Gene Expression, Vol. 16, pp. 163–168
Printed in the USA.

1052-2166/15 $90.00 + .00
DOI: http://dx.doi.org/10.3727/105221615X14399878166113
E-ISSN 1555-3884
www.cognizantcommunication.com

Review

GATA3 in Breast Cancer: Tumor Suppressor or Oncogene?

Motoki Takaku,* Sara A. Grimm,† and Paul A. Wade*

*Epigenetics and Stem Cell Biology Laboratory, National Institute of Environmental Health Science, Research Triangle Park, NC, USA
†Integrated Bioinformatics, National Institute of Environmental Health Science, Research Triangle Park, NC, USA

GATA3 is a highly conserved, essential transcription factor expressed in a number of tissues, including the mammary gland. GATA3 expression is required for normal development of the mammary gland where it is estimated to be the most abundant transcription factor in luminal epithelial cells. In breast cancer, GATA3 expression is highly correlated with the luminal transcriptional program. Recent genomic analysis of human breast cancers has revealed high-frequency mutation in GATA3 in luminal tumors, suggesting "driver" function(s). Here we discuss mutation of GATA3 in breast cancer and the potential mechanism(s) by which mutation may lead to a growth advantage in cancer.

Key words: Oncogene; GATA3; Tumor suppressor; Breast cancer

INTRODUCTION

GATA3 is a member of a family of transcriptional regulators with the capacity to function in determination of cell identity (1). Its expression is associated with cell type specification in the immune system, where its action is integral to the differentiation of multiple cell types (2). In the mammary gland, GATA3 is expressed in the differentiated luminal epithelial cells lining the breast ductal structures, where it is estimated to be the most highly expressed transcription factor (3,4). Conditional deletion of GATA3 around puberty results in failure to form terminal end buds with concomitant failure of mammary gland morphogenesis. Conditional deletion in adult animals results in severe defects in epithelial cells, including loss of luminal identity (3,5). These genetic data are consistent with a fundamental role for GATA3 in establishment and maintenance of luminal cell identity.

In the context of breast cancer, GATA3 is also intimately associated with luminal cell identity and function. In mouse models of breast cancer, GATA3 expression is lost as luminal epithelial cells lose differentiated status and progress toward metastasis (6). In human breast tumors, expression of four transcription factors—GATA3, estrogen receptor α (ER-α), FOXA1, and XBP1—is diagnostic of the Luminal A and Luminal B subtypes (7). Molecular analyses indicate that GATA3, ER-α, and FOXA1 function together at multiple loci in the breast cancer genome (8,9) and that their concerted action may be sufficient for luminal epithelial cell identity (9).

Recent genomic analyses of breast cancer have identified GATA3 as a high-frequency target of mutation in this disease (10–13), a finding previously reported on a smaller scale by Perou and colleagues (14). Mutations in GATA3 were observed in approximately 10% of tumors in all four genomic studies. Importantly, most mutations were limited to a single allele, and expression of both mutated and wild-type alleles was approximately equivalent (10–13). These observations raise important questions regarding the nature of the mutations in GATA3 in breast cancer. In this article, we summarize biochemical and molecular properties of GATA3 and provide a framework for understanding how the mutations observed in breast tumors may provide a growth advantage.

DOMAIN ARCHITECTURE OF GATA3

Shortly after the discovery of GATA3 as an activator of the T-cell receptor (15,16), Engel and colleagues used a series of mutations and gene fusions to define functional domains within the protein (Fig. 1A). Using classic deletion analysis, transcriptional response driven by a GATA

Address correspondence to Paul A. Wade, National Institute of Environmental Health Science, 111 TW Alexander Drive, Mail Drop D4-04, Research Triangle Park, NC 27709, USA. E-mail: wadep2@niehs.nih.gov

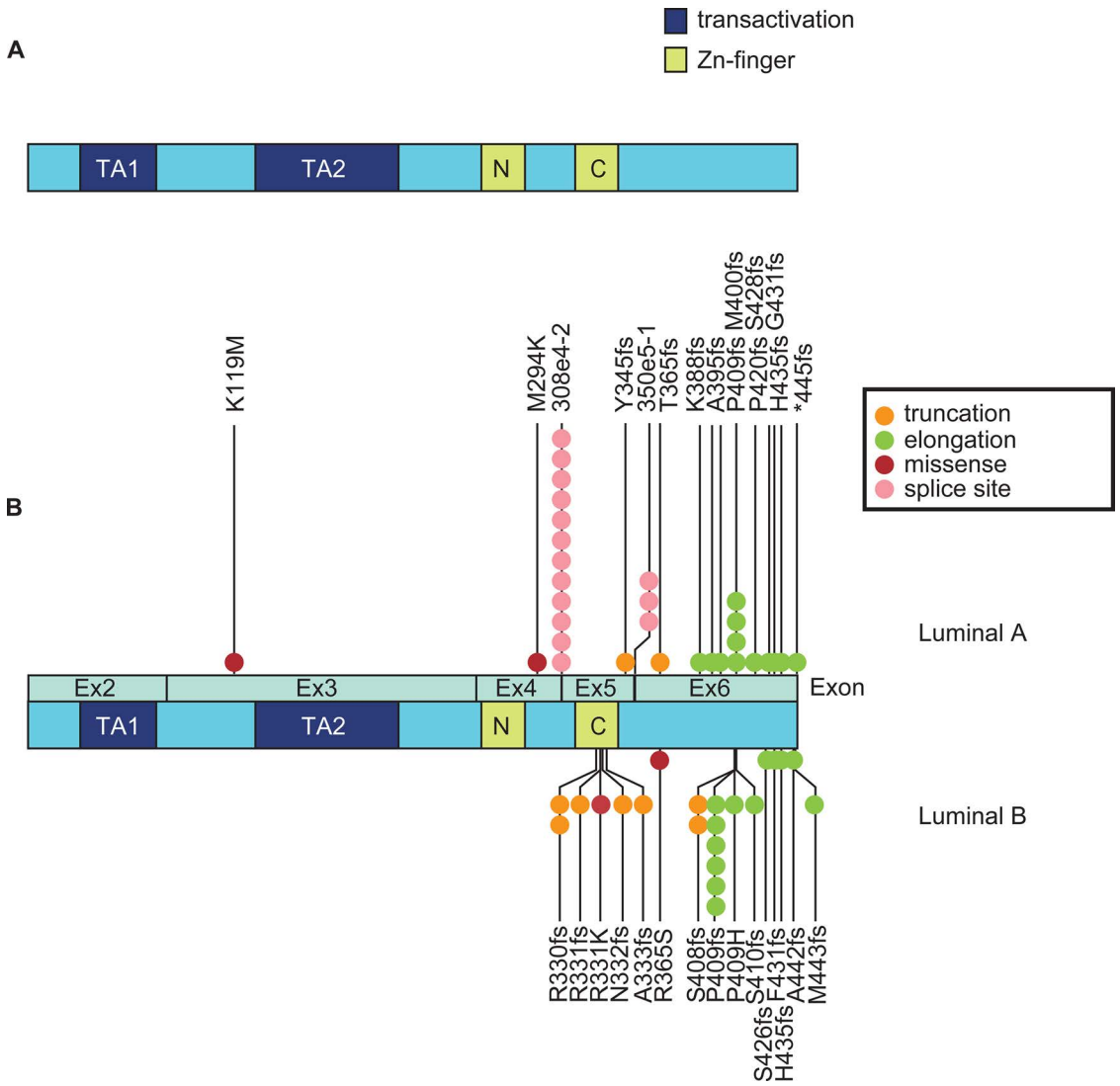

Figure 1. GATA3 domain structure. (A) The model depicts the known functional domains in GATA3. The two transcriptional activation domains (TA1 and TA2) are depicted in dark blue; the two zinc finger motifs (N and C) are depicted in yellow. (B) The model depicts the known functional domains in GATA3 along with the exon structure. Mutations found in GATA3 in breast cancer are indicated in the figure. Mutations found in Luminal A tumors are indicated above the model; mutations found in Luminal B tumors are indicated below the model. Inset indicates symbols for depicting different classes of mutation.

response element was shown to require two separate protein segments. Amino acids 31–59 were required for the function of the first *trans* activation domain (TAI), and amino acids 132–214 defined a second *trans* activation domain (TAII). Loss of the individual activation domains decreased reporter output by 5- to 10-fold, and combined deletion decreased reporter output by 100-fold. Surprisingly, when regions of GATA3 were fused to a heterologous DNA-binding domain, TAI, but not TAII, activated reporter gene expression (17).

All GATA family members, including GATA3, recognize a hexanucleotide response element centered on the GATA motif. In the case of GATA3, the consensus recognition element is (A/T)GATA(A/G) (18). Deletion of the amino-terminal zinc finger moiety (ZF1) had no impact on in vitro binding to a consensus GATA element, but resulted in a modest increase in *trans* activation function. Conversely, deletion of ZF2 abolished both DNA binding and *trans* activation (17).

GATA3 AND BREAST CANCER

GATA3 is a prominent marker of the luminal pattern of gene expression (19); loss of GATA3 expression is associated with tumor types with a propensity for invasive growth and poor prognosis (20). Mechanistically, this relationship seems to spring from two properties of

GATA3. First, GATA3 and ER-α participate in a positive feedback loop, each stimulating expression of the other (9,21). Second, GATA3 has been identified as an important negative regulator of tumor features associated with poor prognosis. In human cell lines (22) and in animal models (23,24), GATA3 suppressed the expression of factors critical to epithelial to mesenchymal transition and metastasis. Enforced GATA3 expression in a mouse model for breast cancer (MMTV-driven polyoma virus middle T antigen) led to increased differentiation and decreased dissemination and metastasis (6). These expression and model system data indicate that GATA3 may play roles in breast cancer associated with promoting growth (through stimulating expression of ER-α) as well as in maintenance of a phenotype with favorable prognosis.

The initial publication by the Cancer Genome Atlas project identified 55 somatic GATA3 mutations (a total of 10.7% of all patients) in their cohort of 512 sequenced tumors (13). Virtually all mutations (53 of 55) were found to impact exons 5 and 6 of the GATA3 protein, which code for the second zinc finger and the relatively understudied carboxyl terminus (Fig. 1B). Three major classes of mutations were described: (1) splice site mutations at the exon 4/5 junction and the exon 5/6 junction, (2) frameshift mutations in exon 6, and (3) frameshift mutations in zinc finger 2.

The splice site mutations at the exon 4/5 junction (total of 12 tumors) are all reported by TCGA to have a CA deletion at the splice acceptor side of the roughly 5-kb intron between exons 4 and 5. This particular CA dinucleotide lies immediately adjacent to the terminal nucleotide of the intron (Fig. 2) within the consensus acceptor site: (Y-rich) – N – **C** – **A** – G –[cut] – G. Examination of RNA-seq reads from these tumors reveals that in 11 of 12 instances, the splice acceptor site is displaced by the CA deletion generating a new splice acceptor site 7 nucleotides downstream (Fig. 2). Utilization of this new splice acceptor results in a new reading frame in exon 5, inclusion of 44 missense amino acids, followed by a stop codon. These events, occurring in 20% of tumors with GATA3 mutation, result in loss of zinc finger 2 and loss of the carboxyl terminus of wild-type GATA3. These splice-acceptor site mutations are exclusive to Luminal A tumors. Given the requirement for ZF2 for productive recognition of the GATA consensus element, this class of mutant proteins is predicted to lack specific DNA binding capacity. Whether protein products are stable in cells and whether any truncated proteins can dimerize with wild-type GATA3 is currently unknown.

The abundant frameshift mutations in exon 6 are particularly interesting and are found in both Luminal A and Luminal B tumor types. There are 25 reported frameshift mutations in exon 6 in the original TCGA breast cancer publication (13)—accounting for roughly 50% of all GATA3 mutations reported. These mutations are scattered throughout exon 6 from amino acid 388 onward, including an example at the stop codon. Slightly under half occur at proline 409 (Fig. 1B). All of these frameshift mutations alter the reading frame at the carboxyl terminus of GATA3 to the same reading frame, which extends the protein from 444 to approximately 500 amino acids (506 amino acids in most cases). It is currently unknown whether these "extension" frameshift mutants produce detectable protein products.

The final class of mutations occurs strictly in Luminal B tumors via frameshifts in zinc finger 2. All of these mutations (just under 10% of all GATA3 mutations in the TCGA report) are predicted to disrupt function of the second zinc finger; all result in premature protein truncation. Unlike the other two major classes of mutation in GATA3 in breast cancer, this class of mutations has been previously studied. Perou and colleagues reported this type of GATA3 mutation in breast tumors a decade ago, including a frameshift mutation in zinc finger two on one allele in the highly utilized MCF7 cell line (14). In this cell line, both wild-type and mutant versions of GATA3 are expressed, with the mutant form appearing to be more abundant at steady state (14,25). As expected, frameshift mutation in the zinc finger has a negative impact on DNA-binding activity. Surprisingly, biochemical fractionation experiments demonstrated that a pool of mutant GATA3 in MCF7 has a salt extraction profile from chromatin identical to wild type, suggesting dimerization of mutant and wild-type proteins (25), a feature also observed in wild-type GATA3 (26). The truncated mutant GATA3 in MCF7 was also shown to be stabilized to physiologic turnover mechanisms induced by estradiol, a novel property that was also conferred on wild-type protein. These observations suggested that this class of mutants may be unusually stable in cells and that the carboxyl terminus serves to regulate physiologic protein turnover (25). This model provides a plausible mechanism by which this class of mutants might provide a growth advantage to tumors. As most luminal tumors are estradiol dependent for growth, and estradiol levels are cyclic in women, an increase in protein half-life of an ESR1-cooperating transcription factor may permit this class of mutants to "bookmark" GATA3-dependent enhancer sites.

A UNIFYING MODEL

GATA3 plays an essential role in the normal development and function of the mammary gland where it promotes a transcriptional program specifying luminal cell identity. In breast cancer, GATA3 represents a useful marker for luminal category tumors. In this abundant class of tumors, GATA3 function is postulated to (1) participate along with estrogen receptor in growth promotion, (2) contribute to differentiated status by regulation

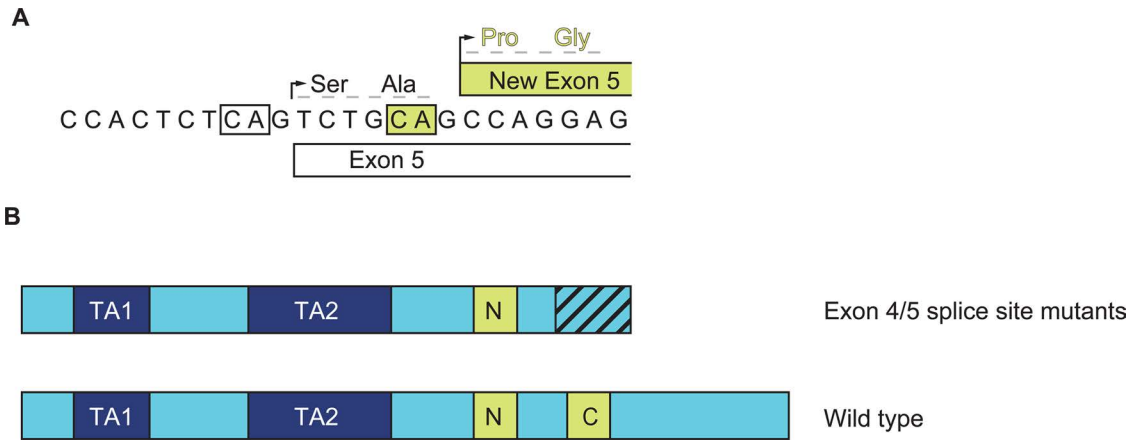

Figure 2. Mutations at the exon 4/5 junction. (A) The figure depicts the sequence of the primary transcript for human GATA3. The CA dinucleotide commonly deleted in breast cancer is boxed. Amino acids encoded by exon 5 are depicted directly above the sequence. The CA dinucleotide utilized by the deletion mutants at the splice site is depicted in the yellow box in the RNA sequence. The new boundary of the new exon 5, along with the missense amino acid sequence, is depicted above the RNA sequence in green. (B) The model depicts the domain structure of wild-type GATA3 along with that of the exon 4/5 splice site mutant proteins. The hatched region in the mutant indicates the 44 missense amino acids that are encoded by the mutant RNA. Note that these mutations cause loss of the C-terminal zinc finger.

of epithelial specification genes, and (3) repress gene products integral to poor prognosis, basal tumor types.

The advent of high-throughput genomic studies of cancer has provided a wealth of new knowledge about GATA3. The gene is mutated in a high frequency of breast tumors, approximately 10%, with both wild-type and mutant products typically expressed (13). The observed frequency of mutation suggests that mutant GATA3 is a driver of breast cancer, while it remains unclear whether GATA3 acts as a tumor suppressor or as an oncogene. Recently, Vogelstein and colleagues addressed molecular definition of oncogenic mutations in the context of genomic surveys of cancer (27). They proposed that oncogenes are characterized by recurrent mutation at diagnostic sites, whereas tumor suppressors are typically altered by truncation mutations throughout their length. In applying this logic to GATA3 in breast cancer, the known mutations are highly clustered in exons 5 and 6. While they do not localize to a few diagnostic amino acids as exemplified by IDH1 mutations (27), it seems likely that clustered mutations within one or two functional protein domains may be consistent with the proposed definition of an oncogene.

The mutational spectrum in GATA3 is somewhat difficult to pinpoint as generating gains or losses of function. The observation that these mutations occur in a heterozygous context, maintaining expression of a fully wild-type allele, implies, but does not prove, gain of function. In the case of two major classes of GATA3 mutations, splice site mutations at the exon 4/5 junction and truncating mutation in zinc finger 2, the obvious prediction would be loss of function. These mutations result in loss of the protein motif required for specific recognition of the GATA site. However, the biochemical data derived from one exemplar mutation, the ZF2 frameshift found in MCF7, are not entirely consistent with these mutations generating a growth advantage through loss of function. While the mutant protein clearly lacks sequence-specific DNA-binding capacity, it is also stabilized to regulation at the protein level (14,25), a feature that is also conferred on the wild-type protein expressed in the same cells.

A unifying model explaining the mechanism by which the mutations observed in breast cancer confer a growth advantage (Fig. 3) posits that protein level regulation of GATA3 is integral to its normal physiology in luminal epithelial cells. Further, this model predicts that the carboxyl terminus of the protein serves a critical role in this posttranscriptional regulation. Mutations that result in premature truncation, including splice site mutations at the exon 4/5 junction and truncating frameshift mutations in exon 5, should result in aberrant protein stability, precisely as observed in MCF7 (14,25). As GATA3 impacts on the ER-α distribution (8), increases in protein half-life alter the time window during which the mitogenic properties of estradiol are manifested at the level of gene transcription by maintaining the chromatin configuration at regulatory regions through which ER-α functions. It is tempting to speculate that the large class of frameshift mutations in exon 6 resulting in extension of the open-reading frame may likewise interfere with regulation at the level of protein stability.

The prediction of this model, regulation of GATA3 at the level of protein turnover, is not without precedent in the current literature. In the context of Th2 differentiation,

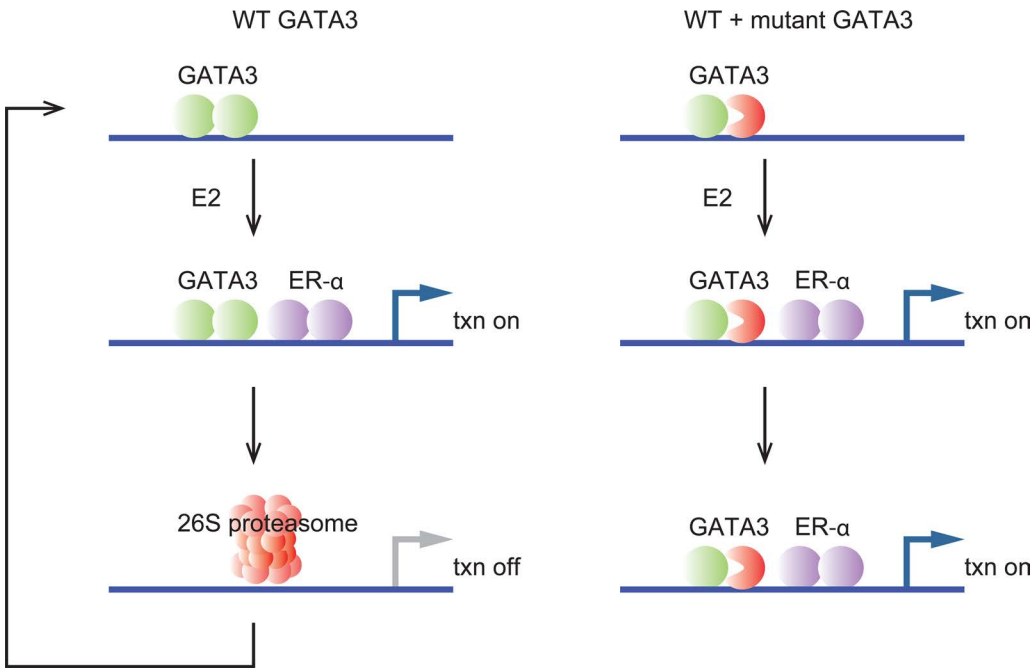

Figure 3. A unifying model for the action of GATA3 mutations in breast cancer. The left-hand panel indicates a model for wild-type GATA3 action in breast cancer. GATA3 influences the binding site selection of estrogen receptor, which activates a set of genes integral to growth in luminal cells. Estrogen receptor is localized in the cytoplasm in the absence of estradiol. Following exposure to hormone, ER-α binds in proximity to GATA3, activates gene expression, and then is rapidly degraded by the 26S proteasome. On the right-hand panel, mutant GATA3 (red) dimerizes with wild-type protein to occupy GATA sites. On exposure to hormone, ER-α binds to these loci and activates gene expression. The GATA3 mutations interfere with normal turnover of GATA3 and of ER-α, lengthening the period during which expression of growth-promoting genes is activated.

GATA3 stability is regulated by a signaling cascade and the ubiquitin-proteasome pathway (28), a finding mirrored in regulatory T cells (29). In prostate cancer, GATA3 expression at the protein level is lost in Pten-deficient tumors, and enforced expression leads to delayed tumor progression to invasive and metastatic growth (30). In adipocyte differentiation induced by rosiglitazone in cell lines, increased GATA3 turnover downstream of prolyl hydroxylase function and eventual ubiquitination are required (31). Finally, stabilization of GATA3 at the protein level is associated with the skewed cytokine profile characteristic of Sezary syndrome (32).

GATA3 has emerged in recent years as a critical factor in normal mammary gland function and in breast cancer. The discovery of multiple mutations in a previously uncharacterized region of the protein in cancer genomic studies may provide yet another example of how the study of cancer provides novel insights into biology. While the impact of cancer-specific GATA3 mutations can be predicted from the existing literature, much work remains to define the impact of these mutations on GATA3 function in cancer and what lessons about biology can be extracted from these studies.

ACKNOWLEDGMENTS: The authors thank the members of the Wade laboratory for many useful discussions through the course of this work. This work was supported, in part, by the NIH, National Institute of Environmental Health Sciences (ES101965 to P.A.W.). M.T. was supported, in part, by a fellowship from the Japan Society for the Promotion of Science.

REFERENCES

1. Rothenberg EV. Transcriptional drivers of the T-cell lineage program. Curr Opin Immunol 2012; 24(2):132–138.
2. Tindemans I, Serafini N, Di Santo JP, Hendriks RW. GATA-3 function in innate and adaptive immunity. Immunity 2014; 41(2):191–206.
3. Kouros-Mehr H, Werb Z. Candidate regulators of mammary branching morphogenesis identified by genome-wide transcript analysis. Dev Dyn 2006; 235(12):3404–3412.
4. 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(5):1041–1055.
5. Asselin-Labat ML, 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(2):201–209.
6. 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(2):141–152.
7. Perou CM, Borresen-Dale AL. Systems biology and genomics of breast cancer. Cold Spring Harb Perspect Biol 2011; 3(2).
8. 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(1):12–22.

9. Kong SL, Li G, Loh SL, Sung WK, Liu ET. Cellular reprogramming by the conjoint action of ERalpha, FOXA1, and GATA3 to a ligand-inducible growth state. Mol Syst Biol 2011; 7:526.
10. Stephens PJ, Tarpey PS, Davies H, Van Loo P, Greenman C, Wedge DC, et al. The landscape of cancer genes and mutational processes in breast cancer. Nature 2012; 486(7403):400–404.
11. Banerji S, Cibulskis K, Rangel-Escareno C, Brown KK, Carter SL, Frederick AM, et al. Sequence analysis of mutations and translocations across breast cancer subtypes. Nature 2012; 486(7403):405–409.
12. Ellis MJ, Ding L, Shen D, Luo J, Suman VJ, Wallis JW, et al. Whole-genome analysis informs breast cancer response to aromatase inhibition. Nature 2012; 486(7403):353–360.
13. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. Nature 2012; 490 (7418):61–70.
14. Usary J, Llaca V, Karaca G, Presswala S, Karaca M, He X, et al. Mutation of GATA3 in human breast tumors. Oncogene 2004; 23(46):7669–7678.
15. Ho IC, Vorhees P, Marin N, Oakley BK, Tsai SF, Orkin SH, et al. Human GATA-3: A lineage-restricted transcription factor that regulates the expression of the T cell receptor alpha gene. EMBO J 1991; 10(5):1187–1192.
16. Ko LJ, Yamamoto M, Leonard MW, George KM, Ting P, Engel JD. Murine and human T-lymphocyte GATA-3 factors mediate transcription through a cis-regulatory element within the human T-cell receptor delta gene enhancer. Mol Cell Biol 1991; 11(5):2778–2784.
17. 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(3):2201–2212.
18. Ko LJ, Engel JD. DNA-binding specificities of the GATA transcription factor family. Mol Cell Biol 1993; 13 (7):4011–4022.
19. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. Nature 2000; 406(6797):747–752.
20. 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(24):11259–11264.
21. Eeckhoute J, Keeton EK, Lupien M, Krum SA, Carroll JS, Brown M. Positive cross-regulatory loop ties GATA-3 to estrogen receptor alpha expression in breast cancer. Cancer Res 2007; 67(13):6477–6483.
22. 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(18):14042–14051.
23. Dydensborg AB, Rose AA, Wilson BJ, Grote D, Paquet M, Giguere V, et al. GATA3 inhibits breast cancer growth and pulmonary breast cancer metastasis. Oncogene 2009; 28(29):2634–2642.
24. Chou J, Lin JH, Brenot A, Kim JW, Provot S, Werb Z. GATA3 suppresses metastasis and modulates the tumour microenvironment by regulating microRNA-29b expression. Nat Cell Biol 2013; 15(2):201–213.
25. 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.
26. Bates DL, Chen Y, Kim G, Guo L, Chen L. Crystal structures of multiple GATA zinc fingers bound to DNA reveal new insights into DNA recognition and self-association by GATA. J MolBiol 2008; 381(5):1292–1306.
27. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Jr., Kinzler KW. Cancer genome landscapes. Science 2013; 339(6127):1546–1558.
28. Yamashita M, Shinnakasu R, Asou H, Kimura M, Hasegawa A, Hashimoto K, et al. Ras-ERK MAPK cascade regulates GATA3 stability and Th2 differentiation through ubiquitin-proteasome pathway. J Biol Chem 2005; 280(33):29409–29419.
29. Zhang J, Chen C, Hou X, Gao Y, Lin F, Yang J, et al. Identification of the E3 deubiquitinase ubiquitin-specific peptidase 21 (USP21) as a positive regulator of the transcription factor GATA3. J Biol Chem 2013; 288(13):9373–9382.
30. Nguyen AH, Tremblay M, Haigh K, Koumakpayi IH, Paquet M, Pandolfi PP, et al. Gata3 antagonizes cancer progression in Pten-deficient prostates. Hum Mol Genet 2013; 22(12):2400–2410.
31. Kim J, Kwak HJ, Cha JY, Jeong YS, Rhee SD, Cheon HG. The role of prolyl hydroxylase domain protein (PHD) during rosiglitazone-induced adipocyte differentiation. J Biol Chem 2014; 289(5):2755–2764.
32. Gibson HM, Mishra A, Chan DV, Hake TS, Porcu P, Wong HK. Impaired proteasome function activates GATA3 in T cells and upregulates CTLA-4: Relevance for Sezary syndrome. J Investig Dermatol 2013; 133(1):249–257.