



GATA3 Truncation Mutants Alter EMT Related Gene Expression via Partial Motif Recognition in Luminal Breast Cancer Cells

Mika Saotome[†], Deepak Balakrishnan Poduval[†], Renju Nair[†], Mikhala Cooper and Motoki Takaku*

Department of Biomedical Sciences, University of North Dakota School of Medicine and Health Sciences, Grand Forks, ND, United States

OPEN ACCESS

Edited by:

Ata Abbas,
Case Western Reserve University,
United States

Reviewed by:

Pabitra Kumar Parua,
Albert Einstein College of Medicine,
United States
Babu Roshan Padmanabhan,
University Hospitals Cleveland Medical
Center, United States
Heeyoun Bunch,
Kyungpook National University, South
Korea

*Correspondence:

Motoki Takaku
motoki.takaku@und.edu

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Cancer Genetics and Oncogenomics,
a section of the journal
Frontiers in Genetics

Received: 23 November 2021

Accepted: 10 January 2022

Published: 28 January 2022

Citation:

Saotome M, Poduval DB, Nair R,
Cooper M and Takaku M (2022)
GATA3 Truncation Mutants Alter EMT
Related Gene Expression via Partial
Motif Recognition in Luminal Breast
Cancer Cells.
Front. Genet. 13:820532.
doi: 10.3389/fgene.2022.820532

GATA3 is known to be one of the most frequently mutated genes in breast cancer. More than 10% of breast tumors carry mutations in this gene. However, the functional consequence of GATA3 mutations is still largely unknown. Clinical data suggest that different types of GATA3 mutations may have distinct roles in breast cancer characterization. In this study, we have established three luminal breast cancer cell lines that stably express different truncation mutants (X308 splice site deletion, C321 frameshift, and A333 frameshift mutants) found in breast cancer patients. Transcriptome analysis identified common and distinct gene expression patterns in these GATA3 mutant cell lines. In particular, the impacts on epithelial-to-mesenchymal transition (EMT) related genes are similar across these mutant cell lines. Chromatin localization of the mutants is highly overlapped and exhibits non-canonical motif enrichment. Interestingly, the A333 frameshift mutant expressed cells displayed the most significant impact on the GATA3 binding compared to X308 splice site deletion and C321fs mutants expressed cells. Our results suggest the common and different roles of GATA3 truncation mutations during luminal breast cancer development.

Keywords: breast cancer, GATA3, GATA-3, somatic mutations, luminal breast cancer, epithelial-to-mesenchymal transition

INTRODUCTION

Breast cancer is the most common cancer among women in the U.S. and the second leading cause of cancer-related deaths. It was projected that there would be about 284,200 new cases of invasive breast cancer and the deaths of approximately 43,600 women in the U.S. from breast cancer in 2021 (American Cancer Society, 2018). GATA3 is a reliable biomarker for breast carcinomas and is frequently used to determine the tissue of origin to confirm a diagnosis (Jensen et al., 2002; Garcia-Closas et al., 2007; Bertucci et al., 1999; Liu et al., 2016; Hoch et al., 1999; Mehra et al., 2005; Sorlie et al., 2003; Chou et al., 2010; Takaku et al., 2015). Recent large-scale molecular profiling of breast carcinomas identified frequent mutations in GATA3 (Usary et al., 2004; The Cancer Genome Atlas Network, 2012). GATA3 is a transcription factor and is known to be involved in multiple developmental pathways and human diseases such as normal mammary gland development, breast cancer progression, and T cell differentiation (Zheng and Flavell, 1997; Kouros-Mehr et al., 2006; Chou et al., 2010). In breast cancer, it has been shown that the cooperative action

between GATA3 and its cofactors Estrogen Receptor alpha (ER- α) and FOXA1 is important for the luminal breast cancer characterization (Kong et al., 2011; Theodorou et al., 2013; Takaku et al., 2020). Based on the METABRIC cohort (Pereira et al., 2016), among 1980 patient cases, 230 breast carcinomas harbored GATA3 mutations (~11.6%). They observed 75% of the mutations in luminal tumors (47% in luminal A, 28% in luminal B). These data suggest that approximately 50,000 new cases of invasive breast cancer in the U.S. will carry GATA3 mutations. While patient genomic data suggests GATA3 mutations as cancer drivers, the functional consequences of GATA3 mutations in breast cancer are underexplored (Adomas et al., 2014; Mair et al., 2016; Gustin et al., 2017; Emmanuel et al., 2018; Takaku et al., 2018). More importantly, the frequency of somatic mutations in GATA3 was even higher in the metastatic breast cancer cohort (Bertucci et al., 2019). GATA3 mutant breast cancer patients had lung, lymph nodes, and brain metastases (Bruna et al., 2016; Bertucci et al., 2019). The GATA3 mutant tumors tend to occur in younger patients (<45 years of age) (Azim et al., 2015; Griffith et al., 2018). The development of tamoxifen resistance correlates with the GATA3 gene silencing (Feng et al., 2014; Bi et al., 2020). These facts highlight the importance of the functional study of GATA3 and its mutations in breast cancer.

We previously identified that patients carrying GATA3 mutations have diverse clinical features. More than 70% of cases are small nucleotide deletions or insertions (indel), while less than 30% are missense mutations. By classifying the GATA3 indel mutations into four groups, we observed distinct clinical features (Takaku et al., 2018). Somatic mutations found in the GATA3 second zinc-finger domain (ZnFn2) are associated with poorer patient outcomes and low survival rates. The ZnFn2 domain is known to be important for the DNA binding and transcription activation activities of the GATA3 protein. ZnFn2 mutations are predominantly found in luminal B breast tumors, while splice site mutations are frequently found in luminal A breast tumors and are associated with better patient survival. These distinct clinical features suggest the differential impacts of GATA3 mutations on breast cancer cells. However, the biological significance and consequences, including the cellular function of these mutants, are not well-studied (Usary et al., 2004; Gustin et al., 2017; Emmanuel et al., 2018; Takaku et al., 2018; Hruschka et al., 2020; Takaku et al., 2020). In this study, we have established three luminal breast cancer cell lines that stably express different GATA3 truncation mutants found in breast cancer. We explored the function of these mutants using genomic approaches. These results revealed the common and distinct impacts of GATA3 truncation mutants on luminal breast cancer cells.

MATERIALS AND METHODS

Cell Line and Cell Culture

T47D cells (originally purchased from ATCC) were maintained in Gibco DMEM high glucose medium (Thermo) supplemented with 10% FBS (Atlanta). GATA3 mutant genes were generated by site-directed mutagenesis using the primers (**Supplementary Figure S1A**) and cloned

into pHAGE lentiviral vectors (a kind gift from Dr. Guang Hu at NIEHS/NIH). The viruses were produced by transient transfection using 293T cells, psPAX2, and pMD2.G packaging vectors (psPAX2 and pMD2.G were gifts from Dr. Didier Trono, Addgene plasmid 12260, 12259). After infection, with lentiviruses encoding each GATA3 mutant, the stable cell pools were established by puromycin selection for 1 week. T47D cells infected with the lentivirus containing the pHAGE empty vector were used as a control cell line.

Cell Proliferation Assay

T47D cells were resuspended in Gibco DMEM high glucose medium (Thermo) containing 10% FBS (Atlanta). Fifty thousand cells per well were plated in a 24-well plate. At Days 1, 3, and 5, cells were harvested and then counted by DeNovix CellDrop.

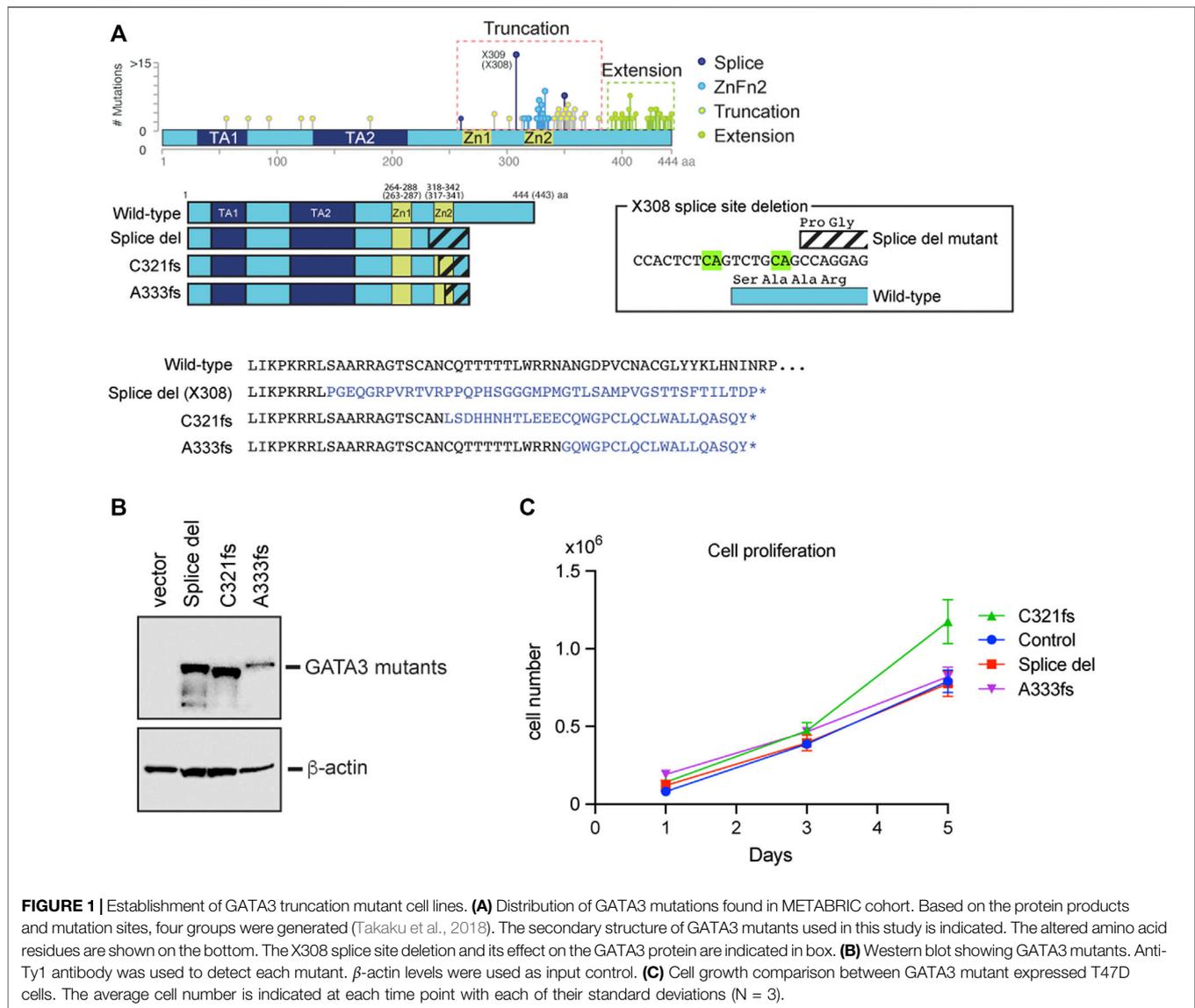
RNA-Seq

RNA was purified by the Direct-zol RNA Miniprep Kit (Zymo). RNA-seq libraries were prepared by the UND genomics core using NEBNext Ultra II RNA-Seq Kit (NEB). TapeStation (Agilent technologies) was used to check the RNA integrity and library quality. Pooled libraries were sequenced in one lane Nova-Seq S4 as 150 bp paired-end reads. RNA-Seq data were analyzed using an in-house pipeline. In the first step, reads were trimmed based on quality, and adapters were removed using Trimmomatic (v0.39) (Bolger et al., 2014). Quality trimmed reads were aligned to the human genome (hg38), and raw read counts were obtained using featureCounts from Rsubread (Liao et al., 2014). Raw read counts were processed using a custom R script, and differential expression levels were quantified using DESeq2 (Love et al., 2014). Differentially expressed genes were annotated using clusterprofiler (Yu et al., 2012) and wikipathways (Slenter et al., 2018).

ChIP-Seq and Peak Analysis

The details of the procedures were previously described (Takaku et al., 2018; Takaku et al., 2020). T47D cells were fixed with 1% formaldehyde at 37°C for 10 min with constant shaking. Fixed cells were incubated with hypotonic buffer containing 10 mM HEPES-NaOH pH 7.9, 10 mM KCl, 1.5 mM MgCl₂, 340 mM sucrose, 10% glycerol, 0.5% Triton X-100 and Halt protease and phosphatase Inhibitor Cocktail (Thermo Fisher Scientific). After centrifugation, the cells were lysed with the sonication buffer containing 20 mM Tris-HCl pH 8.0, 2 mM EDTA, 0.5 mM EGTA, 0.5 mM PMSF, 5 mM sodium butyrate, 0.1% SDS and protease inhibitor cocktail. Nuclei were sonicated by Covaris S220 for 10 min.

Approximately 15 μ g of chromatin was used for immunoprecipitation. For immunoprecipitation of the GATA3 mutants, 2.5 μ g of anti-Ty1 antibody (Diagenode, C15200054) was used. For a total GATA3 pull-down, 2.5 μ L of anti-GATA3 antibody (Cell Signaling, D13C9) was used in each assay. After overnight incubation, Protein G (for Ty1 antibody) or Protein A and G mix (for GATA3 antibody) Dynabeads (Thermo) were added to the chromatin solution. DNAs were purified by AMPure XP (Beckman Coulter). ChIP-seq libraries were prepared by NEXTFLEX Rapid DNA-Seq Kit.



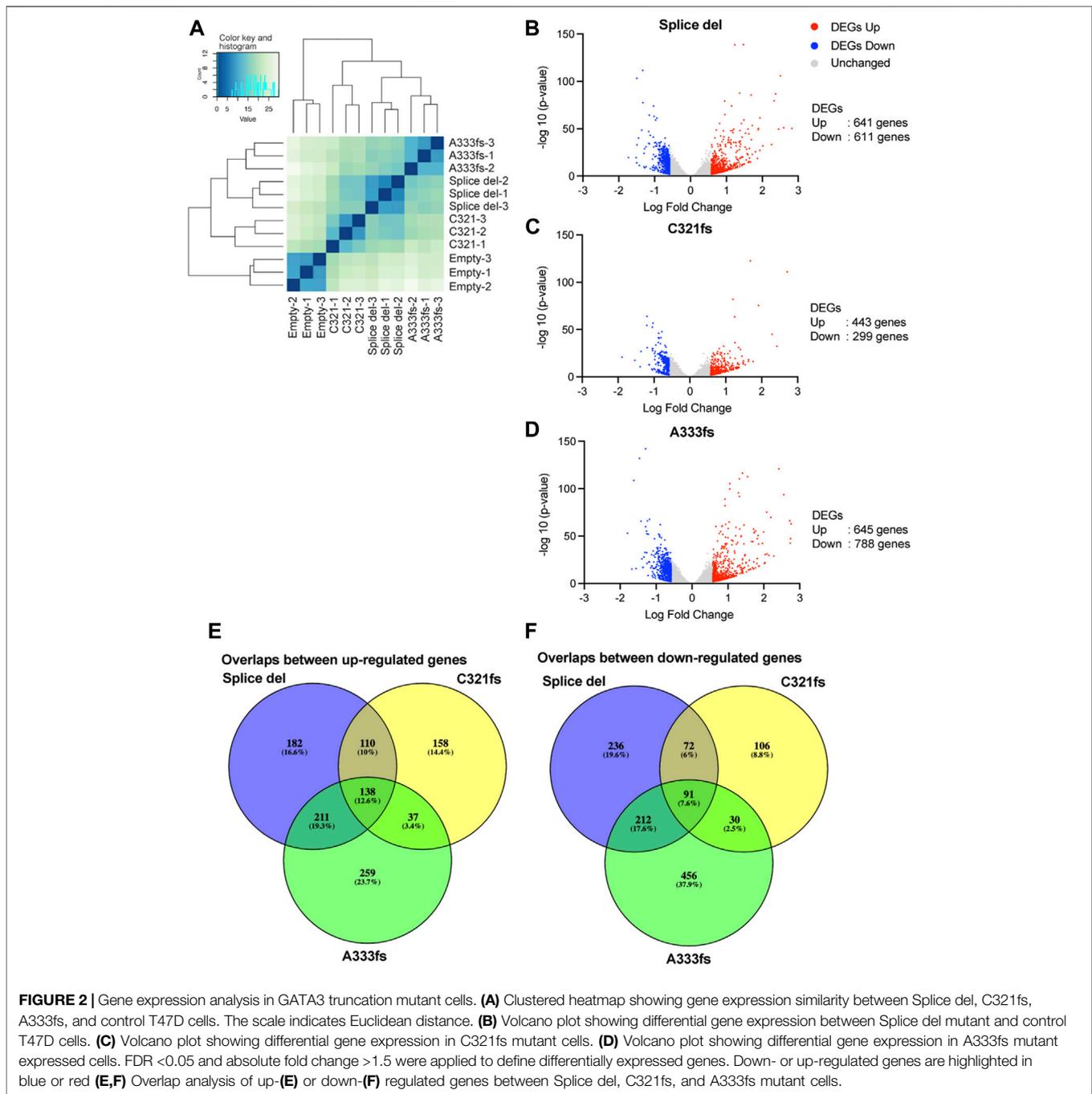
ChIP-Seq reads were quality filtered and adapter trimmed using in-house scripts. The trimmed reads were mapped to the human genome (hg19) using Bowtie (version 1.2.2) (Langmead et al., 2009). Uniquely mapped reads were marked for duplicates with the Picard tools (Broad Institute, 2019). For subsequent analysis, paired reads were merged into single fragments, and coverage tracks were obtained using genomeCoverageBed from bedtools (v2.29.0) (Quinlan and Hall, 2010). ChIP-seq peaks were identified using HOMER (Heinz et al., 2010) with default parameters. Overlap between peak sets was determined using intersectBed from bedtools (v2.29.0). Differential binding events between samples were identified using EdgeR (Robinson et al., 2010) with FDR <0.05 and absolute fold change >1.5. Motif analysis was carried out using MEME-ChIP from The MEME Suite (Machanick and Bailey, 2011). The HOMER tool Perl script (findMotifs.pl) was used to find each motif frequency within the peaks. Read counts were collected in 20 bp bins, then normalized

to 15 million total non-duplicate unique fragments per dataset for regions ± 1 kb relative to peak midpoints.

RESULTS

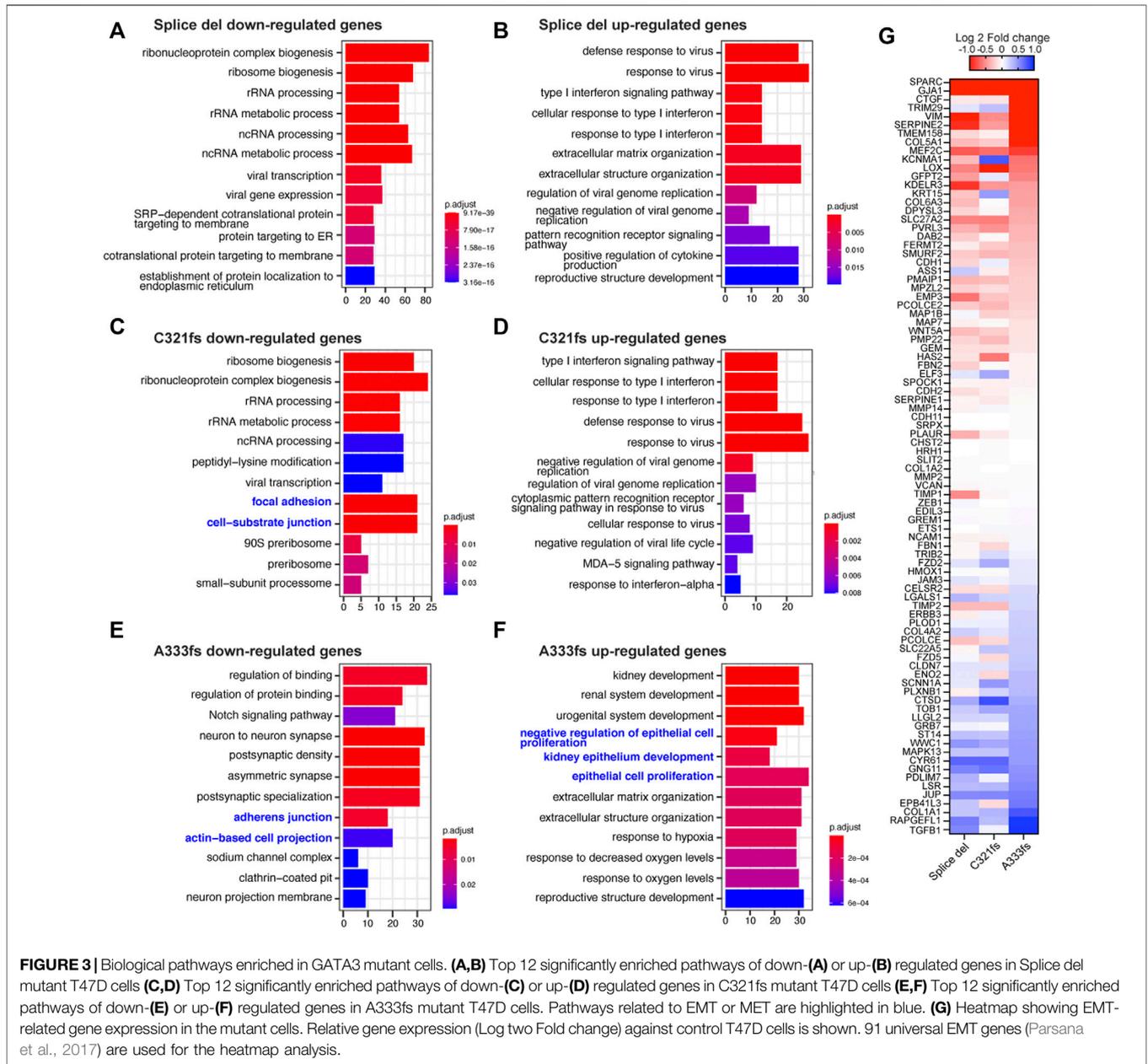
Establishment of GATA3 Truncation Mutant Cell Lines

To better understand the roles of GATA3 mutants in breast cancer cells, we established three luminal breast cancer cell lines stably expressing X308 splice site deletion mutant (Splice del), C321 frameshift mutant (C321fs), and A333 frameshift mutant (A333fs), respectively (Figure 1A). Splice del mutation is one of the hot spot mutations found in breast cancer. Two nucleotide (CA) deletion at the splice acceptor site induces alternative splicing by using a new splice acceptor site seven nucleotides downstream. This results in a



frameshift and protein truncation due to the immature stop codon (Takaku et al., 2015; Hruschka et al., 2020). The protein product entirely loses the original (wild-type) amino acid sequence of the ZnFn2 DNA binding domain. C321fs and A333fs mutations were found in the ZnFn2 domain and belong to the ZnFn2 mutation (**Figure 1A**). Again, these frameshift mutations result in protein truncation due to the immature stop, and they partially lack the ZnFn2 wild-type sequence. We exogenously

expressed these mutants as 3xTy-1 fused protein in T47D cells by lentiviral transduction. After the positive clone selection by Puromycin, we looked at protein expression levels by western blot. The expression levels of Splice del and C321fs were similar, while the A333fs protein level was slightly lower (**Figure 1B**). All GATA3 mutant cell lines exhibited synonymous wild-type GATA3 expression levels compared to the control cell line (**Supplementary Figure 1B**). To investigate the roles of GATA3 mutants in

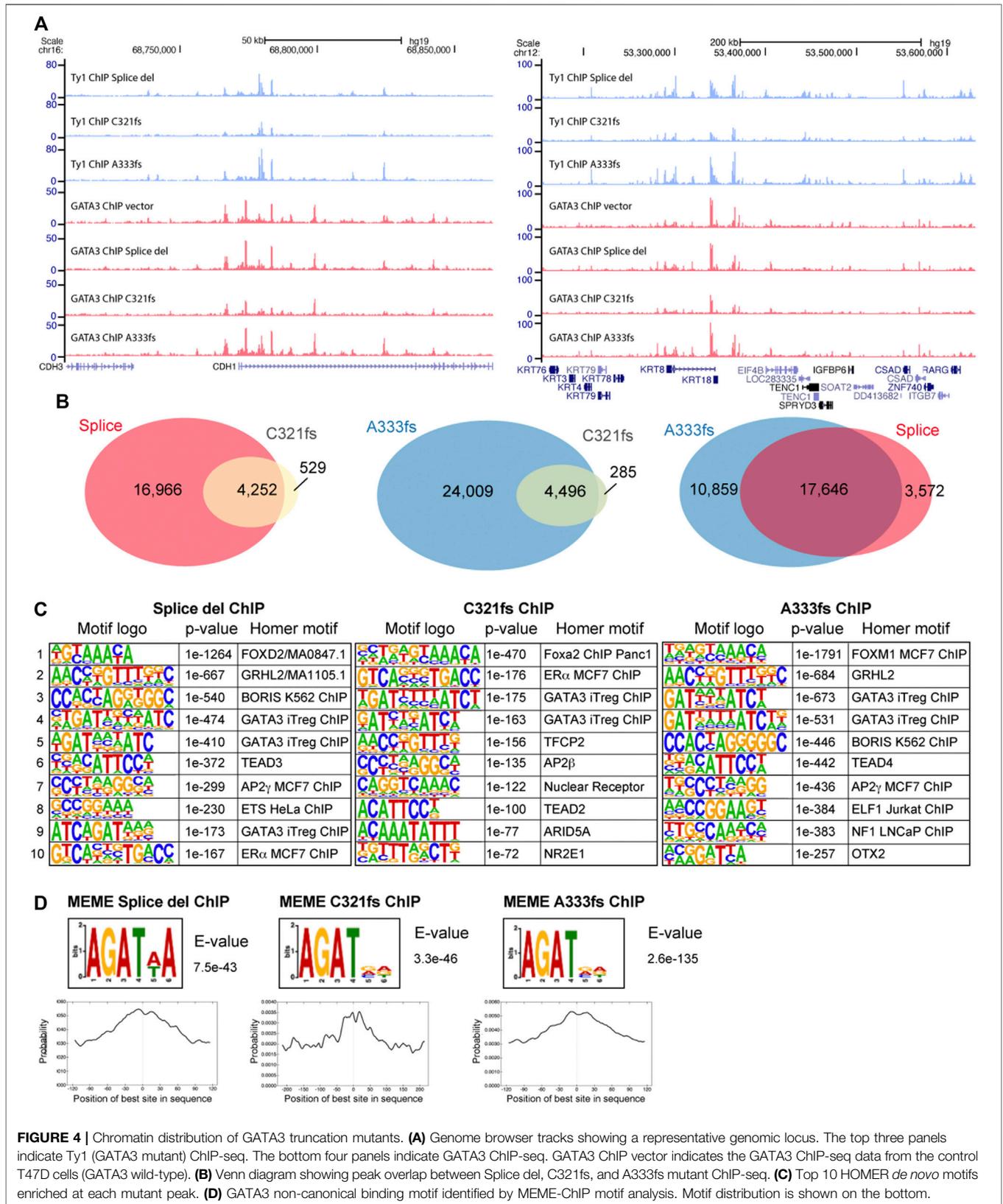


cancer cell proliferation, we performed a cell proliferation assay. Although the cell number of the C321fs expressed cells on Day 5 was slightly higher, overall, the stable expression of these GATA3 truncation mutants did not exhibit a significant impact on T47D cell growth (**Figure 1C**).

GATA3 Truncation Mutants Influence Luminal Breast Cancer Transcriptome

To identify the impacts of GATA3 mutant expression on gene expression, RNA-seq was performed for each of our established cell lines. The PCA plot and clustering analysis showed the high similarity between biological replicates (**Figure 2A**,

Supplementary Figure S1C). The mutant expressing cell lines have distinct transcriptome profiles from the control T47D cells. Despite the weaker expression of A333fs mutant, A333fs and Splice del mutant T47D cells share similarities compared to C321fs mutant expressing cells or control cells. We then defined differentially expressed genes at a false discovery rate (FDR) < 0.01 and |fold change| > 1.5 (**Supplementary Table S1**). In Splice del mutant cells, there were 641 up-regulated genes and 611 down-regulated genes compared to control T47D cells (**Figure 2B**). C321fs mutant expressing cells presented 443 up- and 299 down-regulated genes (**Figure 2C**), while A333fs mutant expressing cells presented 645 up- and 788 down-regulated genes (**Figure 2D**). Among these differentially expressed genes, 138



were up-regulated across all three GATA3 mutant cell lines (Figure 2E), while 91 genes were commonly down-regulated in the mutant cells (Figure 2F). These data suggest that although Splice del, C321fs, and A333fs mutations result in a similar protein truncation, the impacts on luminal breast cancer transcriptome are partially different.

To understand biological pathways enriched in the mutant expressed T47D cells, we performed GO term analysis. In C321fs and A333fs cells, the down-regulated genes were enriched for cell junction pathways (Figures 3C–E). Down-regulated genes in A333fs cells showed enrichment for the Notch signaling pathway and neuron-related pathways, while only Splice del and C321fs cells exhibited significant enrichment of ribosomal RNA processing pathways (Figures 3A,C,E). Among up-regulated genes, immune response pathways, including type I interferon and viral response networks, were enriched in Splice del and C321fs cells (Figures 3B,D). In A333fs cells, epithelial proliferation and development-related genes were enriched in up-regulated genes (Figure 3F). Since GATA3 is known to be a critical regulator of mammary epithelium development, EMT and its reverse process, mesenchymal-to-epithelial transition (MET), we specifically looked at genes involved in these pathways (Figure 3G). 91 universal EMT genes identified by Dr. Battle's group (Lien et al., 2007) were used for this gene expression analysis. Although more distinct impacts on EMT-related genes were observed in A333fs cells, all three GATA3 mutant cells exhibit similar alteration patterns. For instance, some of the mesenchymal marker genes such as VIM, LOX, and SPARC (Lien et al., 2007; El-Haibi et al., 2012) were up-regulated in the mutant cell lines, while the expression levels of other EMT markers such as ZEB1 were not increased in the mutant cells, suggesting a partial EMT phenotype (Aiello and Kang, 2019). Taken together, these transcriptome analyses indicate that Splice del, C321fs, and A333fs mutants can modulate the expression levels of EMT-related genes.

Non-Canonical Chromatin Binding by GATA3 Truncation Mutants

Using ChIP-seq, the chromatin distribution of GATA3 mutants was determined to identify how the truncation mutants alter transcription profiles in T47D cells. Since all mutants expressed as Ty1 epitope tag fusion proteins, we first performed Ty1 ChIP-seq specifically to map the chromatin localization of the mutants. Even though Splice del, C321fs, and A333fs mutations completely or partially lack the intact ZnFn₂ domain, we observed many common peaks and that all three mutants bind to chromatin (Figure 4A). Peak call analysis by HOMER identified 21,218 Splice del peaks, 28,505 A333fs peaks, and 4,781 C321fs peaks. We further confirmed the mutant distributions and peak numbers by comparing the data between biological replicates (Supplementary Figures S2A–C). Since the fewer peaks in the C321fs ChIP-seq data suggest the weaker chromatin binding of C321fs mutant, we compared the signal enrichment in each mutant ChIP-seq data. Metaplot analysis of Ty1 ChIP-seq data clearly suggests that lower enrichment of C321fs compared to Splice del and A333fs (Supplementary Figures S2D,E). Peak

overlap analysis indicated that the genomic distributions of Splice del, C321fs, and A333fs mutants were largely overlapped (Figures 4A,B). We also observed each mutant specific localization. Based on the peak overlap between A333fs and Splice del mutants, 10,859 peaks (~37%) were uniquely enriched in A333fs ChIP-seq data, and 3,572 peaks (~17%) did not overlapped with A333fs peaks. Since the ZnFn₂ is known to be important for the consensus DNA motif binding, and we previously showed the altered DNA binding activities of R330fs and D336fs GATA3 mutants (Adomas et al., 2014; Takaku et al., 2018), we performed *de novo* motif analysis to look at the enriched transcription factor binding motifs (Figure 4C). Interestingly, all mutant ChIP-seq peaks exhibited significantly enriched Forkhead motifs. ERE motifs were also enriched in Splice del (ranked 10th), C321fs (ranked second), and A333fs (ranked 13th, *p*-value = 1e-159) mutant ChIP-seq data. Since FOXA1 and ER- α are well-known as GATA3 co-factors (Kong et al., 2011; Theodorou et al., 2013), the enrichment of these co-factors' motifs may indicate the importance of FOXA1 and ER- α for the chromatin binding activities of the GATA3 mutants (Adomas et al., 2014; Takaku et al., 2020). In addition to these co-factors, other transcription factor binding motifs (e.g., GRHL2, BORIS, TEAD4 AP2 gamma) were also enriched in the multiple data set. Importantly, the enriched GATA3 motifs identified by HOMER seem to lack the consensus motif DNA sequences (WGATAR) but show a partial motif, WGAT. To further confirm the partial motif enrichment, we performed the motif analysis by MEME (Machanick and Bailey, 2011). The MEME *de novo* motif analysis also showed a similar sequence enrichment, AGAT, at all three mutant peaks (Figure 4D). The center of each mutant peak aligned well with distribution of this non-canonical motif.

GATA3 Truncation Mutants Alter GATA3 Chromatin Distribution

The established GATA3 mutant cell lines still express wild-type GATA3, which mimics the situation in breast tumors since most of the GATA3 mutations found in patients are heterozygous. To determine the impact of GATA3 mutant chromatin binding on the overall GATA3 distribution in T47D cells, we performed GATA3 ChIP-seq using the antibody that recognizes both wild-type and mutants (Figure 4A, Supplementary Figure S1B). In control T47D cells (GATA3 wild-type), the consensus GATA3 motif, WGATAA, was the most significantly enriched in MEME-ChIP analysis. Similarly, the most significant enriched sequence in the mutant cells was the same consensus motif, suggesting that the overall sequence preference of GATA3 is maintained in the presence of GATA3 truncation mutants. As shown in Figure 4A, there are frequent overlaps between Ty1 GATA3 mutant ChIP-seq and the total GATA3 ChIP-seq. We also observed Ty1 ChIP-seq unique peaks or total GATA3 ChIP-seq unique peaks (such as around the CDH1 transcription start site). These differential signals might be due to the relatively lower expression of the mutants compared to the wild-type GATA3 protein (Supplementary Figure S1B) or differences in antibody specificity. To identify differential GATA3 binding in the mutant cells, we first defined GATA3 peaks in each cell line

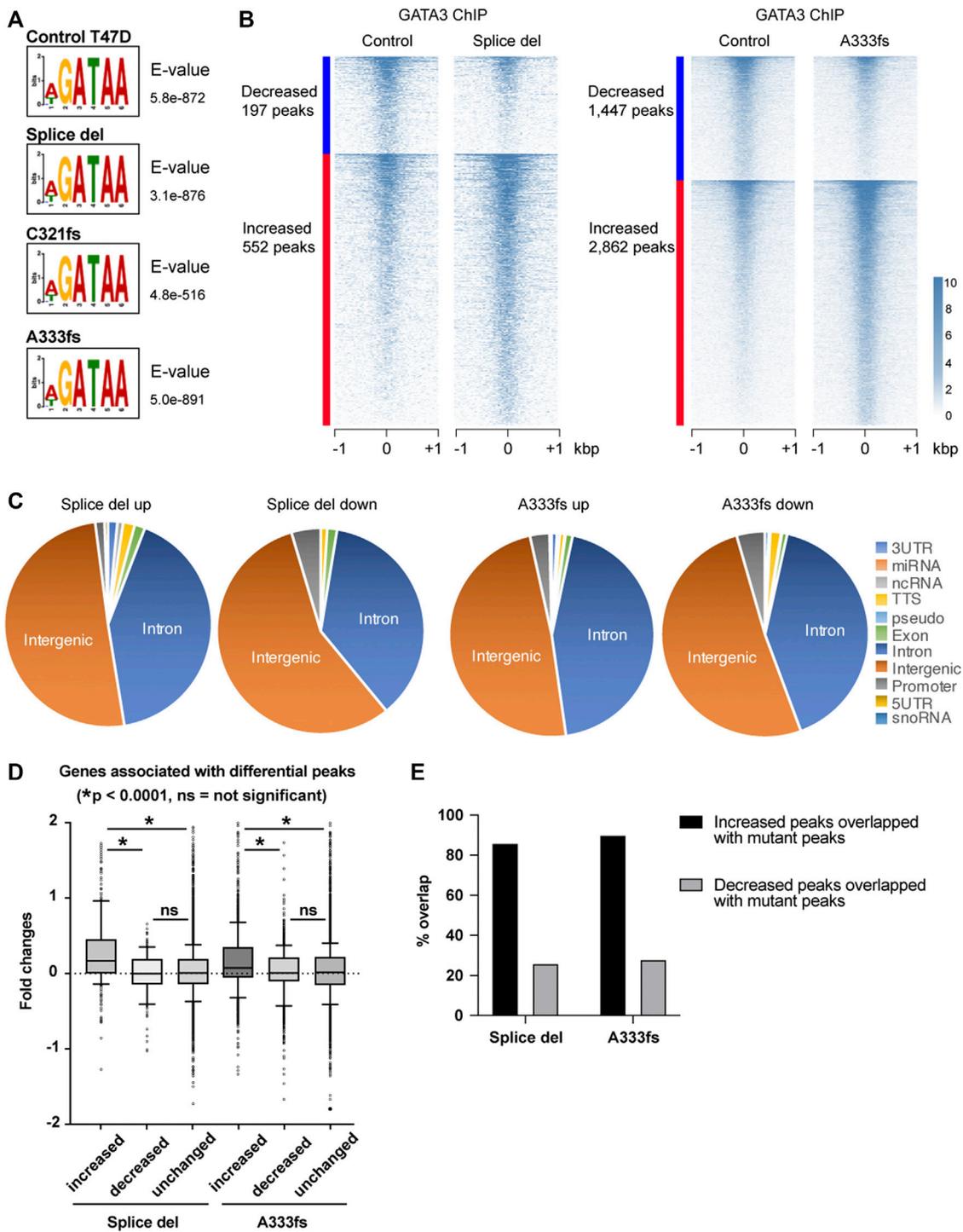


FIGURE 5 | GATA3 re-distribution in GATA3 mutant cells. **(A)** The most significantly enriched motif by MEME ChIP analysis. **(B)** Heatmap showing decreased or increased GATA3 binding in Splice del (left) or A333fs mutant cells. **(C)** Genomic localization of differential peaks. **(D)** Correlation analysis of GATA3 differential peaks and gene expression. Increased, decreased, and unchanged peaks are assigned to nearest genes. Box plots show the fold changes of each GATA3 peak group associated genes. Fold changes were calculated between Splice del and control T47D cells or A333fs and control T47D cells. **(E)** Overlap between differential peaks and Splice del or A333fs peaks.

by HOMER and performed the metaplot analysis at the GATA3 peaks. For the metaplot analysis, we generated the union GATA3 peak set (41,538 peaks) by merging all GATA3 peaks detected in control, Splice del, C321fs, and A333fs T47D cells. Metaplot analysis at the union GATA3 peaks indicates that weaker GATA3 enrichment in C321fs T47D cells and stronger GATA3 enrichment in Splice site or A333fs T47D cells compared to control T47D cells (**Supplementary Figure S3A**). To further identify the impacts of these mutants' expression on the GATA3 binding, we performed EdgeR analysis. The total GATA3 binding was largely unchanged in Splice del, C321fs, and A333fs mutant expressing cells compared to the parental control T47D cells. At FDR <0.05 and absolute fold change >1.5, a small set of GATA3 peaks were defined as differential peaks in the mutant cells (**Figure 5B**, **Supplementary Figure S3B–E**). Among the three mutant cell lines, the A333fs mutant cell line had the most GATA3 differential binding events (~10%) despite having the lowest expression level of the mutant (**Figure 1B**). Based on the EdgeR analysis with two biological replicates, 1,447 peaks showed decreased GATA3 ChIP-seq signals, while 2,862 peaks showed increased GATA3 ChIP-seq intensities in A333fs T47D cells (**Figure 5B**). On the other hand, with the same experimental setting, Splice del mutant expressed cells detected 197 decreased and 552 increased GATA3 binding sites compared to the control wild-type T47D cells. Many differential peaks were located at intergenic regions or introns (**Figure 5C**). To explore the association between differential peaks and gene expression, we assigned the differential peaks to the nearest genes to investigate the changes in gene expression between the mutant cells and control cells more closely. In both A333fs and Splice del mutant cells, increased GATA3 peaks were significantly associated with increased gene expression compared to decreased peaks or unchanged peaks, while decreased peaks didn't exhibit a clear correlation with gene down-regulation (**Figure 5D**). To dissect the mechanism of differential GATA3 binding, we investigated the overlap between differential GATA3 peaks and the mutant-specific peaks defined by Ty1 ChIP-seq (**Figure 5E**). Most of the increased GATA3 peaks overlapped with Splice del mutant peaks in Splice del expressed cells and A333fs mutant peaks in A333fs mutant expressed cells. These data suggest that GATA3 truncation mutants influence GATA3 distributions in luminal breast cancer cells leading to differential gene expression.

DISCUSSION

Worldwide breast cancer genome profiling keeps revealing GATA3 as one of the primary targets for somatic mutations in breast cancer. The systemic analysis defined those GATA3 mutations as cancer drivers. However, our knowledge of GATA3 mutations in tumorigenesis, tumor progression, and acquisition of drug resistance is limited (Kouros-Mehr et al., 2006; Chou et al., 2010; Chou et al., 2013; Liu et al., 2016; Bi et al., 2020). We, and other groups, have previously reported that truncation mutations found around the ZnFn₂ domain possess active roles in breast cancer properties and potentially stimulate tumor growth (Usary et al., 2004; Mair

et al., 2016; Gustin et al., 2017; Emmanuel et al., 2018; Takaku et al., 2018; Takaku et al., 2020). The distribution of GATA3 patient mutations is not focal but widely spread in the C-terminal region of the GATA3 gene (Ciriello et al., 2015; Pereira et al., 2016). Some of them, including splice site mutations, completely lack the wild-type sequence of the ZnFn₂ domain (Takaku et al., 2015). While other frameshift mutations partially or fully have the ZnFn₂ domain (**Figure 1A**). Therefore, it is still unclear whether these slight differences in the GATA3 protein structure can produce different outcomes in luminal breast cancer.

In this study, using our newly established stable cell lines, we analyzed the function of three truncation mutants, X308 splice site deletion (Splice del), C321fs, and A333fs in T47D cells. Transcriptome analysis revealed that although each mutant has a distinct impact on a subset of genes, many differentially expressed genes (>50%) are commonly up- or down-regulated in at least two of the three GATA3 mutant cell lines. Among them, EMT-related genes showed similar alterations across three GATA3 mutant cell lines. We previously examined the impacts of a similar truncation mutant R330fs in the same T47D cells using an exogenous expression model as well as an endogenous expression system by CRISPR Cas9 genome editing (Takaku et al., 2018). Similar to Splice del, C321fs, and A333fs mutations, R330fs mutant expression induced altered expression of the EMT genes such as TWIST1. The function of another truncation mutant, D336fs, was also reported by our group and other groups (Adomas et al., 2014; Gustin et al., 2017; Emmanuel et al., 2018). Both R330fs and D336fs expression stimulated tumor growth of estrogen receptor-positive breast cancer cells in the mouse xenograft model. Therefore, GATA3 splice site or frameshift mutations that partially or fully alter amino acid residues of the ZnFn₂ domain may have similar impacts on breast cancer properties and interrupt luminal transcriptome.

To further identify how GATA3 truncation mutants modulate gene expression, we performed GATA3 mutant specific ChIP-seq and total GATA3 (wild-type and mutant) ChIP-seq. Many mutant binding sites were overlapped across Splice del, C321fs, and A333fs ChIP-seq data. We also observed differential binding sites, particularly between Splice del and A333fs mutants. C321fs ChIP-seq data showed fewer peaks compared to the other mutants. These results may suggest slightly different chromatin binding affinities, protein stability, and/or DNA sequence preference. The motif analysis at the mutant peaks revealed a unique motif enrichment, AGAT, which differs from the consensus motif, WGATAA enriched at the total GATA3 ChIP-seq peaks. Interestingly, the R330fs mutant ChIP-seq data also showed similar non-canonical motif enrichment. These results suggest that the partial or complete disruption of the second zinc-finger domain results in the non-conventional chromatin binding of the GATA3 mutants. Based on the total GATA3 ChIP-seq data, such an altered binding specificity contributes to differential GATA3 binding in the GATA3 mutant cells. The motif

analysis of the mutants also showed significant enrichment of the GATA3 well-known co-factors, FOXA1 and Estrogen Receptor alpha (ER- α). Both R330fs and X308 splice site mutants were previously reported to influence the distributions of these GATA3 co-factors. Thus, the enrichment of the co-factors' motifs suggests that Splice del, C321fs and A333fs may alter FOXA1 and/or ER- α localization (Hruschka et al., 2020).

In addition to observing multiple common events between different types of GATA3 mutations, distinct phenotypes were also observed. For instance, in RNA-seq data, C321fs and A333fs cells exhibited enrichment in cell junction-related pathways, while Splice del mutant cells did not show any enrichment. 10,859 A333fs peaks were not overlapped with Splice del mutant peaks. In the case of R330fs, there was a significant reduction of progesterone receptor (PR) expression, but such reduction of PR expression was not detected in Splice del, C321fs, nor A333fs mutant T47D cells. Further investigation is essential for understanding the common and unique roles of GATA3 mutations in breast cancer. It remains elusive if GATA3 can be a therapeutic target or if these GATA3 mutant breast cancer cells are sensitive to specific chemicals (Mair et al., 2016).

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: GEO: GSE190381.

REFERENCES

- Adomas, A. B., Grimm, S. A., Malone, C., Takaku, M., Sims, J. K., and Wade, P. A. (2014). Breast Tumor Specific Mutation in GATA3 Affects Physiological Mechanisms Regulating Transcription Factor Turnover. *BMC cancer* 14, 278. Epub 2014/04/25PubMed PMID: 24758297; PubMed Central PMCID: PMC4021073. doi:10.1186/1471-2407-14-278
- Aiello, N. M., and Kang, Y. (2019). Context-dependent EMT Programs in Cancer Metastasis. *J. Exp. Med.* 216 (5), 1016–1026. Epub 2019/04/13PubMed PMID: 30975895; PubMed Central PMCID: PMC6504222. doi:10.1084/jem.20181827
- American Cancer Society (2018). Cancer Statistics Center. Available at: <http://cancerstatisticscenter.cancer.org>.
- Azim, H. A., Jr., Nguyen, B., Brohée, S., Zoppoli, G., and Sotiriou, C. (2015). Genomic Aberrations in Young and Elderly Breast Cancer Patients. *BMC Med.* 13, 266. Epub 2015/10/16PubMed PMID: 26467651; PubMed Central PMCID: PMC4606505. doi:10.1186/s12916-015-0504-3
- Bertucci, F., Van Hulst, S., Bernard, K., Loriou, B., Granjeaud, S., Tagett, R., et al. (1999). Expression Scanning of an Array of Growth Control Genes in Human Tumor Cell Lines. *Oncogene* 18 (26), 3905–3912. Epub 1999/08/13PubMed PMID: 10445855. doi:10.1038/sj.onc.1202731
- Bertucci, F., Ng, C. K. Y., Patsouris, A., Droin, N., Piscuoglio, S., Carbuca, N., et al. (2019). Genomic Characterization of Metastatic Breast Cancers. *Nature* 569 (7757), 560–564. Epub 2019/05/24PubMed PMID: 31118521. doi:10.1038/s41586-019-1056-z
- Bi, M., Zhang, Z., Jiang, Y.-Z., Xue, P., Wang, H., Lai, Z., et al. (2020). Enhancer Reprogramming Driven by High-Order Assemblies of Transcription Factors Promotes Phenotypic Plasticity and Breast Cancer Endocrine Resistance. *Nat. Cell Biol.* 22 (6), 701–715. Epub 2020/05/20PubMed PMID: 32424275; PubMed Central PMCID: PMC737911. doi:10.1038/s41556-020-0514-z

AUTHOR CONTRIBUTIONS

MS, DP, and RN equally contributed to this study. Experiments in this study were conceived by MS, DP, RN, and MT. Experiments were performed by MS, DP, RN, and MT. Data analysis was performed by MS, DP, MC, and MT. All authors participated in composing and editing the manuscript.

FUNDING

This project is supported by the National Institutes of Health (P20GM104360, 3P20GM104360-07S1 to MT); start-up funds were provided by the University of North Dakota School of Medicine and Health Sciences, Department of Biomedical Sciences (to MT).

ACKNOWLEDGMENTS

We gratefully acknowledge the UND Genomics core facility for outstanding technical assistance.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.820532/full#supplementary-material>

- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a Flexible Trimmer for Illumina Sequence Data. *Bioinformatics* 30 (15), 2114–2120. Epub 2014/04/04PubMed PMID: 24695404; PubMed Central PMCID: PMC4103590. doi:10.1093/bioinformatics/btu170
- Broad Institute (2019). Picard Toolkit Broad Institute, GitHub Repository. Available at: <https://broadinstitute.github.io/picard/>.
- Bruna, A., Rueda, O. M., Greenwood, W., Batra, A. S., Callari, M., Batra, R. N., et al. (2016). A Biobank of Breast Cancer Explants with Preserved Intra-tumor Heterogeneity to Screen Anticancer Compounds. *Cell* 167 (1), 260–274. e22. Epub 2016/09/20PubMed PMID: 27641504; PubMed Central PMCID: PMC45037319. doi:10.1016/j.cell.2016.08.041
- Chou, J., Provot, S., and Werb, Z. (2010). GATA3 in Development and Cancer Differentiation: Cells GATA Have It! *J. Cell. Physiol.* 222 (1), 42–49. Epub 2009/10/03PubMed PMID: 19798694; PubMed Central PMCID: PMC2915440. doi:10.1002/jcp.21943
- Chou, J., Lin, J. H., Brenot, A., Kim, J.-w., Provot, S., and Werb, Z. (2013). GATA3 Suppresses Metastasis and Modulates the Tumour Microenvironment by Regulating microRNA-29b Expression. *Nat. Cell Biol.* 15 (2), 201–213. Epub 2013/01/29PubMed PMID: 23354167; PubMed Central PMCID: PMC3660859. doi:10.1038/ncb2672
- Ciriello, G., Gatza, M. L., Beck, A. H., Wilkerson, M. D., Rhie, S. K., Pastore, A., et al. (2015). Comprehensive Molecular Portraits of Invasive Lobular Breast Cancer. *Cell* 163 (2), 506–519. Epub 2015/10/10PubMed PMID: 26451490; PubMed Central PMCID: PMC4603750. doi:10.1016/j.cell.2015.09.033
- El-Haibi, C. P., Bell, G. W., Zhang, J., Collmann, A. Y., Wood, D., Scherber, C. M., et al. (2012). Critical Role for Lysyl Oxidase in Mesenchymal Stem Cell-Driven Breast Cancer Malignancy. *Proc. Natl. Acad. Sci.* 109 (43), 17460–17465. Epub 2012/10/04PubMed PMID: 23033492; PubMed Central PMCID: PMC3491529. doi:10.1073/pnas.1206653109
- Emmanuel, N., Lofgren, K. A., Peterson, E. A., Meier, D. R., Jung, E. H., and Kenny, P. A. (2018). Mutant GATA3 Actively Promotes the Growth of Normal and

- Malignant Mammary Cells. *Anticancer Res.* 38 (8), 4435–4441. Epub 2018/08/01PubMed PMID: 30061207; PubMed Central PMCID: PMC6092927. doi:10.21873/anticancer.12745
- Feng, Q., Zhang, Z., Shea, M. J., Creighton, C. J., Coarfa, C., Hilsenbeck, S. G., et al. (2014). An Epigenomic Approach to Therapy for Tamoxifen-Resistant Breast Cancer. *Cell Res* 24 (7), 809–819. Epub 2014/05/31PubMed PMID: 24874954; PubMed Central PMCID: PMC4085766. doi:10.1038/cr.2014.71
- García-Closas, M., Troester, M. A., Qi, Y., Langerod, A., Yeager, M., Lissowska, J., et al. (2007). Common Genetic Variation in GATA-Binding Protein 3 and Differential Susceptibility to Breast Cancer by Estrogen Receptor Tumor Status. *Cancer Epidemiol. Biomarkers Prev.* 16 (11), 2269–2275. Epub 2007/11/17PubMed PMID: 18006915. doi:10.1158/1055-9965.Epi-07-0449
- Griffith, O. L., Spies, N. C., Anurag, M., Griffith, M., Luo, J., Tu, D., et al. (2018). The Prognostic Effects of Somatic Mutations in ER-Positive Breast Cancer. *Nat. Commun.* 9 (1), 3476. Epub 2018/09/06PubMed PMID: 30181556; PubMed Central PMCID: PMC6123466 algorithm. M.J.E. reports ownership in Bioclassifier LLC that licenses PAM50 patents to Nanostring for the Prosigna breast cancer prognostic test. Commercial platforms and algorithms were not used in the analyses reported in this paper. The remaining authors declare no competing interests. doi:10.1038/s41467-018-05914-x
- Gustin, J. P., Miller, J., Farag, M., Rosen, D. M., Thomas, M., Scharpf, R. B., et al. (2017). GATA3 Frameshift Mutation Promotes Tumor Growth in Human Luminal Breast Cancer Cells and Induces Transcriptional Changes Seen in Primary GATA3 Mutant Breast Cancers. *Oncotarget* 8 (61), 103415–103427. Epub 2017/12/22PubMed PMID: 29262572; PubMed Central PMCID: PMC5732738. doi:10.18632/oncotarget.21910
- Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., et al. (2010). Simple Combinations of Lineage-Determining Transcription Factors Prime Cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol. Cell* 38 (4), 576–589. Epub 2010/06/02PubMed PMID: 20513432; PubMed Central PMCID: PMC2898526. doi:10.1016/j.molcel.2010.05.004
- Hoch, R. V., Thompson, D. A., Baker, R. J., and Weigel, R. J. (1999). GATA-3 Is Expressed in Association with Estrogen Receptor in Breast Cancer. *Int. J. Cancer* 84 (2), 122–128. Epub 1999/03/30. PubMed PMID: 10096242. doi:10.1002/(sici)1097-0215(19990420)84:2<122:aid-ijc5>3.0.co;2-s
- Hruschka, N., Kalisz, M., Subijana, M., Graña-Castro, O., Del Cano-Ochoa, F., Brunet, L. P., et al. (2020). The GATA3 X308_Splice Breast Cancer Mutation Is a Hormone Context-dependent Oncogenic Driver. *Oncogene* 39, 5455–5467. Epub 2020/06/27PubMed PMID: 32587399. doi:10.1038/s41388-020-1376-3
- Jensen, T.-K., Kuo, W. P., Stokke, T., and Hovig, E. (2002). Associations between Gene Expressions in Breast Cancer and Patient Survival. *Hum. Genet.* 111 (4–5), 411–420. Epub 2002/10/18PubMed PMID: 12384785. doi:10.1007/s00439-002-0804-5
- Kong, S. L., Li, G., Loh, S. L., Sung, W. K., and Liu, E. T. (2011). Cellular Reprogramming by the Conjoint Action of ERα, FOXA1, and GATA3 to a Ligand-inducible Growth State. *Mol. Syst. Biol.* 7, 526. Epub 2011/09/01PubMed PMID: 21878914; PubMed Central PMCID: PMC3202798. doi:10.1038/msb.2011.59
- Kouros-Mehr, H., Slorach, E. M., Sternlicht, M. D., and Werb, Z. (2006). GATA-3 Maintains the Differentiation of the Luminal Cell Fate in the Mammary Gland. *Cell* 127 (5), 1041–1055. Epub 2006/11/30PubMed PMID: 17129787; PubMed Central PMCID: PMC2646406. doi:10.1016/j.cell.2006.09.048
- Langmead, B., Trapnell, C., Pop, M., and Salzberg, S. L. (2009). Ultrafast and Memory-Efficient Alignment of Short DNA Sequences to the Human Genome. *Genome Biol.* 10 (3), R25. Epub 2009/03/06PubMed PMID: 19261174; PubMed Central PMCID: PMC2690996. doi:10.1186/gb-2009-10-3-r25
- Liao, Y., Smyth, G. K., and Shi, W. (2014). featureCounts: an Efficient General Purpose Program for Assigning Sequence Reads to Genomic Features. *Bioinformatics* 30 (7), 923–930. Epub 2013/11/15PubMed PMID: 24227677. doi:10.1093/bioinformatics/btt656
- Lien, H. C., Hsiao, Y. H., Lin, Y. S., Yao, Y. T., Juan, H. F., Kuo, W. H., et al. (2007). Molecular Signatures of Metaplastic Carcinoma of the Breast by Large-Scale Transcriptional Profiling: Identification of Genes Potentially Related to Epithelial-Mesenchymal Transition. *Oncogene* 26 (57), 7859–7871. Epub 2007/07/03PubMed PMID: 17603561. doi:10.1038/sj.onc.1210593
- Liu, J., Prager-van der Smissen, W. J. C., Look, M. P., Sieuwerts, A. M., Smid, M., Meijer-van Gelder, M. E., et al. (2016). GATA3 mRNA Expression, but Not Mutation, Associates with Longer Progression-free Survival in ER-Positive Breast Cancer Patients Treated with First-Line Tamoxifen for Recurrent Disease. *Cancer Lett.* 376 (1), 104–109. Epub 2016/03/29PubMed PMID: 27018307. doi:10.1016/j.canlet.2016.03.038
- Love, M. I., Huber, W., and Anders, S. (2014). Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2. *Genome Biol.* 15 (12), 550. Epub 2014/12/18PubMed PMID: 25516281; PubMed Central PMCID: PMC4302049. doi:10.1186/s13059-014-0550-8
- Machanic, P., and Bailey, T. L. (2011). MEME-ChIP: Motif Analysis of Large DNA Datasets. *Bioinformatics* 27 (12), 1696–1697. Epub 2011/04/14PubMed PMID: 21486936; PubMed Central PMCID: PMC3106185. doi:10.1093/bioinformatics/btr189
- Mair, B., Konopka, T., Kerzendorfer, C., Sleiman, K., Salic, S., Serra, V., et al. (2016). Gain- and Loss-Of-Function Mutations in the Breast Cancer Gene GATA3 Result in Differential Drug Sensitivity. *Plos Genet.* 12 (9), e1006279. Epub 2016/09/03PubMed PMID: 27588951. doi:10.1371/journal.pgen.1006279
- Mehra, R., Varambally, S., Ding, L., Shen, R., Sabel, M. S., Ghosh, D., et al. (2005). Identification of GATA3 as a Breast Cancer Prognostic Marker by Global Gene Expression Meta-Analysis. *Cancer Res.* 65 (24), 11259–11264. Epub 2005/12/17PubMed PMID: 16357129. doi:10.1158/0008-5472.can-05-2495
- Parsana, P., Amend, S. R., Hernandez, J., Pienta, K. J., and Battle, A. (2017). Identifying Global Expression Patterns and Key Regulators in Epithelial to Mesenchymal Transition through Multi-Study Integration. *BMC cancer* 17 (1), 447. Epub 2017/06/28PubMed PMID: 28651527; PubMed Central PMCID: PMC5485747. doi:10.1186/s12885-017-3413-3
- Pereira, B., Chin, S.-F., Rueda, O. M., Vollan, H.-K. M., Provenzano, E., Bardwell, H. A., et al. (2016). The Somatic Mutation Profiles of 2,433 Breast Cancers Refine Their Genomic and Transcriptomic Landscapes. *Nat. Commun.* 7, 11479. Epub 2016/05/11PubMed PMID: 27161491; PubMed Central PMCID: PMC4866047. doi:10.1038/ncomms11479
- Quinlan, A. R., and Hall, I. M. (2010). BEDTools: a Flexible Suite of Utilities for Comparing Genomic Features. *Bioinformatics* 26 (6), 841–842. Epub 2010/01/30PubMed PMID: 20110278; PubMed Central PMCID: PMC2832824. doi:10.1093/bioinformatics/btq033
- Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor Package for Differential Expression Analysis of Digital Gene Expression Data. *Bioinformatics* 26 (1), 139–140. Epub 2009/11/17PubMed PMID: 19910308; PubMed Central PMCID: PMC2796818. doi:10.1093/bioinformatics/btp616
- Slenter, D. N., Kutmon, M., Hanspers, K., Riutta, A., Windsor, J., Nunes, N., et al. (2018). WikiPathways: a Multifaceted Pathway Database Bridging Metabolomics to Other Omics Research. *Nucleic Acids Res.* 46 (D1), D661–D667. Epub 2017/11/15PubMed PMID: 29136241; PubMed Central PMCID: PMC5753270. doi:10.1093/nar/gkx1064
- Sørle, T., Tibshirani, R., Parker, J., Hastie, T., Marron, J. S., Nobel, A., et al. (2003). Repeated Observation of Breast Tumor Subtypes in Independent Gene Expression Data Sets. *Pnas* 100 (14), 8418–8423. Epub 2003/06/28PubMed PMID: 12829800; PubMed Central PMCID: PMC166244. doi:10.1073/pnas.0932692100
- Takaku, M., Grimm, S. A., and Wade, P. A. (2015). GATA3 in Breast Cancer: Tumor Suppressor or Oncogene? *Gene Expr.* 16 (4), 163–168. Epub 2015/12/08PubMed PMID: 26637396. doi:10.37271/105221615x14399878166113
- Takaku, M., Grimm, S. A., Roberts, J. D., Chrysovergis, K., Bennett, B. D., Myers, P., et al. (2018). GATA3 Zinc finger 2 Mutations Reprogram the Breast Cancer Transcriptional Network. *Nat. Commun.* 9 (1), 1059. Epub 2018/03/15PubMed PMID: 29535312; PubMed Central PMCID: PMC5849768. doi:10.1038/s41467-018-03478-4
- Takaku, M., Grimm, S. A., De Kumar, B., Bennett, B. D., and Wade, P. A. (2020). Cancer-specific Mutation of GATA3 Disrupts the Transcriptional Regulatory Network Governed by Estrogen Receptor Alpha, FOXA1 and GATA3. *Nucleic Acids Res.* 48, 4756–4768. Epub 2020/04/02PubMed PMID: 32232341. doi:10.1093/nar/gkaa179
- The Cancer Genome Atlas Network (2012). Comprehensive Molecular Portraits of Human Breast Tumours. *Nature* 490 (7418), 61–70. Epub 2012/09/25PubMed PMID: 23000897; PubMed Central PMCID: PMC3465532. doi:10.1038/nature11412
- Theodorou, V., Stark, R., Menon, S., and Carroll, J. S. (2013). GATA3 Acts Upstream of FOXA1 in Mediating ESR1 Binding by Shaping Enhancer

- Accessibility. *Genome Res.* 23 (1), 12–22. Epub 2012/11/23PubMed PMID: 23172872; PubMed Central PMCID: PMC3530671. doi:10.1101/gr.139469.112
- Usary, J., Llaca, V., Karaca, G., Presswala, S., Karaca, M., He, X., et al. (2004). Mutation of GATA3 in Human Breast Tumors. *Oncogene* 23 (46), 7669–7678. Epub 2004/09/14PubMed PMID: 15361840. doi:10.1038/sj.onc.1207966
- Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS: A J. Integr. Biol.* 16 (5), 284–287. Epub 2012/03/30PubMed PMID: 22455463; PubMed Central PMCID: PMC3339379. doi:10.1089/omi.2011.0118
- Zheng, W.-p., and Flavell, R. A. (1997). The Transcription Factor GATA-3 Is Necessary and Sufficient for Th2 Cytokine Gene Expression in CD4 T Cells. *Cell* 89 (4), 587–596. Epub 1997/05/16PubMed PMID: 9160750. doi:10.1016/s0092-8674(00)80240-8
- Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Publisher’s Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.
- Copyright © 2022 Saotome, Poduval, Nair, Cooper and Takaku. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*