaling pathways. This instability was also observed in R337H+/R337H+ fibroblast cells, which showed a significant increase in CNVs compared to cells with one wild-type TP53 allele. Altogether, these
results indicate that the presence of the R337H variant is associated with an increased level of genomic instability in the cells. However, its direct role in modulating breast cancer tumorigenicity is unknown.

Materials and methods
Sample collection. A total of 805 breast tissue samples from different patients, predominantly of Euro-
pean descent, who had been diagnosed with breast cancer were collected during primary surgery at Hospi-
tal Nossa Senhora Das Graças (HNSG) and Hospital de Clinicas (HC), both from Curitiba and Uniao Oeste Paranaense de Estudos e Combate ao Cancer (UOPECCAN), Cascavel, southern Brazil. All patients provided signed informed consent. Among these samples, 418 and 105 were of fresh tissue acquired from the Human Cyto

TP53 R337H variant genotyping. Genotyping for the TP53 R337H variant (NM_000546.6(TP53):c.10

TP53 R337H variant status and clinical-histopathological parameters, and survival. The association of the TP53 R337H variant and the clinical-histopathological parameters was performed between the R337H+ and R337H− groups of patients, considering tumor grade and size, lymph node metastasis, and estrogen, progester-
one, and HER2 receptor status. From the R337H− group, 50 patients were selected for this analysis following two main criteria. The first a diagnosis of invasive ductal carcinoma (the same diagnosis as the carrier group). The second was the highest amount of clinical information for the clinical and histopathological parameters above. Student’s t-test was performed to compare the patient groups’ age and tumor size. The chi-square test was used to compare tumor grade and stage, expression of estrogen, progesterone, and HER2 receptors, and lymph node metastasis. Multiple logistic regression analysis was performed using the software GraphPad Prim 8 and taking into consideration the clinical parameters (age at diagnosis, tumor size, lymph node metastasis, and ER, PR, and HER2 receptors as independent variables (X) and the patient R337H genotype (positive or negative for the R337H variant) as the single dependent (Y) variable. Survival data were analyzed using Student’s t and Kaplan Meier tests. Statistical significance was considered at P < 0.05.

**Genome-wide CNA analysis.** To detect CNAs, as a measurement of genomic instability, the DNA from the breast cancer cases positive and negative for the TP53 R337H+ variant were profiled using the SurePrint G3 Human CGH Microarray (Agilent, Santa Clara, CA, USA) according to our previous protocol for FFPE samples57. Nine patients were evaluated from each group of patients using the same protocol. DNA isolated from peripheral blood from multiple normal individuals was used as a control (reference) DNA. Control and case samples were directly labeled using the Bioprimer a-CGH Genomic Labeling kit and hybridized to the arrays for 40 h. The arrays were scanned using the model G2565CA scanner (Agilent). The data were extracted using Feature Extraction software v10.10 (Agilent). The Agilent Cytogenomics v.5.0 software was used to analyze the data using the algorithm ADM-2, threshold of 6.0, and an aberration filter with a minimum of three probes. Copy number gains and losses were defined as the minimum average absolute log2 ratio (intensity of the Cy5 dye (reference DNA)/intensity of the Cy3 dye (test DNA)) value of ≥0.25 and < −0.25, respectively. High copy number gains and losses were considered for log2 ratios ≥2.0 or < 2.0, respectively. The number of “calls” (total significant number of CNAs) and the specifically affected cytobands were obtained from the generated aberration interval base reports (Agilent Cytogenomics v.5.0). Only cytobands affected in >30% of the cases were considered. Statistical analysis of the cytobands and number of calls was performed using the GraphPad Prism software v. 6.0.

**Functional enriched pathways.** For both groups of breast cancer patients (TP53 R337H+ and R337H−) analyzed by array-CGH, the identification of the genes and miRNAs mapped in the cytobands that were mostly affected by CNAs was obtained from the Agilent Cytogenomics v.5.0 interval base reports (based on the analysis parameters described above). DIANA-miPath v.3.027 was used to perform pathway enrichment analysis, based on the KEGG database (https://www.genome.jp/kegg). Only miRNA/mRNA targets that presented a miRNA Target Gene (miTG) score > 0.7 based on the microT-CDS interactions were included. For the selection of the main targets, only those that presented strong evidence in validation methods (luciferase assays, western blotting, and qPCR) were considered, according to miRTarBase v.7.028. A direct integration of the miRNA target genes mapped in the most affected cytobands was performed, as previously described42,44 to determine whether the genes also mapped in these regions were miRNA targets, and therefore could be potentially affected by both CNAs and miRNA expression regulation.

**Kaplan–Meier plot analysis.** The KM Plotter Tool (https://kmplot.com/analysis/) was used to calculate hazard ratios, confidence intervals, and log-rank P values for the selected genes resulting from the integration of the genes that were miRNA targets and also affected by copy number alterations (CNAs). This analysis was performed in relation to survival in the aggregated breast cancer clinical studies extracted from The Cancer Genome Atlas and Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) databases (breast cancer cases in general and selected for TP53 variants).

**DNA damage induction in normal fibroblasts homozygous and heterozygous for the TP53 R337H variant.** Skin biopsies from two TP53 R337H+/R337H+ homozygous boys who had adrenocortical cancer and their heterozygous Wt/R337H+ mothers. The skin biopsies collected were from surgical removal of the foreskin for therapeutic and prophylactic reasons with the authorization by their parents. Fibroblast cultures were treated with doxorubicin twice, at passage 6 for 5 days at a concentration of 0.025 μM, and after recovering for 15 days in normal medium at a concentration of 0.025 μM for 5 days. The cells were allowed to recover for 8 days, and DNA was isolated together with controls (untreated cells-passage 5). Each cell type, including untreated controls and doxorubincin treated (R337H+/R337H+, Wt/R337H+) were prepared in duplicate, considering that the variation of the results in most of these assays were not significant. In few cases, where differences in the cell counting were observed, the entire assay was repeated.

**Genome-wide copy number variation (CNV) analysis.** Genomic DNA was isolated from all the samples (n = 4 for each cell type, with two doxorubincin treated and two untreated samples) for CNV analysis, using the Affymetrix 6.0 array. The data were analyzed using the Affymetrix Genotyping Console software in the Affymetrix Power Tool (https://www.affymetrix.com/partners_programs/programs/developer/tools/powertools.affx), using the following criteria: < 90% genotype call rate or minor allele frequency < 5% or Hardy–Weinberg equilibrium exact P value < 0.05 in cases or controls. The CNVs were estimated using two software programs: APT and PENN CNV. CNV analyses were performed for each of the chromosomes, or for specific cytobands (9q, 9q33–34, 11p, 11p15, 17p and 17p13) considering their relevance to the ACTs29,30 and breast cancer31,32, as we previously described. CNVs identified in cases with > 10% overlap with CNVs identified in the controls were
not considered. The CNVs identified were checked in the Database of Genomic Variants (https://projects.tcag.ca/variation). The comparison of the CNV mean values among the treated and non-treated cases was performed using the two-tailed Student’s t-test. A P-value < 0.05 was considered significant.

Received: 19 March 2020; Accepted: 17 August 2020
Published online: 06 October 2020

References
1. DiGiammarino, E. L. et al. A novel mechanism of tumorigenesis involving pH-dependent destabilization of a mutant p53 tetramer. Nat. Struct. Biol. 9(1), 12–26 (2002).
2. Ribeiro, R. C. et al. An inherited p53 mutation that contributes in a tissue-specific manner to pediatric adrenal cortical carcinoma. Proc. Natl. Acad. Sci. U. S. A. 98, 9330–9335 (2001).
3. Figueiredo, B. C. et al. Penetration of adrenocortical tumours associated with the germline TP53 R337H mutation. J. Med. Genet. 43, 91–96 (2006).
4. Costa, T. E. J. et al. Penetration of the TP53 R337H mutation and pediatric adrenocortical carcinoma incidence associated with environmental influences in a 12-year observational cohort in Southern Brazil. Cancers (Basel). 11(1), 1804–1820 (2019).
5. Custódio, G. et al. Impact of neonatal screening and surveillance for the TP53 R337H mutation on early detection of childhood adrenocortical tumors. J. Clin. Oncol. 31, 2619–2626 (2013).
6. Pinto, E. M. et al. Founder effect for the highly prevalent R337H mutation of tumor suppressor p53 in Brazilian patients with adrenocortical tumors. Arq. Bras. Endocrinol. Metabol. 48, 647–650 (2004).
7. Custódio, G. et al. Increased incidence of choroid plexus carcinoma due to the germline TP53 R337H mutation in southern Brazil. PLoS ONE 6(3), e18015 (2011).
8. Garritano, S. et al. Determining the effectiveness of High Resolution Melting analysis for SNP genotyping and mutation scanning at the TP53 locus. BMC Genet. 10, 1–12 (2009).
9. Seidinger, A. L. et al. Association of the highly prevalent TP53 R337H mutation with pediatric choroid plexus carcinoma and osteosarcoma in Southeast Brazil. Cancer 117, 2228–2235 (2011).
10. Seidinger, A. L. et al. Occurrence of neuroblastoma among TP53 p.R337H carriers. PLoS ONE 10, 1–15 (2015).
11. Mastellaro, M. J. et al. Contribution of the TP53 R337H mutation to the cancer burden in southern Brazil: insights from the study of 55 families of children with adrenocortical tumors. Cancer 123, 3150–3158 (2017).
12. Giacomazzi, J. et al. Prevalence of the TP53 p.R337H mutation in breast cancer patients in Brazil. PLoS ONE 9, 6–12 (2014).
13. Achatz, M. I. W. et al. The TP53 mutation, R337H, is associated with Li-Fraumeni and Li-Fraumeni-like syndromes in Brazilian families. Cancer Lett. 245, 96–102 (2007).
14. Gomes, M. C. B. et al. The TP53 R337H mutation in TP53 and breast cancer in Brazil. Hered. Cancer Clin. Pract. 10, 3 (2012).
15. Cury, N. M., Ferraz, V. E. F. & Silva, W. A. TP53 p.R337H prevalence in a series of Brazilian hereditary breast cancer families. Hered. Cancer Clin. Pract. 12, 1–8 (2014).
16. Silva, F. C. et al. Hereditary breast and ovarian cancer: assessment of point mutations and copy number variations in Brazilian patients. BMC Med. Genet. 15, 1–11 (2014).
17. da Formiga, M. N. C., De Andrade, K. C., Kowalski, L. P. & Achatz, M. I. Frequency of thyroid carcinoma in Brazilian TP53 p.R337H carriers with Li Fraumeni syndrome. JAMA Oncol. 3, 1400–1402 (2017).
18. Macedo, G. S. et al. p53 signaling pathway polymorphisms, cancer risk and tumor phenotype in TP53 R337H mutation carriers. Fam. Cancer 17, 269–274 (2018).
19. Ferreira, A. M. et al. Clinical spectrum of Li-Fraumeni syndrome/Li-Fraumeni-like syndrome in Brazilian individuals with the TP53 p.R337H mutation. J. Steroid Biochem. Mol. Biol. 190, 250–255 (2019).
20. Giacomazzi, J. et al. The TP53 R337H mutation is a conditional cancer-predisposing mutation: further evidence from a homozygous patient. BMC Cancer 13, 1–8 (2013).
21. Andrade, K. C. et al. Early-onset breast cancer patients in the South and Southeast of Brazil should be tested for the TP53 p.R337H mutation. Genet. Mol. Biol. 39, 199–202 (2016).
22. Palmero, E. I. et al. Detection of R337H, a germline TP53 mutation predisposing to multiple cancers, in asymptomatic women participating in a breast cancer screening program in Southern Brazil. Cancer Lett. 261, 21–25 (2008).
23. Jordan, J. C. et al. Altered-function p53 missense mutations identified in breast cancers can have subtle effects on transactivation. Mol. Cancer Res. 8, 701–716 (2010).
24. Fitarelli-Kiehl, M. et al. The breast cancer immunophenotype of TP53-p.R337H carriers is different from that observed among other pathogenic TP53 mutation carriers. Fam. Cancer 14, 333–336 (2015).
25. Hanahan, D. & Weinberg, R. A. The hallmarks of cancer. Cell 100, 57–70 (2000).
26. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. Cell 144, 646–674 (2011).
27. vlachos, I. S. et al. DIANA-miRPath v3.0: deciphering microRNA function with experimental support. Nucleic Acids Res. 43, W460–W466 (2015).
28. Huang, H. Y. et al. miRtarBase 2020: updates to the experimentally validated microRNA-target interaction database. Nucleic Acids Res. 48, D148–D154 (2020).
29. Figueiredo, B. C. et al. Comparative genomic hybridization analysis of adrenocortical tumors of childhood 1. J. Clin. Endocrinol. Metab. 84, 1116–1121 (1999).
30. Pinto, E. M. et al. Genomic landscape of paediatric adrenocortical tumours. Nat. Commun. 6, 1–10 (2015).
31. Cavalli, L. R. et al. Genetic and epigenetic alterations in sentinel lymph nodes metastatic lesions compared to their corresponding primary breast tumors. Cancer Genet. Cytogenet. 146, 33–40 (2003).
32. Santos, C. L. et al. Patterns of DNA copy number changes in sentinel lymph node breast cancer metastases. Cytogenet. Genome Res. 122, 16–21 (2008).
33. Calin, G. A. et al. Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. Proc. Natl. Acad. Sci. 101, 2999–3004 (2004).
34. Zhang, L. et al. microRNAs exhibit high frequency genomic alterations in human cancer. Proc. Natl. Acad. Sci. 103, 9136–9141 (2006).
35. Landau, D.-A. & Slack, F. J. MicroRNAs in mutagenesis, genomic instability, and DNA repair. Semin. Oncol. 38, 743–751 (2011).
36. Vincent, K., Pichler, M., Lee, G. W. & Ling, H. MicroRNAs, genomic instability and cancer. Int. J. Mol. Sci. 15, 14475–14491 (2014).
37. Oliveto, S., Mancino, M., Manfrini, N. & Biffo, S. Role of microRNAs in translation regulation and cancer. World J. Biol. Chem. 8, 45 (2017).
38. Elston, R. & Inman, G. J. Crosstalk between p53 and TGF-β Signalling. J. Signal Transduct. 2012, 1–10 (2012).
39. Liu, W. T. et al. TGF-β upregulates the translation of USP15 via the PI3K/ AKT pathway to promote p53 stability. Oncogene 36, 2715–2723 (2017).
40. Suzuki, H. I. MicroRNA control of TGF-β signaling. Int. J. Mol. Sci. 19, 1901 (2018).
41. Li, K. et al. An integrated approach to reveal miRNAs’ impacts on the functional consequence of copy number alterations in cancer. Sci. Rep. 5, 1–13 (2015).
42. Sugita, B. et al. Differentially expressed microRNAs in triple negative breast cancer between African-American and non-Hispanic white women. Oncotarget 7, 79274–79291 (2016).
43. Soh, J., Cho, H., Choi, C. H. & Lee, H. Identification and characterization of microRNAs associated with somatic copy number alterations in cancer. Cancers (Basel) 10, 1–18 (2018).
44. Sugita, B. M. et al. Integrated copy number and miRNA expression analysis in triple negative breast cancer of Latin American patients. Oncotarget 10, 6184–6203 (2019).
45. Luo, Z., Cui, R., Tili, E. & Croce, C. Friend or foe: microRNAs in the p53 network. Cancer Lett. 419, 96–102 (2018).
46. Donzelli, S. et al. MicroRNA-128-2 targets the transcriptional repressor E2F5 enhancing mutant p53 gain of function. Cell Death Differ. 19, 1038–1048 (2012).
47. Dong, P. et al. Mutant p53 gain-of-function induces epithelial-mesenchymal transition through modulation of the miR-130b-ZEB1 axis. Oncogene 32, 3286–3295 (2013).
48. Wang, W., Cheng, B., Miao, L., Me, Y. & Wu, M. Mutant p53–R273H gains new function in sustained activation of EGFR signaling via suppressing miR-27a expression. Cell Death Dis. 4, 1–11 (2013).
49. Masciarelli, S. et al. Gain-of-function mutant p53 downregulates miR-223 contributing to chemoresistance of cultured tumor cells. Oncogene 33, 1601–1608 (2014).
50. Cristina, I. et al. MIR605 rs2043556 is associated with the occurrence of multiple primary tumors in TP53 p. (Arg337His) mutation carriers. Cancer Genet. 240, 54–58 (2020).
51. Matsu, A. et al. Oxidation resistance 1 prevents genome instability through maintenance of G2/M arrest in gamma-ray-irradiated cells. J. Radiat. Res. 61, 1–13 (2020).
52. Rosati, R. et al. High frequency of loss of heterozygosity at 11p15 and IGF2 overexpression are not related to clinical outcome in childhood adrenocortical tumors positive for the R337H TP53 mutation. Cancer Genet. Cytogenet. 186, 19–24 (2008).
53. Silva, A. G., Achatz, I. M. W., Krepischi, A. C., Pearson, P. L. & Rosenberg, C. Number of rare germline CNVs and TP53 mutation types. Orphanet J. Rare Dis. 7, 1 (2012).
54. Letouzé, E. et al. SNP array profiling of childhood adrenocortical tumors reveals distinct pathways of tumorigenesis and highlights candidate driver genes. J. Clin. Endocrinol. Metab. 97, 1284–1293 (2012).
55. Silva, A. G. et al. The profile and contribution of rare germline copy number variants to cancer risk in Li-Fraumeni patients negative for TP53 mutations. Orphanet J. Rare Dis. 9, 2–7 (2014).
56. Letouzé, E. et al. Identity by descent mapping of founder mutations in cancer using high-resolution tumor SNP data. PLoS ONE 7, 1–11 (2012).
57. Torresan, C. et al. Increased copy number of the DLX4 homeobox gene in breast axillary lymph node metastasis. Cancer Genet. 207, 177–187 (2014).
58. Bengtsson, H., Simpson, K., Bullard, J. & Hansen, K. aroma.affymetrix: a generic framework in R for analyzing small to very large Affymetrix data sets in bounded memory. Methods, Technical Report, 1–9 (2008).
59. Wang, K. et al. PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. Genome Res. 17, 1665–1674 (2007).

Acknowledgements
The authors would like to thank the personnel from the Genomics and Epigenomics Shared Resource of the Lombardi Comprehensive Cancer Center for performing the washes and scanning of the array-CGH slides. We are grateful to José Renato Sales Barbosa for his technical support. This study was partially funded by PRONEX, and PPSUS – Fundação Araucária-CNPq-MS. Scholarship to CM was provided by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (CAPES/PROAP)—Finance Code 001. We also acknowledge the support of the Genomics and Epigenomics Shared Resources (GEsR) at LCCC, partially supported by the NIH/NCI Grant P30-CA051008.

Author contributions
B.C.F., I.J.C., C.A.U., L.R.C., and E.M.F.S.R. conceived the study; C.M., H.K. and A.S.F. conducted the experiments; C.M., S.B., A.C., C.P. and I.J.C. performed the data analysis, R.S.L., C.A.U., A.P.S., and E.W.S.S. assisted in the samples collection; C.M., S.B., A.C., B.C.F., I.J.C., L.R.C., and E.M.F.S.R. prepared the original draft; B.C.F., I.J.C., C.A.U., L.R.C., and E.M.F.S.R. performed the review and editing of the final manuscript version. All authors have read and agreed to the published version of the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-73282-y.
Correspondence and requests for materials should be addressed to L.R.C.
Reprints and permissions information is available at www.nature.com/reprints.
Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2020