Journal of Mammary Gland Biology and Neoplasia (2019) 24:271–284
https://doi.org/10.1007/s10911-019-09432-4

GATA3 Truncating Mutations Promote Cistromic Re-Programming In Vitro, but Not Mammary Tumor Formation in Mice

Lisette M. Cornelissen[1] · Roebi de Bruijn[1,2] · Linda Henneman[1,3] · Yongsoo Kim[2,4] · Wilbert Zwart[4,5] · Jos Jonkers[1]

Received: 28 February 2019 / Accepted: 31 May 2019 / Published online: 19 June 2019

Abstract

Heterozygous mutations in the transcription factor GATA3 are identified in 10–15% of all breast cancer cases. Most of these are protein-truncating mutations, concentrated within or downstream of the second GATA-type zinc-finger domain. Here, we investigated the functional consequences of expression of two truncated GATA3 mutants, in vitro in breast cancer cell lines and in vivo in the mouse mammary gland. We found that the truncated GATA3 mutants display altered DNA binding activity caused by preferred tethering through FOXA1. In addition, expression of the truncated GATA3 mutants reduces E-cadherin expression and promotes anchorage-independent growth in vitro. However, we could not identify any effects of truncated GATA3 expression on mammary gland development or mammary tumor formation in mice. Together, our results demonstrate that both truncated GATA3 mutants promote cistromic re-programming of GATA3 in vitro, but these mutants are not sufficient to induce tumor formation in mice.

Keywords GATA3 · Breast cancer · Truncating mutations · Cistromic re-programming

Introduction

Breast cancer is the most common female cancer worldwide and can be categorized into four molecular subtypes: luminal A and B, HER2+ and triple negative breast cancer (TNBC) [1, 2]. The luminal subtypes, characterized by expression of estrogen receptor alpha (ERα), are associated with a more favorable prognosis compared to the HER2+ and TNBC subtypes [1, 3]. Expression of the transcription factor GATA3 is highly correlated with ERα levels in luminal breast cancer [4, 5]. In breast cancer cells, this positive correlation between ERα and GATA3 expression is shown to be caused by a positive cross-regulatory loop between these factors [6]. In addition, GATA3 is able to adapt enhancer accessibility at ERα target sites, thereby facilitating ERα-regulated gene expression [7]. High GATA3 levels are also associated with lower histologic grade and better survival, while low levels correlate with higher histologic grade and poor survival [8, 9]. Similarly, high expression of GATA3 in breast cancer cells is associated with a differentiated and epithelial phenotype, whereas GATA3 loss is associated with the acquisition of a less differentiated and more invasive phenotype [10, 11].

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10911-019-09432-4) contains supplementary material, which is available to authorized users.

✉ Wilbert Zwart
w.zwart@nki.nl

✉ Jos Jonkers
j.jonkers@nki.nl

[1] Division of Molecular Pathology, Oncode Institute, The Netherlands Cancer Institute, Plesmanlaan 121, 1066CX Amsterdam, The Netherlands

[2] Division of Molecular Carcinogenisis, Oncode Institute, The Netherlands Cancer Institute, Plesmanlaan 121, Amsterdam 1066CX, The Netherlands

[3] Mouse Clinic for Cancer and Aging – Transgenic facility, The Netherlands Cancer Institute, Plesmanlaan 121, Amsterdam 1066CX, The Netherlands

[4] Division of Oncogenomics, Oncode Institute, The Netherlands Cancer Institute, Plesmanlaan 121, 1066CX Amsterdam, The Netherlands

[5] Laboratory of Chemical Biology and Institute for Complex Molecular systems, Department of Biomedical Engineering, Eindhoven University of Technology, PO Box 513, Eindhoven, The Netherlands

GATA3 is a member of the GATA-like zinc-finger transcription factor family. These GATA factors contain two zinc-finger DNA-binding domains able to recognize the consensus DNA motif [A/T]GATA[A/G] with high affinity [12, 13]. GATA3 regulates gene expression in cell fate decisions in several hematopoietic lineages, but also during the development of multiple epithelial tissues such as mammary gland, kidney, skin and ear [14–19]. Loss of GATA3 in the mouse mammary gland blocks luminal differentiation, resulting in impaired outgrowth of the mammary ductal-tree [14, 19]. GATA3 haploinsufficiency in humans results in congenital hypothyroidism, sensorineural deafness and renal disease (HDR) syndrome [20–22].

Mutations in *GATA3* are identified in about 10–15% of all breast cancer cases and seem to be enriched in invasive ductal carcinomas, compared to invasive lobular carcinomas [23–25]. These mutations are heterozygous and concentrated within the second GATA-type zinc-finger or further downstream [21, 26]. Most mutations are predicted to result in a truncated protein, whereas a cluster of mutations at the end of the coding region is predicted to result in a longer GATA3 protein due to additional missense amino acids [27]. The breast cancer cell line MCF7 is known to carry a heterozygous D336fs mutation in *GATA3* [26]. CRISPR-mediated correction of this mutation resulted in reduced in vivo growth and smaller tumors in xenografts [28]. Conversely, introducing a mutation in GATA3 wild-type (WT) luminal breast cancer cell lines T47D or CAMA-1, or overexpression of truncated GATA3 mutants in ZR-75-1 cells, resulted in increased tumor growth in vivo [28–30]. These data indicate that GATA3 truncations are sufficient to drive proliferation in xenografts.

In this study, we set out to characterize the functional consequences of expression of two truncated GATA3 mutants that either lack the C-terminus alone, or both the second GATA-type zinc-finger domain and the C-terminus, associated with truncating *GATA3* mutations. We compared genome-wide genomic interactions between these truncated GATA3 mutants and wild-type GATA3 in vitro. In addition, we investigated whether expression of the truncated GATA3 mutants affects mammary gland development or results in mammary tumor formation in mice.

Materials and Methods

Cell Lines and Culturing Conditions

Cells were cultured at 37 °C in a 5% CO_2 incubator. T47D cells were cultured in Dulbecco's modified Eagle's medium (DMEM)/F12-Glutamax (31331–093, ThermoFisher Scientific, Bleiswijk, The Netherlands) containing 10% fetal bovine serum (FBS; F0926, Sigma-Aldrich, Zwijndrecht, The Netherlands), 50 IU/ml penicillin, 50 μg/ml streptomycin (15070–063, ThermoFisher Scientific) and 5 μg/ml insulin (I0516, Sigma-Aldrich). MCF7 and MDA-MB-231 cells were cultured in DMEM/F12-Glutamax containing 10% FBS, 50 IU/ml penicillin and 50 μg/ml streptomycin. HEK293T cells for virus production were cultured in Iscove's modified Dulbecco's medium (I3390, Sigma-Aldrich) containing 10% FBS, 50 IU/ml penicillin and 50 μg/ml streptomycin. Purified primary mouse mammary epithelial cells (MMECs) were isolated as described previously [31], and cultured in DMEM/F12-Glutamax containing 10% FBS, 50 IU/ml penicillin, 50 μg/ml streptomycin, 5 μg/ml insulin, 5 ng/ml EGF (E4127, Sigma-Aldrich) and 5 ng/ml cholera-toxin (Inaba 569B, Gentaur, Kampenhout, Belgium). All cell lines were routinely tested for mycoplasma contamination using the MycoAlert Mycoplasma Detection Kit (LT07–318, Lonza, Breda, The Netherlands).

Generation of Cell Lines Expressing Human GATA3 Mutants

Human *GATA3* sequences were isolated from T47D cDNA using specific PCR primers with XbaI and EcoRI overhangs and a 5' HA-tag using Pwo DNA polymerase (11644947001, Sigma-Aldrich) and subsequently cloned into a zero TOPO blunt vector (450245, ThermoFisher Scientific). Sequence-verified cDNAs were inserted as XbaI-EcoRI fragments into the lentiviral pCDH-CMV-MCS-EF1-puro or pCDH-CMV-MCS-EF1-copGFP (CD510B-1 or CD511B-1 respectively, System Biosciences, CA, USA) vector. Used primers are listed in supplementary Table S1. Lentiviral particles were produced by co-transfection of four plasmids in HEK293T cells as previously described [32]. Conditioned media containing the lentiviral particles were collected 48 h after transfection and used for transduction. Stable cell lines were obtained after puromycin (1.8 μg/ml; P7255, Sigma-Aldrich) selection or FACS-sorting for GFP.

siRNA Transfection

Cells were transfected with RNA oligonucleotides using DharmaFECT 1 (T-2001, GE Dharmacon/Horizon Discovery, Lafayette, CO, USA), according to manufacturer's instructions. ON-TARGETplus human siRNAs for FOXA1 (LQ-010319-00-0002, GE Dharmacon) and a non-targeting control (D-001810-10-05, GE Dharmacon) were used. A final concentration of 25 nM RNA oligonucleotides was used. Cells were harvested 48 h after transfection.

Chromatin Immunoprecipitation (ChIP) Sequencing

ChIP was performed as described previously, with adaptations [33]. In short, cells were crosslinked in solution A (pH 7.4, 50 mM Hepes, 100 mM NaCl, 1 mM EDTA, 0.5 M EGTA) containing 2 mM DSG for 35 min, formaldehyde was added to a final concentration of 1% and incubated for another

10 min. After addition of glycine (final concentration of 125 mM) to quench the crosslinking reaction and washing with PBS, cells were collected. The Bioruptor Pico (Diagenode SA, Seraing, Belgium) was used for sonication. Antibodies used were anti-HA-tag and anti-GATA3 (supplementary Table S2) with 100 μl Protein A magnetic beads (ThermoFisher Scientific).

Immunoprecipitated DNA was processed for library preparation using the KAPA library preparation kit (Part#0801–0303, KAPA Biosystems, Amsterdam, The Netherlands). Sequences were generated by the Illumina HiSeq2500 (using 65 bp reads) and mapped to GRCh37 reference genome using Burrows-Wheeler Aligner (BWA, v0.7.5a) with a mapping quality >20. Peak calling was performed using MACS (v1.4 [34]) and DFilter (v1.5 [35]), where only peaks were considered that were shared by the two peak callers. Genome browser snapshots were generated using IGV (v2.4.9) and heatmaps were generated using SeqMiner (v1.3.4 [36]). Peaks identified in two replicates were used for downstream analysis. Enriched motifs were obtained using the SeqPos motif tool in Galaxy Cistrome [37].

Anchorage-Independent Growth

The basal layer consisted of normal growth medium containing 1.5% low gelling agar (A4018, Sigma-Aldrich). Cells were mixed with normal growth medium containing 0.3% low gelling agar and seeded as the upper layer. After 7 and 14 days, plates were imaged using the GelCount (Oxford optronix, Abingdon, UK) and the number of colonies was quantified using the GelCount Software (v1.1.2.0).

Cell Migration

When cells were grown to confluency, a scratch was created across the center of the well using a 10 μl pipette tip. Images were taken immediately after the scratch (day 0) and at day 2, 6 and 10 after the scratch was made. Using ImageJ Software (v1.28e), the area of the scratch was determined at each time point.

Cell Proliferation Assay

Cells were seeded in 384-wells plates and well confluency was recorded every 4 h for each well using an IncuCyte Zoom Live-Cell Analysis System (Essen Bioscience). The images were analyzed using the IncuCyte Zoom Software (v2015a).

Protein Analysis

Protein lysates were prepared using a RIPA lysis buffer (10 mM Tris-HCl at pH 7.5, 150 mM NaCl, 5 mM EDTA, 0.1% SDS, 1% Triton X-100, 1% deoxycholate in milli-Q) supplemented with protease inhibitors (11836145001, Sigma-Aldrich). Protein lysates were quantified using the BCA Protein Assay Kit (23,225, ThermoFisher Scientific). Equal amounts of protein were separated by 4–12% NuPAGE gradient gel (NP0321, ThermoFisher Scientific) and transferred onto nitrocellulose membrane (162–0112, Bio-Rad, Veenendaal, The Netherlands) in transfer buffer (25 mM Tris, 2 M Glycine, 20% Methanol in demineralized water). Membranes were blocked in 5% non-fat dry milk in TBS-T (pH 7.6, 20 mM Tris, 138 mM NaCl, 0.05% Tween-20 in demineralized water) after which they were probed overnight at 4 °C with primary antibodies as listed in supplementary Table S2 and 1 h with HRP-conjugated secondary antibodies (1:2000; Dako/Agilent, Amstelveen, The Netherlands). Protein was visualized using ECL (32209; ThermoFisher Scientific) on film or using the ChemiDoc MP (Bio-Rad). Protein quantifications were performed using Image Lab Software (v6.0 Bio-Rad).

RNA Isolation and Sequencing

Total RNA was extracted from cells using TRIzol reagent (15596026, ThermoFisher Scientific) according to manufacturer's instructions. Quality and quantity of RNA was assessed by the 2100 Bioanalyzer using a Nano chip (Agilent) and samples having RIN > 8 were subjected to library generation. Library preparation for Illumina sequencing was performed using the KAPA Stranded RNA-seq Library Preparation Kit (KR0934, KAPA Biosystems) according to the manufacturer's instruction. The libraries were analyzed on a 2100 Bioanalyzer using a 7500 chip (Agilent), diluted and pooled equimolar into a 10 nM sequencing stock solution. Libraries were sequenced with 65 base single reads on a HiSeq2500 using V4 chemistry (Illumina Inc., San Diego, CA, USA). The reads were mapped to the GRCh38 reference genome using TopHat (v2.0.12 [38]; Bowtie v1.0.0 [39]; Samtools 0.1.19 [40]) with settings transcriptome-index and prefilter-multihits after filtering out low complexity and repetitive reads. Gene expression counts were generated by Icount, which is based on HTSeq-count [41], using gene definitions from Ensembl GRCh38. Read counts were corrected for differences in sequencing depth by using DESeqs median-of-ratios [42] and the normalized counts were log2-transformed. Unsupervised hierarchical clustering was done on Euclidean distance with complete linkage. DESeq2 was used to perform differential expression and genes with FDR-adjusted p-values of <0.05 were considered to be significant [43].

Mouse Models

Mouse *Gata3* sequences were isolated from *Gata3* cDNA (mmm1013–202859543, GE Dharmacon) using specific

PCR primers with FseI and PmeI overhangs using Pwo DNA polymerase (11644947001, Sigma-Aldrich) and subsequently cloned into a zero TOPO blunt vector (450245, ThermoFisher Scientific). The sequence-verified cDNAs were inserted as FseI-PmeI fragments into the *Frt-inv*CAG-IRES-*Luc* vector [44], resulting in *inv*CAG-mGATA3^{WT}-IRES-*Luc*, *inv*CAG-mGATA3^{TR1}-IRES-*Luc* and *inv*CAG-mGATA3^{TR2}-IRES-*Luc*. Flp-mediated integration of the shuttle vectors in *Col1a1*$^{frt/+}$ ESC clones (FVB) and subsequent blastocyst injections of the modified ESCs were performed as previously described [44]. The resulting chimeric animals were crossed with either *MMTV-cre;mT/mG* (FVB) animals or *Wap-Cre* (FVB) animals to generate the cohorts. The *Col1a1*$^{invCAG\text{-}mGATA3\text{-}WT\text{-}IRES\text{-}Luc}$, *Col1a1*$^{invCAG\text{-}mGATA3\text{-}TR1\text{-}IRES\text{-}Luc}$, *Col1a1*$^{invCAG\text{-}mGATA3\text{-}TR2\text{-}IRES\text{-}Luc}$, *Col1a1*^{WT}, *MMTV-cre*, *mT/mG* and *Wap-Cre* alleles were confirmed by multiplex PCR using MyTaq HS Red Mix (BIO-25048, Bioline, Waddinxveen, The Netherlands) with an annealing temperature of 60 °C, according to manufacturer's instructions. All used primers are listed in supplementary Table S1. In vivo bioluminescence imaging was performed as described previously [45]. Signal intensity was measured over the region of interest and quantified as flux (p/s/cm^2/sr). All mouse experiments were approved by the Animal Ethics Committee of the Netherlands Cancer Institute and performed in accordance with institutional, national and European guidelines for animal care and use.

Histology, Immunohistochemistry and Carmine Stainings

Mouse tissues were formalin-fixed in 10% neutral buffered formalin, embedded in paraffin, sectioned and stained with hematoxylin and eosin (H&E). Immunohistochemical stainings for GFP were processed as described previously [45, 46], using Tris/EDTA buffer-based antigen retrieval, anti-GFP (supplementary Table S2) and HRP anti-rabbit Envision (K4011, Dako/Agilent). All slides were digitally processed using the Aperio ScanScope (Aperio, Vista, CA, USA) and captured using ImageScope software (v12.0.0, Aperio). Whole-mount carmine stainings were performed as described previously [31] on one inguinal mammary gland per mouse. Glands were imaged using a Stereo microscope (Olympus SZX12, Zoeterwoude, The Netherlands). Images were analyzed using ImageJ Software to measure length of the mammary fat pad and the length of ductal invasion.

Statistics

Statistical analyses were performed with GraphPad Prism (v7.03). Statistical tests used were T-test, one-way ANOVA, two-way ANOVA, Log-rank (Mantel-Cox) test and Fisher's exact test. *P*-values of <0.05 were considered to be significant.

All sequencing data generated in this study are available on GEO repository: GSE122849.

Results

Two Truncated GATA3 mutants, Lacking either the C-Terminus or both the GATA-Type Zinc-Finger and the C-Terminus, Are Stably Expressed in T47D Cells

Combining all breast cancer-associated *GATA3* mutations listed in cBio Cancer genomics Portal reveals clustering of truncating mutations in the C-terminus of GATA3 (Fig. 1a) [47, 48]. Even though many distinct mutations have been identified, the majority of these mutations can be divided into two clusters. Cluster 1 contains truncating mutations predicted to affect the C-terminal part of GATA3, to which no functional domains are annotated. Cluster 2 contains protein-truncating mutations predicted to affect the second GATA-type zinc-finger domain (Fig. 1a). To study the functional consequences of disruption of these domains, we generated two HA-tagged truncated GATA3 mutants. One of the GATA3 mutants lacks the C-terminus, encoded by exon 6, resulting in expression of a truncated hGATA3$^{1–350}$ protein, hereafter referred to as HA-hGATA3^{TR1} (Fig. 1b). The second GATA3 mutant lacks both the second GATA-type zinc-finger and C-terminus, encoded by exon 5 and 6, resulting in expression of a truncated hGATA3$^{1–308}$ protein, hereafter referred to as HA-hGATA3^{TR2} (Fig. 1b). In addition, HA-hGATA3^{WT} was generated to use as a wild-type control (Fig. 1b).

These GATA3 variants were expressed in T47D luminal breast cancer cells with endogenous wild-type GATA3 expression (Fig. 1c). Both truncated GATA3 mutants were expressed at similar or higher levels compared to endogenous GATA3, while HA-hGATA3^{WT} expression was modest. Since GATA3 is shown to be a tumor suppressor that inhibits growth, increasing GATA3 levels might be selected against, thereby explaining the relatively low levels of HA-hGATA3^{WT} [10, 11]. In addition, the *GATA3*^{D336fs} mutation in MCF7 cells is shown to be associated with increased stability of the truncated protein, resulting in higher levels compared to WT [49]. This increased stability could explain the higher expression levels of both exogenously introduced truncated GATA3 mutants compared to WT. Together, this shows that both truncated GATA3 mutants, lacking either the C-terminus or both the second GATA-type zinc-finger domain and the C-terminus, can be stably expressed in GATA3 proficient cells.

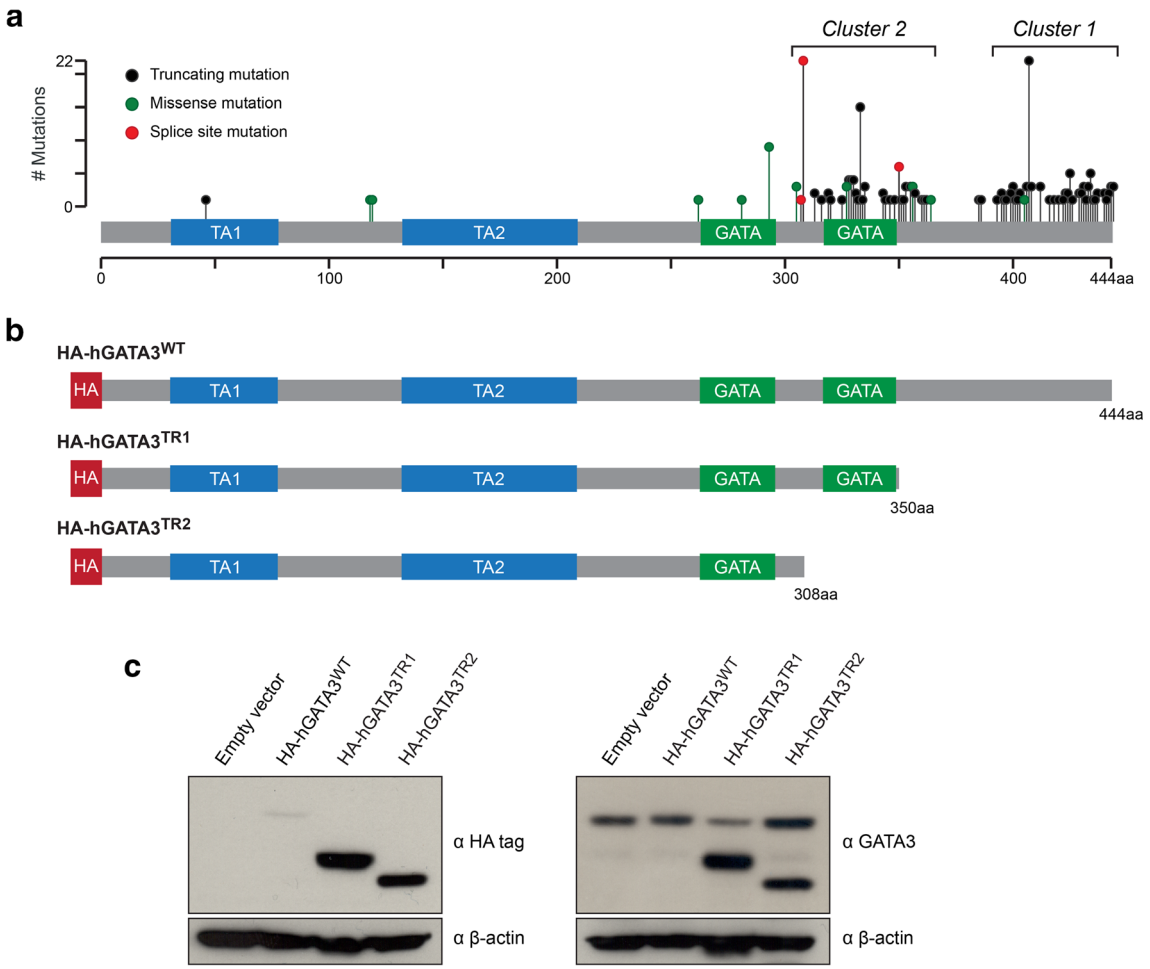
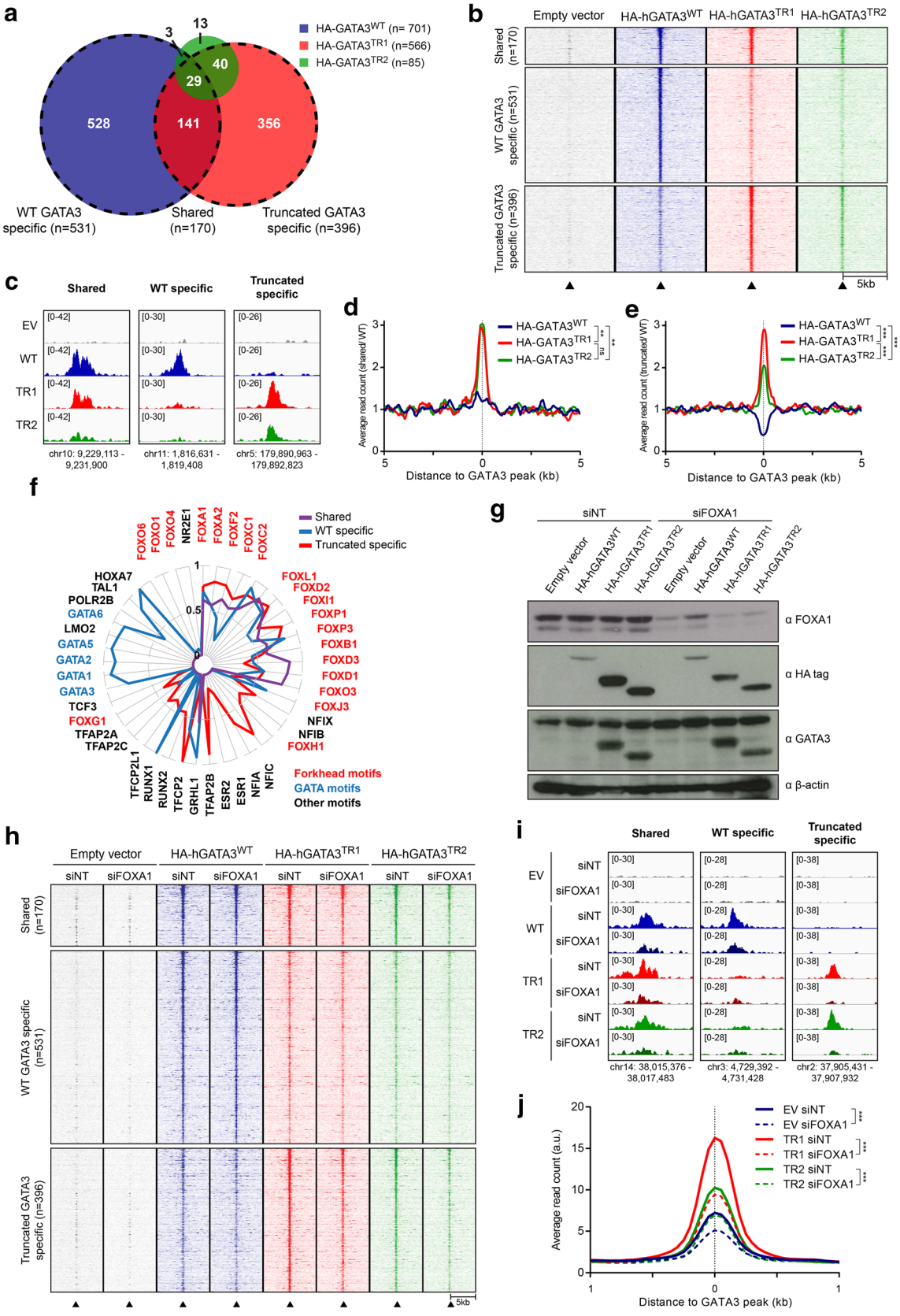


Fig. 1 Stable expression of truncated GATA3 mutants in T47D cells. (**a**) Mutations identified in *GATA3* listed in the cBio Cancer Genomics Portal, clustering primarily in the C-terminus. Cluster 1, amino acids 386 to 444; cluster 2, amino acids 305 to 364. TA1/TA2, transactivating domains 1 and 2; GATA, GATA-type zinc-finger DNA-binding domain. (**b**) Overview of the GATA3 variants used for downstream analysis. HA-hGATA3^{WT}; WT human GATA3 sequence (amino acids 1–444), HA-hGATA3^{TR1}, truncated mutant 1 (amino acids 1–350), HA-hGATA3^{TR2}; truncated mutant 2 (amino acids 1–308). HA, N-terminal fused HA-tag. (**c**) Western blot analysis showing T47D cells transduced with any of the GATA3 constructs with modest expression of HA-hGATA3^{WT} and high expression of either HA-hGATA3^{TR1} or HA-hGATA3^{TR2} compared to endogenous GATA3 levels. β-actin was used as loading control

Truncation of GATA3 Induces Genomic Re-Programming by Tethering through FOXA1 Rather than Interacting Directly with the DNA

HA-hGATA3^{TR2} lacks the second GATA-type zinc-finger that is essential in DNA interactions, while this domain is not affected in HA-hGATA3^{TR1} and HA-hGATA3^{WT} [50]. Therefore, we were interested to determine whether these truncated mutants are able to interact with the DNA and whether the genomic distribution of these mutants is comparable to WT GATA3. To analyze DNA-binding activity of the exogenously introduced GATA3 variants, ChIP-sequencing analyses were performed using an anti-HA antibody. The number of detected binding sites was comparable between HA-hGATA3^{WT} and HA-hGATA3^{TR1}, while for HA-hGATA3^{TR2} less peaks were observed (Fig. 2a). A total of 566 sites were occupied by HA-hGATA3^{TR1}, of which 69 were also identified for HA-hGATA3^{TR2}. Comparing genomic distribution between HA-hGATA3^{WT} and HA-hGATA3^{TR1} (Fig. 2a; dotted lines) revealed that about 25% of sites identified for HA-hGATA3^{WT} were also occupied by HA-hGATA3^{TR1} (shared sites, $n = 170$), while at the other 75% of sites HA-hGATA3^{TR1} binding was decreased compared to HA-hGATA3^{WT} (WT GATA3-specific sites, $n = 531$; Fig. 2b-e). Interestingly, HA-hGATA3^{TR1} was able to interact with an additional 396 sites that were only weakly bound by HA-hGATA3^{WT} (truncated GATA3-specific sites, $n = 396$; Fig. 2b-e), suggesting that loss of the C-terminal domain results in re-targeting of GATA3 to other genomic sites. Even though not all these sites were identified by peak-calling algorithms in HA-hGATA3^{TR2} cells, we did observe an increased raw ChIP-seq signal compared to HA-hGATA3^{WT} (Fig. 2b, c). The difference in DNA-binding intensity between the two mutants might be explained by differences in



expression levels, or less stable DNA binding of HA-hGATA3^{TR2}. The total number of binding sites identified for the GATA3 variants is relatively low compared to total GATA3 ChIP-seq in T47D cells [49] due to the lower

Fig. 2 Cistromic re-programming of truncated GATA3 caused by a shift from direct interaction with the canonical GATA motif to tethering through FOXA1. **a** Venn diagram illustrating the overlap in binding sites identified for HA-hGATA3^{WT} (blue), HA-hGATA3^{TR1} (red) and HA-hGATA3^{TR2} (green) shared in two independent ChIP-seq experiments. Dotted lines represent the comparison between HA-hGATA3^{WT} and HA-hGATA3^{TR1} that was used for further analyses. **b** Heatmap indicating raw binding peak intensity of ChIP signal of the cell lines separated in shared, WT GATA3-specific and truncated GATA3-specific sites. A window of 5 kb around the peak is shown. **c** Genome browser snapshots with examples of peaks that are shared, WT GATA3-specific or truncated GATA3-specific. Genomic locations and read counts are indicated. (**d**, **e**) Average read count profiles of the binding peak signal at (**d**) shared sites and (**e**) truncated GATA3-specific sites normalized to the signal at WT GATA3-specific sites. Average read counts at the center of the peaks were compared, ANOVA: ** $p < 0.01$; *** $p < 0.001$; ns $p > 0.05$ (**f**) Differential enrichment of the fraction of Forkhead and GATA DNA motifs enriched in shared, WT-specific and truncated-specific binding sites visualized by a radar plot. (**g**) Western blot analysis showing an efficient silencing of FOXA1 upon siRNA-mediated knockdown. β-actin was used as loading control. (**h**) Heatmap indicating raw peak intensity of HA-tag ChIP-seq of the different cell lines comparing siNT versus siFOXA1 conditions separated in shared, WT GATA3-specific and truncated GATA3-specific sites. A window of 5 kb around the peak is shown. (**i**) Genome browser snapshots with examples of binding sites that are shared, WT GATA3-specific or truncated GATA3-specific comparing siNT versus siFOXA1 conditions. Genomic locations and read counts are indicated. (**j**) Average read count profiles of the peak signal at truncated GATA3-specific sites. Average read counts at the center of the peaks were compared, ANOVA: *** $p < 0.001$. A.U., arbitrary units

efficiency of the anti-HA antibody. To confirm cistromic re-programming of truncated GATA3 mutants, we compared total GATA3 binding between GATA3-wild-type T47D cells and MCF7 cells carrying a heterozygous *D336fs* mutation (Fig. S1A). Quantification of the GATA3 ChIP-seq signal intensity in T47D cells revealed a lower signal at truncated GATA3-specific sites compared to WT-specific sites, while these signals were comparable in MCF7 cells (Fig. S1B,C). Together, these results indicate that absence of the C-terminal domain results in cistromic re-programming of mutant GATA3 compared to WT GATA3 and that these truncated mutants interact with the DNA, independent of the presence of the second GATA-type zinc-finger.

It has been shown that interaction of WT GATA3 with the DNA is primarily mediated by the second GATA-type zinc-finger through the [A/T]GATA[A/G] motif [13, 50]. To determine how this interaction is affected upon truncation of the protein, we analyzed the enriched DNA motifs for the three different groups of binding sites (Fig. 2f). As expected, HA-hGATA3^{WT}-specific binding sites were enriched for the consensus GATA motif and Forkhead motifs, indicating that WT GATA3 is able to interact directly with the DNA trough [A/T]GATA[A/G] motifs, but also indirectly via binding to FOXA1 [7, 51]. However, the shared sites between WT and truncated GATA3 as well as the truncated GATA3-specific sites showed a clear enrichment for the Forkhead motifs, but no GATA motifs. Upon siRNA-mediated knockdown of FOXA1 expression, DNA binding at the shared and truncated GATA3-specific sites was largely decreased (Fig. 2g-j; Fig. S1D), while binding at WT GATA3-specific sites was only modestly affected (Fig. 2g-i; Fig. S1E). These results indicate that WT GATA3 has the capacity to interact with the DNA directly using GATA motifs and by tethering through FOXA1 using Forkhead motifs, but upon loss of the C-terminus and/or the second GATA-type zinc-finger the direct interaction using GATA motifs is lost and binding is re-programmed to sites with Forkhead motifs.

Truncated GATA3 Expression Does Not Affect Cell Proliferation, Migration and Anchorage-Independent Growth in T47D Cells

Since we observed cistromic re-programming of HA-hGATA3^{TR1} and HA-hGATA3^{TR2}, we were interested to evaluate the phenotypic consequences of expression of these mutants in T47D cells. Comparing cell proliferation between cells expressing HA-hGATA3^{TR1} or HA-hGATA3^{TR2} to WT or endogenous GATA3 did not reveal any significant differences (Fig. S2A). Cell migration, as measured by a scratch assay (Fig. S2B,C), and anchorage-independent growth (Fig. S2D,E) were also not significantly changed upon expression of these truncated GATA3 mutants. RNA-sequencing of these cell lines revealed a separation between cells expressing HA-hGATA3^{WT} and cells expressing either HA-hGATA3^{TR1} or HA-hGATA3^{TR2} in unsupervised hierarchical clustering and principal component analysis (Fig. S3A,B). However, only a small number of differentially expressed genes were identified upon expression of any of the truncated GATA3 mutants (log2FC >1/<−1; Fig. S3C; WT vs EV, 8 up/3 down; TR1 vs EV, 82 up/42 down; TR2 vs EV, 15 up/6 down) and we did not identify enrichment of differentially expressed genes in pathways or gene sets from the BioCarta, KEGG and Reactome pathway databases (data not shown). To evaluate whether differential binding of the truncated GATA3 mutants could explain the alterations in gene expression, each peak was assigned to a gene when the peak was within 20 kb of the transcription start site or within the gene body of that specific gene. Nonetheless, the differential binding events did not induce obvious changes in expression of these genes (Fig. S3D-F). Together, this indicates that expression of either of the two truncated GATA3 mutants does not induce any distinct phenotypic effects or gene expression changes in T47D cells.

Expression of Truncated GATA3 Mutants Reduces E-Cadherin Expression and Promotes Anchorage-Independent Growth in GATA3-Negative MDA-MB-231 Cells

GATA3 is shown to induce epithelial differentiation, while GATA3 loss is linked to mesenchymal transformation [10,

11]. As reported previously, exogenous expression of GATA3 in the GATA3-negative MDA-MB-231 breast cancer cell line results in induction of E-cadherin expression and thereby a more epithelial phenotype and less invasive growth [11, 52, 53]. Introduction of HA-hGATA3^{WT} in MDA-MB-231 cells resulted in increased E-cadherin levels, while no increase was observed upon introduction of HA-hGATA3^{TR1} or HA-hGATA3^{TR2} (Fig. 3a, b). In contrast, both truncated GATA3 mutants were able to further reduce the low levels of endogenous E-cadherin. Both mutants also reduced the up-regulation of E-cadherin in MDA-MB-231 cells transduced with HA-hGATA3^{WT} (Fig. 3a, b). No obvious differences in cell proliferation were observed between the different cell lines (Fig. 3c). Expression of HA-hGATA3^{WT} in MDA-MB-231 cells reduced their capacity to form colonies in soft agar (Fig. 3d, e). Interestingly, HA-hGATA3^{TR1} expression was associated with an increased potential to form colonies (Fig. 3d, e). Moreover, the decreased potential to form colonies in soft agar upon HA-hGATA3^{WT} expression could be reversed significantly by HA-hGATA3^{TR1} expression (Fig. 3d, e). A similar effect was observed upon expression of HA-hGATA3^{TR2}, but this did not reach statistical significance. Collectively, these results indicate that expression of the truncated GATA3 mutants, in contrast to WT GATA3, reduces E-cadherin expression and promotes anchorage-independent growth in the GATA3-negative MDA-MB-231 cell line.

Expression of Truncated GATA3 Does Not Influence Mouse Mammary Gland Development

GATA3 is shown to be important in the development of the mammary gland [14, 19]. Therefore, we set out to test the effect

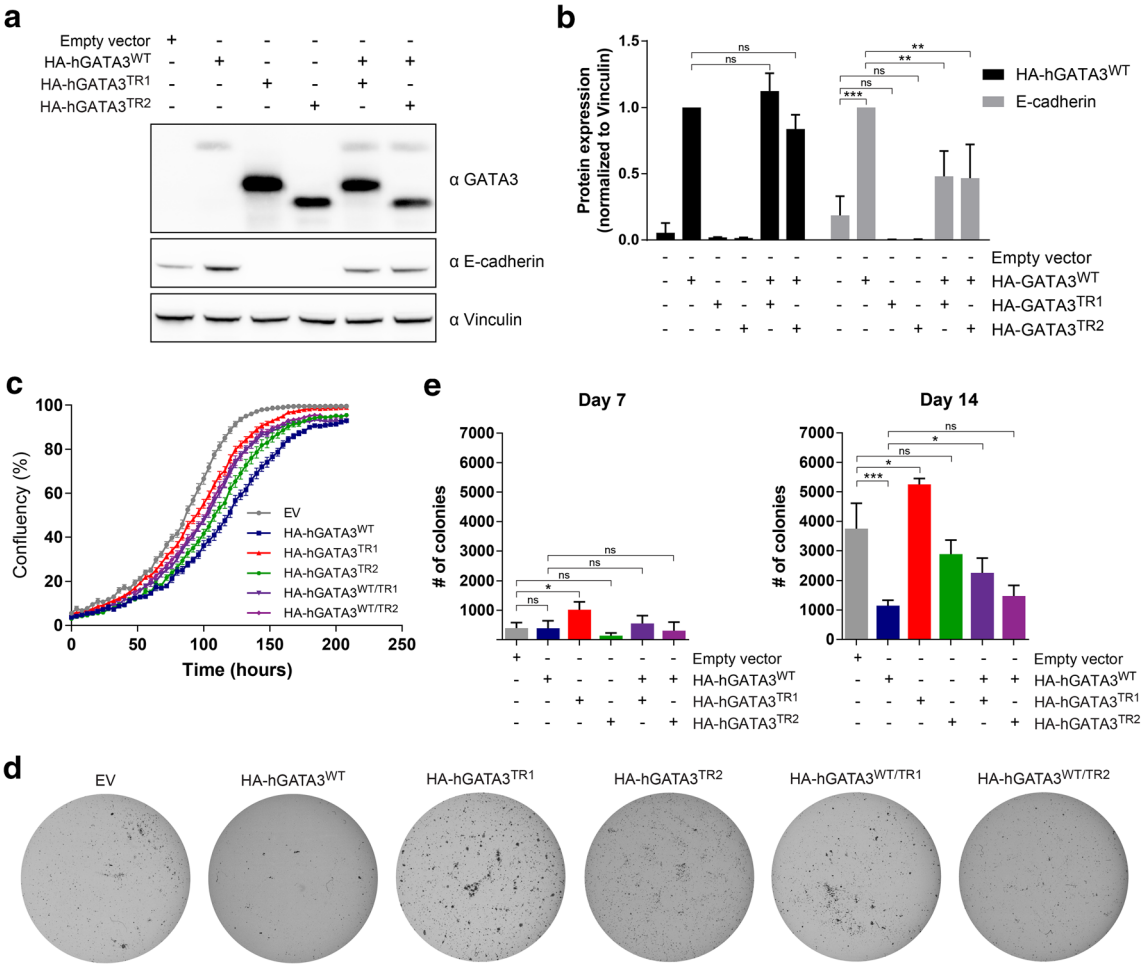


Fig. 3 Truncated GATA3 expression decreases E-cadherin expression and promotes anchorage-independent growth in MDA-MB-231 cells. **a** Western blot analysis of E-cadherin expression in MDA-MB-231 cells transduced with HA-hGATA3^{WT}, HA-hGATA3^{TR1} and/or HA-hGATA3^{TR2}. Vinculin was used as loading control. **b** Quantification of protein expression levels of HA-hGATA3^{WT} and E-cadherin normalized to Vinculin. Data represent mean + SD, $n = 2$. ANOVA: ** $p < 0.01$; *** $p < 0.001$; ns $p > 0.05$. **c** Cell proliferation of MDA-MB-231 cells expressing either WT or truncated GATA3 or one of the combinations remained similar, as quantified using IncuCyte imaging for 208 h. Data represent mean ± SEM. **d** Representative images of soft agar colony formation at day 14 after plating of MDA-MB-231 cells expressing WT or truncated GATA3 or one of the combinations. **e** Quantification of a representative experiment at day 7 and 14 after plating, as measured using the GelCount software. Data represent mean + SD, $n = 3$. ANOVA: * $p < 0.05$; *** $p < 0.001$; ns $p > 0.05$

of expression of GATA3 lacking the C-terminus alone or together with the second GATA-type zinc-finger on mammary gland development by generating mice with mammary gland-specific expression of the GATA3 variants. Even though human and mouse GATA3 show 96% amino-acid sequence identity, mouse *Gata3* cDNA was used to rule out any species specificity. The truncated GATA3 mutant that lacks the C-terminus alone, encoded by exon 6, results in mGATA3$^{1–349}$, hereafter referred to as mGATA3^{TR1}. The second GATA3 mutant that lacks both the second GATA-type zinc-finger and the C-terminus, encoded by exon 5 and 6, results in mGATA3$^{1–307}$ hereafter referred to as mGATA3^{TR2}. We generated transgenic mice with Cre-inducible expression of one of the mouse GATA3 variants in combination with firefly luciferase, by targeting *Frt-inv*CAG-mGATA3^{WT}-IRES-*Luc*, *Frt-inv*CAG-mGATA3^{TR1}-IRES-*Luc* or *Frt-inv*CAG-mGATA3^{TR2}-IRES-*Luc* alleles to the *Col1a1* locus of mouse embryonic stem cells (mESCs) (Fig. 4a) [44]. Through blastocyst injections of these genetically modified mESCs, chimeric mice were generated and crossed with mice carrying the *MMTV-cre* transgene in combination with the *mT/mG* reporter allele [54], to produce *MMTV-cre;mT/mG;Col1a1*$^{Frt\text{-}invCAG\text{-}mGATA3\text{-}WT\text{-}IRES\text{-}Luc}$ (*MMTV-cre;mT/mG;mGATA3*^{WT}), *MMTV-cre;mT/mG;Col1a1*$^{Frt\text{-}invCAG\text{-}mGATA3\text{-}TR1\text{-}IRES\text{-}Luc}$ (*MMTV-cre;mT/mG;mGATA3*^{TR1}) and *MMTV-cre;mT/mG;Col1a1*$^{Frt\text{-}invCAG\text{-}mGATA3\text{-}TR2\text{-}IRES\text{-}Luc}$ (*MMTV-cre;mT/mG;mGATA3*^{TR2}) mice. In these mice, mammary gland-specific expression of Cre results in co-expression of one of the GATA3 variants with luciferase, in addition to recombination of the *mT/mG* allele resulting in loss of tdTomato (mT) and induction of GFP (mG) expression.

To confirm that Cre expression results in expression of the GATA3 variants, primary mouse mammary epithelial cells (MMECs) were derived from these models and in vitro transduced with adenovirus encoding Cre (AdCre). While we could not distinguish between endogenous GATA3 and mGATA3^{WT} expression, mGATA3^{TR1} and mGATA3^{TR2} were expressed at higher levels compared to endogenous GATA3 (Fig. 4b). Analysis of carmine-stained mammary glands from *MMTV-cre;mT/mG;mGATA3*^{WT}, *MMTV-cre;mT/mG;mGATA3*^{TR1} and *MMTV-cre;mT/mG;mGATA3*^{TR2} mice did not reveal any obvious differences in ductal outgrowth. (Fig. 4c, d; Fig. S4A-F). As demonstrated by the high percentage of GFP-positive cells, Cre-driven recombination occurred in most epithelial cells that form the ductal-tree, indicating that the specific GATA3 variants were expressed in these cells. These data illustrate that expression of the truncated GATA3 mutants does not interfere with mammary gland development.

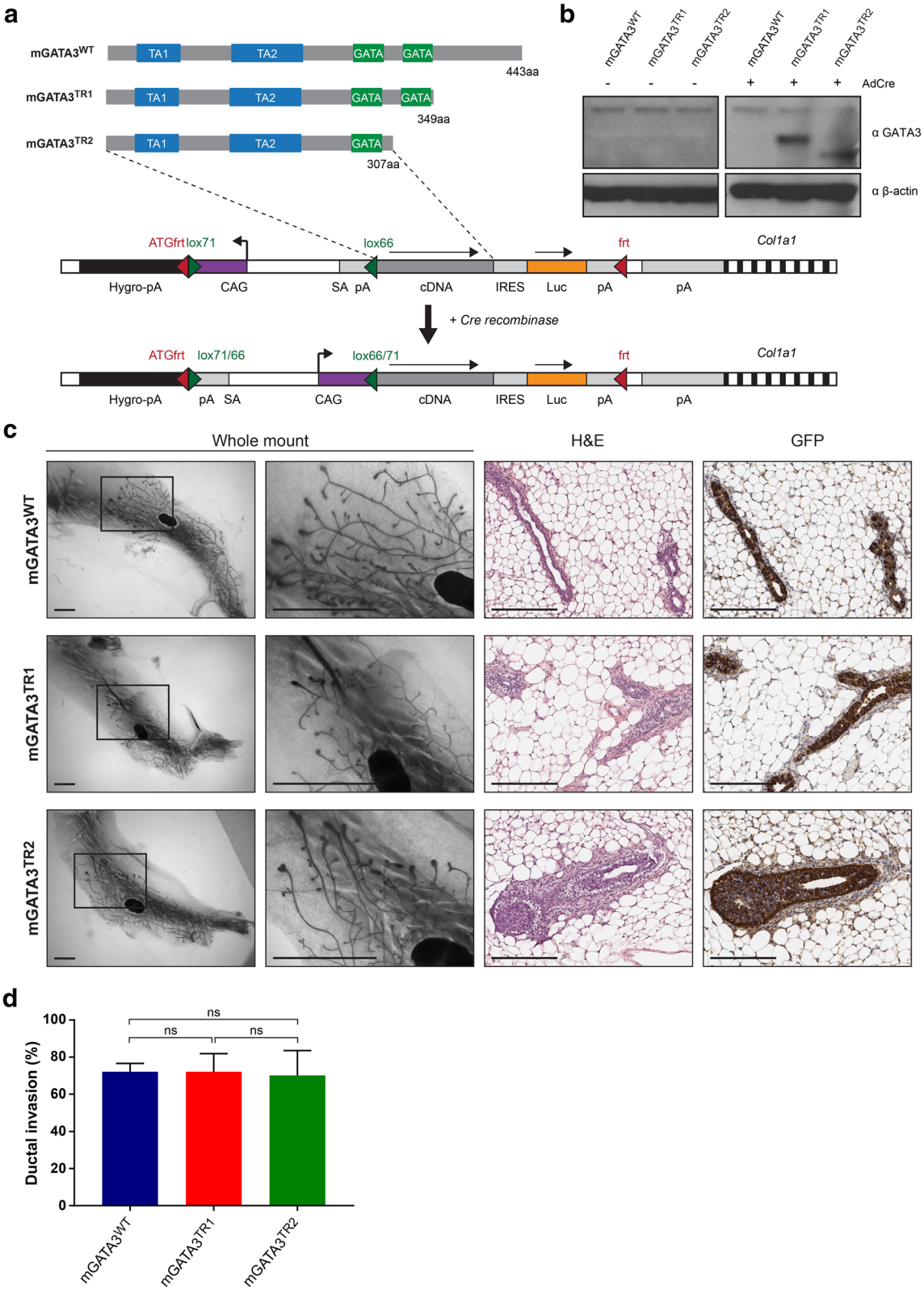
Expression of Truncated GATA3 Does Not Promote Mammary Tumor Formation in Mice

To determine whether expression of any of the two truncated GATA3 mutants results in mammary tumor formation, the *mGATA3*^{TR1} and *mGATA3*^{TR2} alleles were crossed with mammary gland-specific Cre recombinase expressing *Wap-Cre* mice, resulting in *Wap-Cre;mGATA3*^{TR1} and *Wap-Cre;mGATA3*^{TR2} mice. Even though the mammary-specific bioluminescence signals (indicative of expression of the GATA3 mutants) remained high in aged mice, we did not observe any palpable mammary tumors (Fig. 5a, b). Furthermore, histology did not reveal any early-stage mammary hyperplasias or tumors. We observed dilated mammary ducts in most glands, as expected in aged mice (Fig. S5A,B) [55]. The median overall survival of *Wap-Cre;mGATA3*^{TR2} mice was decreased compared to *Wap-Cre;mGATA3*^{TR1} mice (Fig. S5C; Mantel-Cox, 427 days versus >565 days, $p = 0.0007$). However, the distribution between animals that were sacrificed due to illness and animals that reached the endpoint was not significantly different between the two models (Fig. S5D, Fisher's exact, $p = 0.2701$). Together, these data indicate that mammary gland-specific expression of the truncated GATA3 mutants in mice does not induce mammary tumor formation.

Discussion

In this study we set out to determine the functional consequences of expression of two truncated GATA3 mutants that lack the C-terminus by itself or together with the second GATA-type zinc-finger. We demonstrated that these truncated GATA3 mutants lose the capacity to associate with the DNA through canonical GATA motifs, whilst retaining the capacity to tether through FOXA1, resulting in cistromic re-programming. Furthermore, we showed that the truncated GATA3 mutants, in contrast to WT GATA3, reduce E-cadherin levels and promote anchorage-independent growth. However, in spite of their activity in breast cancer cell lines, the truncated GATA3 mutants did not exert any effects on mammary gland development or mammary tumor formation in mice.

Interestingly, both truncated GATA3 mutants were able to interact with the DNA, even though HA-hGATA3^{TR2} lacks the second GATA-type zinc-finger. The second GATA-type zinc-finger domain of GATA3 is most important for DNA recognition using the consensus [A/T]GATA[A/G] motif [13, 50]. Even though the intact N-terminal zinc-finger is shown to recognize the AGAT[C/G/T] motif and directs DNA binding in the presence of the *GATA3*^{R330fs} mutation [30, 56], we did not observe enrichment for this motif for both mutants. Herein we have shown that HA-hGATA3^{TR1} is not able to bind the DNA through the consensus [A/T]GATA[A/G] motif, even though the annotated second GATA-type zinc-finger domain remains intact. In addition, *GATA3* mutations are not only identified within the second GATA-type zinc-finger



domain, but also extend beyond this domain (Fig. 1a; amino acids 351–364). Together, this suggests that additional amino acids downstream of the annotated zinc-finger domain are required for functioning of this domain.

Since truncating *GATA3* mutations are often heterozygous, functional consequences of these mutations might be attributed to activity gained by the mutant allele and/or dosage reduction of the wild-type allele [21, 26]. We have shown that

Fig. 4 Expression of truncated GATA3 does not affect mammary gland development in mice. **a** Schematic overview of the Cre-conditional *Frt-inv*CAG-mGATA3^{WT}-IRES-*Luc* (mGATA3^{WT}), *Frt-inv*CAG-mGATA3^{TR1}-IRES-*Luc* (mGATA3^{TR1}) and *Frt-inv*CAG-mGATA3^{TR2}-IRES-*Luc* (mGATA3^{TR2}) knock-in alleles at the *Col1a1* locus. Expression of Cre recombinase will result in inversion of the CAG promoter, thereby driving expression of the GATA3 variants in combination with luciferase expression. (**b**) Primary MMECs derived from the different models show truncated GATA3 expression upon recombination using Cre adenovirus (AdCre). β-actin was used as loading control. (**c**) Outgrowth of the ductal-tree was determined for *MMTV-cre;mT/mG;mGATA3^{WT}* (mGATA3^{WT}), *MMTV-cre;mT/mG;mGATA3^{TR1}* (mGATA3^{TR1}) and *MMTV-cre;mT/mG;mGATA3^{TR2}* (mGATA3^{TR2}) mice at six weeks of age. Representative images of carmine-stained mammary glands are shown (left; scale bars, 2.5 mm). Representative microscopic images of H&E staining and for GFP expression by immunohistochemistry are shown (right; scale bars, 200 μm). (**d**) Quantification of ductal invasion in mGATA3^{WT} ($n = 5$), mGATA3^{TR1} (n = 3) and mGATA3^{TR2} ($n = 4$) mice (one inguinal gland per mouse) at six weeks of age. Ductal invasion was determined as the length of the mammary fat pad invaded by the ductal structure as a percentage of the total length of the mammary fat pad, measured using ImageJ. Data represent mean + SD, no significant differences (ANOVA)

exogenous expression of GATA3 that lacks the C-terminus alone or in combination with the GATA-type zinc-finger is sufficient to promote anchorage-independent growth of GATA3-negative MDA-MB-231 cells, indicating that this is solely induced by gain-of-function activity of the mutant allele. Even though expression of the GATA3 mutants in GATA3-positive T47D cells does not induce distinct phenotypic effects or gene expression changes, expression of both GATA3 mutants does reduce the effect of wild-type GATA3 expression on E-cadherin levels and anchorage-independent growth potential in MDA-MB-231 cells. Since GATA3 expression is shown to be important in maintaining an epithelial differentiation state [10, 11], the level of wild-type GATA3 expression is relevant. While expression levels of the exogenously introduced truncated GATA3 mutants in T47D are comparable to endogenous wild-type GATA3 levels, expression levels of the two GATA3 mutants in wild-type GATA3-expressing MDA-MB-231 cells are much higher compared to wild-type GATA3. Together, this indicates that expression of either one of the truncated GATA3 mutants we investigated herein has oncogenic potential, but the associated dose-reduction of wild-type protein allows for more pronounced effects.

GATA3 expression is essential in the mouse mammary gland during the formation of the ductal-tree and for maintenance of the luminal epithelium in the adult mammary gland [14, 19]. In contrast to the pronounced phenotypes observed upon loss of GATA3, expression of mGATA3^{TR1} or mGATA3^{TR2} did not affect mammary gland development in our experiments. Our data are in contrast to a recent study from Emmanuel et al. [29], who showed that transgenic expression of truncated human GATA3 in *MMTV-GATA3^{335fs}* mice induces a precocious lobuloalveolar phenotype in the developing mammary gland. The differences between this study and our findings might be due to the use of (i) human versus mouse GATA3, (ii) the GATA3^{335fs} mutation versus mGATA3^{TR1} (GATA3$^{1–349}$) and mGATA3^{TR2} (GATA3$^{1–307}$) or (iii) MMTV-driven expression of truncated GATA3 in *MMTV-GATA3^{335fs}* mice versus (Cre-conditional) CAG-driven expression in our mouse models. The oncogenic potential of truncating GATA3 mutations has been implicated by several in vivo-xenograft studies, reporting a correlation between expression of mutated GATA3 and increased proliferation [28–30]. In contrast, aging of *MMTV-GATA3^{335fs}* mice (Emmanuel et al., [29]) or *Wap-Cre;mGATA3^{TR1}* and *Wap-Cre;mGATA3^{TR2}* mice (this study) did not induce a proliferative phenotype. Together these results indicate that current approaches to model breast cancer-associated *GATA3* mutations in mice do not recapitulate the effects observed in in vitro cell line or in vivo xenograft studies.

Limitations of current in vivo approaches to study the effects of *GATA3* mutations in genetically engineered mouse models (GEMMs) might include the introduction of truncated GATA3 expression without affecting endogenous GATA3 levels, resulting in insufficient levels of the truncated protein to induce measurable effects. This limitation could be

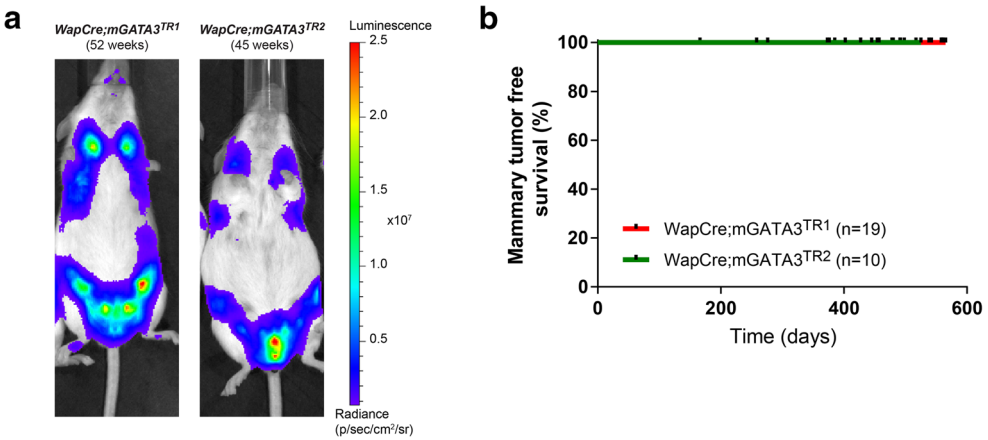


Fig. 5 Expression of truncated GATA3 does not promote mammary tumor formation in mice. **a** Representative images of in vivo bioluminescence imaging of luciferase expression in *Wap-Cre;mGATA3^{TR1}* and *Wap-Cre;mGATA3^{TR2}* mice of approximately one year of age. **b** Kaplan-Meier analysis of mammary tumor free survival of *Wap-Cre;mGATA3^{TR1}* ($n = 19$) and *Wap-Cre;mGATA3^{TR2}* ($n = 10$) mice

addressed by crossing *Wap-Cre;mGATA3^{TR1}* and *Wap-Cre;mGATA3^{TR2}* mice with heterozygous *Gata3$^{+/flox}$* mice [14, 19], resulting in mammary gland-specific expression of truncated GATA3 in combination with half-dose levels of endogenous GATA3. Another possibility might be that truncated GATA3 does not contribute to tumor initiation, but may be beneficial during tumor progression. In support of this, GATA 3 loss is not tolerated during early stages of tumor development, whereas it has been linked to tumor progression [57]. Combining our mouse models with existing mammary tumor models carrying breast cancer driver genes that are frequently co-mutated with GATA3 in human breast cancers might provide insight into whether truncated GATA3 promotes tumor progression. A final limitation might be that *GATA3* mutations are identified primarily in ERα-positive luminal breast cancers, indicating that ERα-signaling is essential for the tumorigenic potential of truncated GATA3 [23, 26, 58]. Even though GEMMs exist that resemble luminal breast cancer, the majority of GEMMs develop ERα-negative tumors [59–61]. Reasons for selective outgrowth of ERα-negative tumors in mice remain elusive, but indicate that mice may not be the most suitable organism to determine the role of *GATA3* mutations in ERα-positive breast cancer. Therefore, there is a need for more refined breast cancer models using different organisms to address the functional consequences of *GATA3* mutations in vivo in ERα-positive breast cancer.

In conclusion, we have shown that two truncated GATA3 mutants, lacking either the C-terminus alone or in combination with the GATA-type zinc-finger, show differential DNA binding compared to WT GATA3. Even though expression of these truncated GATA3 mutants decreases E-cadherin levels and promotes anchorage-independent growth in vitro, expression of the mutants in the mouse mammary gland does not influence development of the gland or induces tumor formation. Reducing endogenous GATA3 expression, introducing additional cancer drivers or using a more suitable model organism might give more insight in the functional consequences of truncated GATA3 expression in vivo.

Acknowledgements We are grateful to Anne Paulien Drenth, Eline van der Burg and Eva Schut for their technical support with the animal studies. We thank the Netherlands Cancer Institute Genomics Core Facility, Mouse Clinic for Cancer and Aging, Animal Facility, and Animal Pathology Facility for their expert technical support.

GEO accession number GSE122849.

Statement of Author Contributions LMC, WZ and JJ designed research. LMC and LH performed research. LMC, RB, YK, WZ and JJ analyzed data. LMC, WZ and JJ wrote the paper.

Funding Sources This work was financially supported by the Oncode Institute; the Center for Translational Molecular Medicine (CTMM) Breast Care Project; the Netherlands Organization for Scientific Research (NWO: Cancer Genomics Netherlands (CGCNL), Cancer Systems Biology Center (CSBC), Zenith 93512009, Vici 91814643); the EU Seventh Framework Program (EurocanPlatform project 260791); the European Research Council (ERC Synergy project CombatCancer); and a National Roadmap grant for Large-Scale Research facilities from NWO.

Compliance with Ethical Standards

Conflict of Interest The authors declare that there are no conflicts of interest

References

1. Perou CM, Sørlie T, Eisen MB, van de RM, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. Nature. 2000;406:747–52.
2. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. CA Cancer J Clin. 2015;65:87–108.
3. Sørlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A. 2001;98:10869–74.
4. Hoch RV, Thompson DA, Baker RJ, Weigel RJ. GATA-3 is expressed in association with estrogen receptor in breast cancer. Int J Cancer. 1999;84:122–8.
5. Voduc D, Cheang M, Nielsen T. GATA-3 expression in breast Cancer has a strong association with estrogen receptor but lacks independent prognostic value. Cancer Epidemiol Prev Biomark. 2008;17:365–73.
6. Eeckhoute J, Keeton EK, Lupien M, Krum SA, Carroll JS, Brown M. Positive cross-regulatory loop ties GATA-3 to estrogen receptor α expression in breast Cancer. Cancer Res. 2007;67:6477–83.
7. Theodorou V, Stark R, Menon S, Carroll JS. GATA3 acts upstream of FOXA1 in mediating ESR1 binding by shaping enhancer accessibility. Genome Res. 2013;23:12–22.
8. Mehra R, Varambally S, Ding L, Shen R, Sabel MS, Ghosh D, et al. Identification of GATA3 as a breast Cancer prognostic marker by global gene expression meta-analysis. Cancer Res. 2005;65:11259–64.
9. Yoon NK, Maresh EL, Shen D, Elshimali Y, Apple S, Horvath S, et al. Higher levels of GATA3 predict better survival in women with breast cancer. Hum Pathol. 2010;41:1794–801.
10. Chou J, Lin JH, Brenot A, Kim J, Provot S, Werb Z. GATA3 suppresses metastasis and modulates the tumour microenvironment by regulating microRNA-29b expression. Nat Cell Biol. 2013;15:201–13.
11. Yan W, Cao QJ, Arenas RB, Bentley B, Shao R. GATA3 inhibits breast Cancer metastasis through the reversal of epithelial-mesenchymal transition. J Biol Chem. 2010;285:14042–51.
12. Lowry JA, Atchley WR. Molecular evolution of the GATA family of transcription factors: conservation within the DNA-binding domain. J Mol Evol. 2000;50:103–15.
13. Merika M, Orkin SH. DNA-binding specificity of GATA family transcription factors. Mol Cell Biol. 1993;13:3999–4010.
14. Asselin-Labat M-L, Sutherland KD, Barker H, Thomas R, Shackleton M, Forrest NC, et al. Gata-3 is an essential regulator of mammary-gland morphogenesis and luminal-cell differentiation. Nat Cell Biol. 2007;9:201–9.
15. Duncan JS, Lim KC, Engel JD, Fritzsch B. Limited inner ear morphogenesis and neurosensory development are possible in the absence of GATA3. Int J Dev Biol. 2011;55:297–303.
16. Grote D, Boualia SK, Souabni A, Merkel C, Chi X, Costantini F, et al. Gata3 acts downstream of β-catenin signaling to prevent

ectopic Metanephric kidney induction. PLoS Genet. 2008;4: e1000316.
17. Ho IC, Tai TS, Pai SY. GATA3 and the T-cell lineage: essential functions before and after T-helper-2-cell differentiation. Nat Rev Immunol. 2009;9:125–35.
18. Kaufman CK, Zhou P, Amalia Pasolli H, Rendl M, Bolotin D, Lim K-C, et al. GATA-3: an unexpected regulator of cell lineage determination in skin. Genes Dev. 2003;17:2108–22.
19. Kouros-Mehr H, Slorach EM, Sternlicht MD, Werb Z. GATA-3 maintains the differentiation of the luminal cell fate in the mammary gland. Cell. 2006;127:1041–55.
20. Ali A, Christie PT, Grigorieva IV, Harding B, Van Esch H, Ahmed SF, et al. Functional characterization of GATA3 mutations causing the hypoparathyroidism-deafness-renal (HDR) dysplasia syndrome: insight into mechanisms of DNA binding by the GATA3 transcription factor. Hum Mol Genet. 2006;16:265–75.
21. Gaynor KU, Grigorieva IV, Allen MD, Esapa CT, Head RA, Gopinath P, et al. GATA3 mutations found in breast cancers may be associated with aberrant nuclear localization, reduced transactivation and cell invasiveness. Horm Cancer. 2013;4:123–39.
22. Van Esch H, Groenen P, Nesbit MA, Schuffenhauer S, Lichtner P, Vanderlinden G, et al. GATA3 haplo-insufficiency causes human HDR syndrome. Nature. 2000;406:419–22.
23. Ciriello G, Gatza ML, Beck AH, Wilkerson MD, Rhie SK, Pastore A, et al. Comprehensive molecular portraits of invasive lobular breast Cancer. Cell. 2015;163:506–19.
24. Pereira B, Chin SF, Rueda OM, Vollan H-KM, Provenzano E, Bardwell HA, et al. The somatic mutation profiles of 2,433 breast cancers refine their genomic and transcriptomic landscapes. Nat Commun. 2016;7:11479.
25. Stephens PJ, Tarpey PS, Davies H, Loo PV, Greenman C, Wedge DC, et al. The landscape of cancer genes and mutational processes in breast cancer. Nature. 2012;486:400–4.
26. Usary J, Llaca V, Karaca G, Presswala S, Karaca M, He X, et al. Mutation of GATA3 in human breast tumors. Oncogene. 2004;23: 7669–78.
27. Mair B, Konopka T, Kerzendorfer C, Sleiman K, Salic S, Serra V, et al. Gain- and loss-of-function mutations in the breast Cancer gene GATA3 result in differential drug sensitivity. PLoS Genet. 2016;12: e1006279.
28. Gustin JP, Miller J, Farag M, Rosen DM, Thomas M, Scharpf RB, et al. GATA3 frameshift mutation promotes tumor growth in human luminal breast cancer cells and induces transcriptional changes seen in primary GATA3 mutant breast cancers. Oncotarget. 2017;8: 103415–27.
29. Emmanuel N, Lofgren KA, Peterson EA, Meier DR, Jung EH, Kenny PA. Mutant GATA3 actively promotes the growth of Normal and malignant mammary cells. Anticancer Res. 2018;38: 4435–41.
30. Takaku M, Grimm SA, Roberts JD, Chrysovergis K, Bennett BD, Myers P, et al. GATA3 zinc finger 2 mutations reprogram the breast cancer transcriptional network. Nat Commun. 2018;9:1059.
31. Boelens MC, Nethe M, Klarenbeek S, de Ruiter JR, Schut E, Bonzanni N, et al. PTEN loss in E-cadherin-deficient mouse mammary epithelial cells rescues apoptosis and results in development of classical invasive lobular carcinoma. Cell Rep. 2016;16:2087–101.
32. Follenzi A, Ailles LE, Bakovic S, Geuna M, Naldini L. Gene transfer by lentiviral vectors is limited by nuclear translocation and rescued by HIV-1 *pol* sequences. Nat Genet. 2000;25:217–22.
33. Schmidt D, Wilson MD, Spyrou C, Brown GD, Hadfield J, Odom DT. ChIP-seq: using high-throughput sequencing to discover protein-DNA interactions. Methods San Diego Calif. 2009;48: 240–8.
34. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 2008;9:R137.
35. Kumar V, Muratani M, Rayan NA, Kraus P, Lufkin T, Ng HH, et al. Uniform, optimal signal processing of mapped deep-sequencing data. Nat Biotechnol. 2013;31:615–22.
36. Ye T, Krebs AR, Choukrallah M-A, Keime C, Plewniak F, Davidson I, et al. seqMINER: an integrated ChIP-seq data interpretation platform. Nucleic Acids Res. 2011;39:e35.
37. Liu T, Ortiz JA, Taing L, Meyer CA, Lee B, Zhang Y, et al. Cistrome: an integrative platform for transcriptional regulation studies. Genome Biol. 2011;12:R83.
38. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. Bioinforma Oxf Engl. 2009;25:1105–11.
39. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009;10:R25.
40. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. Bioinforma Oxf Engl. 2009;25:2078–9.
41. Anders S, Pyl PT, Huber W. HTSeq–a Python framework to work with high-throughput sequencing data. Bioinforma Oxf Engl. 2015;31:166–9.
42. Anders S, Huber W. Differential expression analysis for sequence count data. Genome Biol. 2010;11:R106.
43. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15:550.
44. Huijbers IJ, Del Bravo J, Bin Ali R, Pritchard C, Braumuller TM, van Miltenburg MH, et al. Using the GEMM-ESC strategy to study gene function in mouse models. Nat Protoc. 2015;10:1755–85.
45. Henneman L, van MMH, Michalak EM, Braumuller TM, Jaspers JE, Drenth AP, et al. Selective resistance to the PARP inhibitor olaparib in a mouse model for BRCA1-deficient metaplastic breast cancer. Proc Natl Acad Sci U S A. 2015;112:8409–14.
46. Doornebal CW, Klarenbeek S, Braumuller TM, Klijn CN, Ciampricotti M, Hau C-S, et al. A preclinical mouse model of invasive lobular breast Cancer metastasis. Cancer Res. 2013;73: 353–63.
47. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio Cancer genomics portal: an open platform for exploring multidimensional Cancer genomics data. Cancer Discov. 2012;2:401–4.
48. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal. 2013;6:pl1.
49. Adomas AB, Grimm SA, Malone C, Takaku M, Sims JK, Wade PA. Breast tumor specific mutation in GATA3 affects physiological mechanisms regulating transcription factor turnover. BMC Cancer. 2014;14:278.
50. Yang Z, Gu L, Romeo PH, Bories D, Motohashi H, Yamamoto M, et al. Human GATA-3 trans-activation, DNA-binding, and nuclear localization activities are organized into distinct structural domains. Mol Cell Biol. 1994;14:2201–12.
51. Kong SL, Li G, Loh SL, Sung W-K, Liu ET. Cellular reprogramming by the conjoint action of ERα, FOXA1, and GATA3 to a ligand-inducible growth state. Mol Syst Biol. 2011;7:526.
52. Graff JR, Herman JG, Lapidus RG, Chopra H, Xu R, Jarrard DF, et al. E-cadherin expression is silenced by DNA Hypermethylation in human breast and prostate carcinomas. Cancer Res. 1995;55: 5195–9.
53. Pan Y, Li J, Zhang Y, Wang N, Liang H, Liu Y, et al. Slug-upregulated miR-221 promotes breast cancer progression through suppressing E-cadherin expression. Sci Rep. 2016;6:25798.

54. Muzumdar MD, Tasic B, Miyamichi K, Li L, Luo L. A global double-fluorescent Cre reporter mouse. Genesis. 2007;45:593–605.
55. Raafat A, Strizzi L, Lashin K, Ginsburg E, McCurdy D, Salomon D, et al. Effects of age and parity on mammary gland lesions and progenitor cells in the FVB/N-RC mice. PLoS One. 2012;7: e43624.
56. Pedone PV, Omichinski JG, Nony P, Trainor C, Gronenborn AM, Clore GM, et al. The N-terminal fingers of chicken GATA-2 and GATA–3 are independent sequence-specific DNA binding domains. EMBO J. 1997;16:2874–82.
57. Kouros-Mehr H, Bechis SK, Slorach EM, Littlepage LE, Egeblad M, Ewald AJ, et al. GATA-3 links tumor differentiation and dissemination in a luminal breast Cancer model. Cancer Cell. 2008;13: 141–52.
58. Arnold JM, Choong DYH, Thompson ER, Waddell N, Lindeman GJ, Visvader JE, et al. Frequent somatic mutations of GATA3 in non-BRCA1/BRCA2 familial breast tumors, but not in BRCA1-, BRCA2- or sporadic breast tumors. Breast Cancer Res Treat. 2009;119:491.
59. Dabydeen SA, Furth PA. Genetically engineered ER - positive breast cancer mouse models. Endocr Relat Cancer. 2014;21:R195–208.
60. Herschkowitz JI, Simin K, Weigman VJ, Mikaelian I, Usary J, Hu Z, et al. Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. Genome Biol. 2007;8:R76.
61. Özdemir BC, Sflomos G, Brisken C. The challenges of modeling hormone receptor-positive breast cancer in mice. Endocr Relat Cancer. 2018;25:R319–30.