Soy-derived phytochemical genistein modifies chromatome topology to restrict cancer cell proliferation

Dutta, Bamaprasad; Park, Jung Eun; Qing, Ivan Toh Yi; Kon, Oi Lian; Sze, Siu Kwan

2018

Dutta, B., Park, J. E., Qing, I. T. Y., Kon, O. L., & Sze, S. K. (2018). Soy-derived phytochemical genistein modifies chromatome topology to restrict cancer cell proliferation. *Proteomics*, 18(16), 1700474-. doi:10.1002/pmic.201700474

<https://hdl.handle.net/10356/139282>

<https://doi.org/10.1002/pmic.201700474>

This is the accepted version of the following article: Dutta, B., Park, J. E., Qing, I. T. Y., Kon, O. L., & Sze, S. K. (2018). Soy-derived phytochemical genistein modifies chromatome topology to restrict cancer cell proliferation. *Proteomics*, 18(16), 1700474-., which has been published in final form at 10.1002/pmic.201700474. This article may be used for non-commercial purposes in accordance with the Wiley Self-Archiving Policy [<https://authorservices.wiley.com/authorresources/Journal-Authors/licensing/self-archiving.html>].

Downloaded on 26 Apr 2023 21:57:33 SGT

1   **Soy-derived phytochemical genistein modifies chromatome topology to**
2   **restrict cancer cell proliferation**

3   Bamaprasad Dutta¹, Jung Eun Park¹, Ivan Toh Yi Qing¹, Oi Lian Kon² and Siu Kwan Sze^{1*}

4   ¹School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, Singapore 637551.

5   ²Division of Medical Sciences, National Cancer Centre Singapore, 11 Hospital Drive, Singapore 169610.

7   Running title: Genistein inhibits cell proliferation via epigenetic reprogramming.

11   ***Correspondence:**

12   Siu Kwan SZE, PhD

13   School of Biological Sciences

14   Division of Structural Biology and Biochemistry

15   Nanyang Technological University,

16   60 Nanyang drive, Singapore 637551

17   Tel: (+65) 6514-1006

18   Fax: (+65) 6791-3856

19   Email: sksze@ntu.edu.sg

Abstract

Epidemiological data indicate that human cancer risk is significantly reduced by consumption of soy-based foods containing the 'phytoestrogen' genistein, which can signal via host cell estrogen receptors. While additional chemoprotective effects of genistein induced by epigenetic factors have also been reported, the key molecules and mechanisms involved are poorly defined. We therefore investigated genistein effects on chromatin-bound proteins in the estrogen receptor-deficient cell line MDA-MB-231 which is insensitive to phytoestrogen signaling. After exposure to low-dose genistein for >1-month, MDA-MB-231 cells exhibited stable epigenetic alterations that were analyzed via partial MNase digestion and TMT-based quantitative proteomics. We identified 3177 chromatin-bound proteins with high confidence, including 882 molecules that displayed altered binding topology after cell conditioning with genistein. Prolonged phytochemical exposure conferred heritable changes in the binding topology of key epigenetic regulators including ATRX, SUV39H1/H2, and HP1BP3 that were preserved in untreated progeny, resulting in sustained downregulation of proliferation genes and reduced cell growth. These data indicate that soy derivative genistein exerts complex estrogen receptor-independent effects on the epigenome likely to influence tumorigenesis by restricting cell growth.

Keywords: Genistein, epigenetics, epigenome, epigenetic reprogram, chromatin, chromatome, euchromatin, heterochromatin, cell proliferation, TMT, quantitative proteomics.

1   **Introduction**

2   Genistein is a soy-derived ‘phytoestrogen’ that has been reported to exert potent
3   chemoprotective effects both *in vitro* and *in vivo*^[1]. Accordingly, several epidemiological studies
4   have reported that high consumption of soy-based foods is associated with reduced incidence of
5   breast cancer^[2]. There is also substantial experimental evidence that genistein is effective in
6   protecting against several other types of cancer including leukemia, prostate, colon and bladder
7   carcinoma^[3, 4]. While the anti-tumor effects of genistein have been largely attributed to signaling
8   via estrogen receptors, this phytochemical is also capable of influencing human cell function via a
9   range of alternative mechanisms. Genistein can disrupt oncogenic signaling and reduce
10   genotoxicity in human cells by inhibiting the activity of enzymes including tyrosine kinases and
11   topoisomerase II^[3, 5]. Other chemoprotective effects of genistein exposure have been reported to
12   include promotion of tumor cell cycle arrest and apoptosis^[6], as well as disruption of
13   angiogenesis^[7] and cell invasion^[8], but the molecular basis of these beneficial effects in human
14   cells is unclear.

15   Recent studies in rodent models have demonstrated that repeated dietary exposure to
16   genistein confers permanent alterations in DNA methylation patterns and gene expression that may
17   reduce cancer risk^[9, 10]. Dietary genistein can prevent carcinogen-induced neoplasia via epigenetic
18   modulation of oncogenic gene expression^[11], but also modify the activity of epigenetic regulatory
19   enzyme enhancer of zeste homolog 2 (EZH2) to promote uterine tumorigenesis via differential
20   methylation of histone H3K27^[12]. In prostate cancer cells, genistein-induced changes in
21   methylation and acetylation marks at histone H3K9 can reactivate tumor suppresser genes^[13], and
22   dietary exposure to this phytochemical may also be capable of inhibiting DNA
23   methyltransferase^[14]. Antitumor effect of genistein through epigenetic regulation also been

1 reported in several recent studies^[15]. However, the key molecular mediators of these disparate
2 effects remain unknown and their dependence on phytoestrogen signaling has not been established.

3 To identify whether genistein exhibits significant estrogen receptor-independent
4 chemoprotective properties, we used our optimized proteomic methods^[16, 17] to assess
5 phytochemical-induced changes in the chromatinome of the estrogen receptor-deficient breast
6 cancer cell line MDA-MB-231. After prolonged treatment with low-dose genistein to mimic
7 dietary exposure, MDA-MB-231 cells displayed durable changes in epigenetic profile and gene
8 expression that were heritable by progeny cells even after phytochemical withdrawal. Using partial
9 Micrococcal nuclease (MNase) digestion together with Tandem mass tags (TMT) based
10 quantitative proteomic techniques, we observed that genistein conditioning induced stable
11 modification in the binding topology of a subset of chromatin-organizing proteins including
12 Transcriptional regulator ATRX (ATRX), Chromobox protein homolog 2 (CBX2), Chromobox
13 protein homolog 5 (CBX5), Methyl-CpG-binding domain protein 3 (MBD3), Histone-lysine N-
14 methyltransferase SUV39H1 (SUV39H1), Histone-lysine N-methyltransferase SUV39H1
15 (SUV39H2), and Heterochromatin protein 1-binding protein 3 (HP1BP3). Genistein exposure
16 promoted compaction of genetic material into heterochromatin and downregulated the expression
17 of cell proliferation genes to suppress cell growth. Future attempts to exploit dietary genistein for
18 its chemoprevention effects on epigenetic regulation

19 **Materials and methods**

20 **Reagents**

21 All reagents were purchased from Sigma-Aldrich (Sigma-Aldrich, WI) unless otherwise
22 indicated. Antibodies against GAPDH (sc-32233) was from Santa Cruz Biotechnology Inc., USA.

Anti-beta Actin (ab8227), -histone H2A (ab13923), histone H4 (ab10158), and HP1BP3 (ab98894) antibodies were purchased from Abcam, UK. The anti-actin antibody (MAB1501) was purchased from Millipore, USA. Anti DNMT1 (#5032) and CHD1 (#4351) were purchased from Cell Signaling Technology, MA, USA.

Experimental methodology

Cell culture and genistein treatment

The MDA-MB-231 and MCF7 cell line were purchased from ATCC and maintained in DMEM supplemented with 10% FBS. For each experiment, 2×10^6 cells were seeded into a 10cm Petri dish and incubated at 37°C in 5% CO₂ in the presence of 4μM genistein (treated) or Dimethyl sulfoxide (DMSO) (vehicle control). The cells were subculture in every 2days interval and treatment was continued for a total of 8 weeks. In some experiments, the cells were cultured in genistein-supplemented medium for 4 weeks only then cultured for a further 4 weeks in DMEM (genistein withdrawal).

Chromatin isolation and digestion

Chromatin isolation was carried out as described previously^[16] except for minor modifications. In brief, cells were homogenized in nuclei extraction buffer A (10mM Tris-HCl pH 7.4, 15mM NaCl, 60mM KCl, 150mM Sucrose, 1mM Ethylenediaminetetraacetic acid (EDTA), 1% TritonX-100, 0.1% NP40, 1mM Dithiothreitol (DTT), and protease inhibitor cocktail from Roche Diagnostics, Mannheim, Germany) prior to isolation of nuclei via centrifugation of the homogenate at 500×g for 10min at 4°C. The nuclei were further purified using sucrose cushion separation before being lysed in extraction buffer B (10mM Tris-HCl pH 7.4, 15mM NaCl, 60mM KCl, 150mM Sucrose, 1mM EDTA, 0.1% Triton X-100, 1% NP40, 1mM DTT, and protease

inhibitor cocktail). The chromatin fraction was isolated by centrifugation of the nuclear lysate at 20000×g for 10min at 4°C. The chromatin pellet was washed three times with wash buffer (10mM Tris-HCl pH 7.4, 15mM NaCl, 60mM KCl, 1mM EDTA, 1mM DTT and protease inhibitor cocktail) to remove the soluble nuclear contaminants. The purified chromatin pellet was suspended in 50µl MNase digestion buffer (50mM Tris-HCl pH 7.9, 5mM CaCl₂,) and digested with 10KU MNase (New England Biolabs, Singapore) for 60min at 37°C. Digestion was terminated by addition of 1mM EDTA and the supernatant fraction (fraction S) collected by centrifugation for 30min at 20,000×g, 4°C. The remaining undigested pellet (fraction P) was dissolved in 0.5% SDS using sonication. The protein content of the supernatant and pellet fractions were each quantified using the bicinchoninic acid (BCA) assay method. Cells from three biological replicates were pooled prior to the chromatin isolation in order to minimize the effects of processing error. The details of the experimental design are outlined in [Fig. S1](#).

In-gel protein digestion and TMT labeling

A total of 100µg protein from each sample was loaded into 15% SDS-PAGE gels for electrophoresis at 100V for 30min. Each sample lane was cut into small pieces and washed with 25mM triethylammonium bicarbonate (TEAB) in 50% acetonitrile (ACN). The gel pieces were then dehydrated with 100% ACN and vacuum dried. Reduction was carried out using 5mM Tris 2-carboxyethyl phosphine hydrochloride (TCEP) in 50mM TEAB buffer at 60°C for 30min, followed by alkylation with 55mM iodoacetamide (IAA) in 50mM TEAB buffer at room temperature for 45min. TCEP and IAA were removed by performing alternate washes with 50mM TEAB buffer and 50mM TEAB in 50% ACN. The gel pieces were then dehydrated and dried for a second time. The gel pieces were completely rehydrated with 10ng/ml of sequencing-grade modified trypsin solution (Promega Corporation, Madison, WI. USA) in 50mM TEAB (protein to

trypsin ratio of 1:50) on ice and incubated at 37°C for overnight digestion. Tryptic peptides were extracted using 50% ACN and 1% acetic acid before being dried by vacuum centrifugation (Eppendorf, USA). Labeling was performed using TMT reagent multiplex kits according to the manufacturer's protocol (Thermo Scientific, USA). Supernatant and pellet fractions from control, treated, and genistein withdrawal samples were labeled with TMT tags as specified in [Table S1](#). The TMT-labeled peptides were pooled and desalted using Sep-Pak C18 Vac cartridges (Waters, Milford, MA) before being vacuum-centrifuged to dryness.

Fractionation of labeled peptides

TMT-labeled peptides were fractionated by reversed-phase chromatography at ~pH 8. Briefly, the labeled peptides were reconstituted in 200µl buffer A (0.02% NH₄OH in water) and fractionated using X-Bridge C18 columns (4.6×200mm, 5µm particle size, 130Å pore size) (Waters, Milford, MA) on a HPLC unit (Prominence™, Shimadzu, Kyoto, Japan) at flow rate of 1ml/min. The HPLC gradient comprised 100-95% buffer A (0.02% NH₄OH in water) for 3min; 5-35% buffer B (80% ACN and 0.02% NH₄OH) for 40min; then 35-70% buffer B for 12min, followed by 70% buffer B for 5min. The chromatograms were recorded at 280nm. A total of 64 fractions were collected within the 65min elution period. The eluted peptides were combined into 16 distinct fractions using a concatenated pooling strategy and concentrated to dryness using vacuum centrifugation.

LC-MS/MS and data analysis

LC-MS/MS analysis was performed using a Q Exactive mass spectrometer coupled with online nano-HPLC system (Thermo Scientific; USA). Data were acquired using Xcaliber 2.2 software (Thermo Scientific; USA). Protein identification and quantification were performed

using Proteome Discoverer 1.4.1.14 (Thermo Fisher, MA) together with the Sequest and Mascot search engines (parameters included cysteine alkylation with IAA and TMT6plex at peptide N-terminus as static modification; methionine oxidation, and asparagine and glutamine deamidation and TMT6plex at lysine as dynamic modifications; full trypsin digestion and maximum 2 missed cleavages as digestion parameters; 10ppm precursor mass and 0.02Da fragment mass tolerance). TMT 6-plex was selected as the quantification method. The UniProt Knowledgebase (UniProtKB) of human proteins (downloaded on 25th July 2016, including 70,849 sequences and 23,964,784 residues) was used as a search database with a false discovery rate cut-off of <0.05% at the peptide level. The resulting data set was automatically bias-corrected to eliminate potential variation due to unequal mixing when combining the samples. All LC-MS/MS analysis was performed in triplicate. All technical replicates were search together and statistical calculations were performed using experimental duplicates. Cluster analysis of the identified proteins was performed using the online bioinformatics tool Gene Pattern (<http://genepattern.broadinstitute.org>) followed by hierarchical clustering and Pearson correlation.

Western blot analysis

Chromatin isolates were dissolved in RIPA buffer (25mM Tris-HCl, pH 7.4; 150mM NaCl, 0.1% SDS, 0.5% sodium deoxycholate, 1% NP-40 and protease inhibitor cocktail) and the protein content determined using the BCA method. Equal amounts of protein from each sample were used for Western blot analysis. Protein samples were resolved on 10% SDS-PAGE gels and transferred onto nitrocellulose membranes. Immunoblotting was performed using anti-protein antibodies together with the ECL detection system (Invitrogen, USA).

Chromatin compaction assay

Cellular chromatin was isolated from cells incubated under control, treated, and genistein withdrawal conditions according to the protocol outlined above. The chromatin compaction assay was performed as described previously^[16] with minor modifications. Isolated chromatins were digested with different concentration of MNase and DNA fragments were extracted using phenol-chloroform. DNA content was measured using a Nano Drop 2000 UV-Vis Spectrophotometer (Thermo Fisher Scientific, USA). Equal amounts of DNA from each sample were then loaded and resolved on 1% agarose gels for analysis.

Histone Solubility Assay

Total nonchromatin-associated histone was measured using histone solubility assay as described by Michael *et al.*^[18]. In brief, nuclei were isolated by using extraction buffer A (0.3M sucrose, 15mM Tris-HCl (pH 7.6), 60mM KCl, 15mM NaCl, 5mM MgCl₂, 0.1mM EDTA, 0.5mM DTT, 0.4% NP-40 and protease inhibitor cocktail) at 4°C and purified using sucrose cushion (1.2M sucrose, 60mM KCl, 15mM NaCl, 5mM MgCl₂, 0.1mM EDTA, 0.5mM DTT and protease inhibitor cocktail). Soluble nonchromatin-associated histone fraction was extracted from the purified nuclei with histone extraction buffer (320mM sucrose, 50mM Tris (pH 7.6), 4mM MgCl₂, 1mM CaCl₂, 1% NP-40 and protease inhibitor cocktail). Nonchromatin-associated histone H4 in MCF7 and MDA-MB-231 cells was measured by immunoblotting while beta actin was used as loading control.

MTT assay

Cells cultured under different conditions for 8 weeks (control, treated, or genistein withdrawal) were transferred into medium with or without genistein accordingly in a 96 well plate at a density of 1×10^4 cells/well and grown for 24h prior to addition of 0.5mg/ml MTT reagent to

each well. After 2h incubation at 37°C, the culture medium was removed, and resultant formazan crystals were dissolved in DMSO. Optical density was measured at 570nm using a micro-plate reader with a reference wavelength of 630nm (Tecan Magellan™, Switzerland). All the experiments were performed in triplets.

Clonogenic assay

Cells grown under different conditions (control, treated, or genistein withdrawal) for 8 weeks were subsequently seeded at a density of 500 cells/well into 6-well plates in DMEM supplemented with 10% FBS and incubated at 37°C in 5% CO₂ for 7d. For treated group, 4μM genistein was added into the medium. After 7d, culture medium was removed, and the cell colonies were washed, fixed with 95% ethanol, and stained with 5% crystal violet. The number of colonies formed per well was counted for quantification.

Cell cycle analysis

Cells were harvested, fixed and stained with propidium iodide for assessment of DNA content using a BD FACS-Calibur™ flow-cytometer (BD Biosciences, Singapore) and data analysis using Weasel software (version 3.3.2; copyright (c) 2015 Frank Battye).

Quantitative reverse transcription polymerase chain reaction (qRT-PCR)

RNA was extracted from cells using TRIzol® reagent (Ambion, Life Technologies, CA, USA) according to the manufacturer's protocol. First strand cDNA was synthesized from 2μg total RNA using RevertAid H Minus First Strand cDNA Synthesis Kits (Thermo Fisher Scientific). The primer sequences are detailed in supplementary [Table S2](#). The primers for E-Box genes were

designed as described by Xiong *et al.*^[19]. Real-time PCR amplification was performed using KAPA SYBR® FAST qPCR Kits (KAPA Biosystems, Boston, USA) with a CFX Connect™ Real-Time PCR Detection System (Bio-Rad Laboratories, Singapore). Cycling temperatures were as follows: enzyme activation at 95°C for 3min, denaturation at 95°C for 10sec, annealing and extension at 60°C for 30sec. Data were normalized to the expression level of 18S ribosomal RNA in each sample. All experiments were performed in triplicate.

Results

Genistein-conditioned cell culture model and proteomic analysis

Epidemiological data indicate that human cancer risk is reduced by regular intake of soy-based foods containing the ‘phytoestrogen’ genistein, which signals via estrogen receptors and can reach concentrations of up to 6μM in the peripheral blood of human consumers^[20]. To determine the impact of this compound on cancer epigenetics, we assessed the effect of long-term exposure to physiological levels of genistein (4μM for 8 weeks) in the estrogen receptor-deficient cell line MDA-MB-231. In some experiments, the cells were cultured in genistein-supplemented medium for 4 weeks and then transferred into normal medium for a further 4 weeks to test the effects of genistein withdrawal. Equal culture duration (8 weeks) was used for all experimental conditions and the dose of genistein administered did not exceed 4μM, which is >50-fold lower than the EC50 value in MDA-MB-231 cells (Fig. S2).

Environmental stimuli alter the binding topology of chromatin-associated (chromatome) proteins to induce changes in gene expression that are heritable by progeny cells^[21]. To understand how this ‘epigenetic remodeling’ influences human cell biology, we have developed a proteomics-based approach to assessing change in chromatome dynamics induced by altered environmental

conditions^[16, 17]. As outlined in figure 1, this method employs partial MNase digestion of chromatin samples isolated from cells cultured under different experimental conditions prior to sample analysis by TMT-based quantitative proteomics. By identifying dynamic changes in the composition of both the euchromatin-associated proteins (solubilized fraction) and predominantly heterochromatin-associated proteins (undigested fraction), this approach allows unbiased analysis of chromatin topology and how it is modified by environmental factors.

Identification and quantification of the genistein-modified chromatin

In order to identify the effects of genistein exposure on chromatin dynamics, MDA-MB-231 cells were subjected to 8 weeks of sustained genistein treatment (T), a brief genistein exposure of 4 weeks followed by phytochemical withdrawal (W), or culture with DMSO vehicle control (C). All cells were cultured for 8 weeks total duration prior to chromatin extraction and partial MNase digestion to separate euchromatin-bound proteins from the undigested heterochromatin, nucleolus and nuclear matrix. Chromatin purity was confirmed by western blot (Fig. S3). For proteomic analysis, chromatin samples were subjected to tryptic digestion and the resulting peptides were labeled with 6-plex TMT reagents. Labeled peptides were separated via high pH reversed-phase fractionation and then analyzed using a Nano LC-coupled hybrid quadrupole-orbitrap mass spectrometer. Proteins were identified and quantified using Proteome Discoverer software v1.4.1.14. Proteomic data were generated from two independent experimental replicates that each analyzed pooled samples from three biological replicates (Total 6 biological replicates) (Fig. S1). Using combined Mascot and Sequest searches, we identified a total of 5316 proteins with more than one component peptide with a false discovery rate (FDR) <1%. All the proteins and peptides identified are listed in the supplementary Table S3-S5. Experimental replicates were performed several months apart so the number of identified proteins in the experimental duplicate

was varied and only proteins detected in both experimental replicates (3177 total) were considered for further analysis (Fig. S4a). Bioinformatics analyses were performed using Uniprot and Nextprot databases; 75% of the identified proteins were localized in nucleus and ~1% had no classified subcellular localization (Fig. S4b). Accordingly, these proteins were found to be primarily involved in nuclear activities including transcriptional regulation (21%), mRNA processing (13%), chromatin organization (12%), replication and DNA repair (6%), cell cycle progression (5%), nuclear transport (4%) and rRNA processing and ribosome biogenesis (4%) (Fig. S4c). Among these were 882 proteins that displayed a significant shift in chromatin binding topology either in the genistein treatment or withdrawal conditions. (≥ 1.2 -fold change in relative abundance between the pellet [heterochromatin] and supernatant fraction [euchromatin] expressed as a ratio [H/E]) (Fig. S4d).

Genistein conditioning modifies cellular chromatin organization

Our proteomic data indicated that a distinct subset of chromatinome proteins displayed altered binding topology after genistein conditioning of estrogen receptor-deficient MDA-MB-231 cells (Fig. 2a). Cluster analysis revealed marked enrichment of these proteins in the undigested pellet fractions of genistein-treated samples compared with controls, indicating a general shift towards heterochromatin binding. Notably, these proteins included several key epigenetic regulators (ATRX, CBX2, CBX5, MBD3, HP1BP3), that maintained preferential binding to heterochromatin even after genistein withdrawal (Fig. 2b). Among these, HP1BP3 displayed an increase in total levels of chromatin (both heterochromatin and euchromatin) association in cells subjected to genistein treatment or withdrawal (Fig. S6). The proteomic results also indicated that genistein exposure promoted durable heterochromatin binding by histone methyltransferase

1 enzymes SUV39H1, SUV39H2, and Histone-lysine N-methyltransferase MLL4 (WBP7), as well
2 as histone demethylase enzymes Lysine-specific demethylase 5A (KDM5A) and Lysine-specific
3 demethylase 5B (KDM5B) (Fig. 2a and Fig. 2b). Indeed, heterochromatin-specific histones
4 including H3 Histone Family Member 3A (H3F3A) and Core histone macro-H2A.1 (H2AFY)
5 were also present in greater quantities in the undigested pellet fractions of cells subjected to either
6 genistein treatment or withdrawal (Fig. 2a and Fig. 2b). These data suggested that prolonged
7 genistein exposure promotes heritable chromatin compaction likely to exert a major influence on
8 gene expression.

9 Consistent with this concept, MNase finger-printing indicated that chromatin from cells
10 subjected to genistein treatment or withdrawal displayed increased resistance to nuclease digestion
11 (Fig. 3a). To further strengthen our proteomic findings, we also measure the soluble histone
12 fraction to determine the total nonchromatin-associated histone which reflect the chromatin
13 compaction state and results showed prominent reduction in nonchromatin-associated histones H4
14 amount in treated and withdrawn cells compare to the untreated MDA-MB-231 cells (Fig. 3b-c).
15 Decrease in unincorporated histone H4 upon genistein treatment reduced the pool of free histones
16 due to incorporation of more nucleosomes^[18] which ultimately increase overall chromatin
17 compaction level in treated and withdrawn cells.

18 Further cluster analysis revealed that chromatin from genistein-treated cells displayed
19 enhanced MNase mediated release of epigenome-editing enzymes including DNA methyl
20 transferases, histone deacetylases and histone demethylase (Fig. 2c and Fig. S5). While most of
21 these proteins displayed a prominent reduction in MNase-mediated release upon genistein
22 withdrawal, a select group exhibited further increases in nuclease sensitivity after phytochemical
23 removal (Histone deacetylase 1 (HDAC1), Histone deacetylase 2 (HDAC2), Lysine-specific

histone demethylase 1A (KDM1A), and DNA (cytosine-5)-methyltransferase 1 (DNMT1)) (Fig. 2c). Together, these data indicate that genistein exerts a wide range of epigenetic effects that display variable persistence and heritability after compound withdrawal. Given that histone deacetylases/demethylases and DNA methyltransferases deactivate transcription via rearrangement of epigenetic marks^[22, 23], increased binding of these enzymes to euchromatin may help maintaining gene suppression even after genistein removal. Collectively, these results indicate that genistein exposure promotes epigenetic mediated gene silencing and likely exerts potent effects on cell functions that may persist in untreated progeny.

Genistein inhibits growth of via downregulation of c-MYC

Our chromatin mapping results indicated that genistein conditioning alters the chromatin binding topology of cell cycle regulatory proteins that exert critical roles in epigenetic control of proliferation^[24, 25]. In particular, key proteins including Retinoblastoma-associated protein (RB1) and Retinoblastoma-like protein 1 (RBL1) were observed to increase their association with euchromatin in both genistein-treated and previously exposed cells while total chromatin binding of these two proteins remained unchanged (Fig. 4a-b). The cyclin D–RB1 signaling pathway is a master regulator of G1/S phase progression and cell proliferation^[25, 26]. Since dissociation of RB1 protein from transcription site is critical for activation of E2F-dependent S phase transcriptional events^[26], increased RB1 binding to euchromatin might delay G1/S transition and disrupt proliferation. Consistent with this concept, our cell cycle analysis revealed the presence of more proliferative S phase cells in the untreated culture condition compared with the genistein-treated or withdrawal conditions in both estrogen receptor positive MCF7 and negative MDA-MB-231 cell lines (Fig. 4c). Other investigators have observed that genistein can also induce G0/G1 cell

cycle arrest in MCF-7 and HB4a cells^[27]. Accordingly, our MTT-based proliferation assay revealed a ~22% and ~19% reduction in MDA-MB-231 cell proliferation after genistein treatment and withdrawal respectively (Fig. 4d). For MCF7 cells, cell proliferation increased (Fig S6) at the beginning due to the estrogenic effect of genistein in below 10 μ M dose^[28] but long term exposure to low dose of genistein no longer exhibits its estrogenic effect and after 8 week treatment period the proliferation was reduced by ~28% and ~24% in treatment and withdrawal samples (Fig. 4d and Fig. S6). Consistence with our MTT and cell cycle assessment results colony forming assay also showed that genistein treatment significantly reduced the cell viability in MDA-MB-231 cells and the colony forming efficiency was reduced by ~1.6-folds and ~1.3-folds in treated and withdrawn condition (Fig. 4e-f). A comparable effect was also observed in MCF7 cells and colony forming efficiency was declined by ~1.7-folds and ~1.4-folds upon genistein treatment and withdrawal (Fig. 4e-f).

Previous data have indicated that transcriptional control of E-box-regulated genes by c-MYC plays a critical role in driving cell proliferation^[19], and our quantitative RT-PCR analysis confirmed that c-MYC was downregulated in cells exposed to genistein (Fig. 5). We also detected a corresponding genistein-induced decrease in expression of E-box genes including CAD, CCND2, E1F2A, LDHA, NCL and ODC in both MCF7 and MDA-MB-231 cells (Fig. 5). These results indicate that heritable, genistein-induced changes in the epigenome negatively regulate E-Box genes which may contribute to suppression of cell proliferation irrespective of antiestrogen effect of genistein.

Discussion

1   The eukaryotic genome is contained within cell nuclei in the form of chromatin, which
2 exhibits dynamic changes in structure that regulate DNA-dependent cellular functions. Chromatin
3 exists in two major interchangeable forms that exhibit variable levels of compaction, with
4 environmental factors stimulating dynamic shifts between accessible euchromatin and highly
5 condensed heterochromatin. Chromatin-binding (chromatome) proteins play major roles in the
6 structural organization of the genome into euchromatin/heterochromatin and are therefore essential
7 regulators of gene expression and epigenetic inheritance^[21, 29]. Recent studies have reported potent
8 effects of the dietary soy derivative genistein on epigenetic regulation^[9, 11, 13, 15], but until now the
9 key molecules and mechanisms underpinning this have remained unclear.

10   Our proteomic strategy for mapping the human cell chromatome revealed that genistein
11 conditioning promotes heterochromatin binding of several key epigenetic regulators of gene
12 suppression, including ATRX, H3.3 and HP1BP3. ATRX/Daxx is an epigenetic editor complex
13 that contributes to the maintenance of heterochromatin via recruitment of histone variant H3.3^[30],
14 which together with methylation of H3K9me3 are hallmarks of transcriptional silencing^[31]. Also
15 detected a shift towards increased heterochromatin binding by HP1BP3 which organizes higher-
16 order chromatin structure and influences cancer progression via cellular reprogramming^[17]. We
17 further observed a genistein-induced accumulation of heterochromatin-specific proteins including
18 CBX2, CBX5 and MBD3, increased association with gene silencing histone variant macro-
19 H2A.1^[32], and elevated binding of histone methyltransferases SUV39H1/2 which promote genome
20 stability^[33]. Importantly, the preferential heterochromatin binding displayed by these proteins after
21 genistein conditioning of estrogen receptor negative MDA-MB-231 cells was maintained upon
22 phytochemical withdrawal. These data therefore indicate that prolonged genistein exposure
23 induces estrogen receptor-independent epigenetic changes that confer heterochromatinization that

are transmissible to unexposed progeny. Long-term genistein exposure in culture also altered gene expression patterns via recruitment of transcriptional silencers inducing HDAC1, HDAC2, KDM1A and DNMT1 to active regions^[22]. These effects were accompanied by increased total chromatin association of cell cycle regulatory protein RB1 and key pathway component KDM5B^[34] which stall cell proliferation in G0/G1 phase^[26, 35]. Additional effects of genistein conditioning included negative regulation of c-MYC and a corresponding decrease in expression of E-Box genes, further suggesting a potent ability of this phytochemical to suppress cell growth^[19]. Indeed, in previous studies we observed a similar increase in chromatin binding of HP1BP3 in association with cell cycle arrest and growth restriction, albeit under low-oxygen conditions^[17]. Together, these data demonstrate that genistein conditioning induces heritable epigenetic changes that negatively regulate gene transcription and suppress cell proliferation in an estrogen receptor-independent manner. Overall our findings suggest that genistein-induced growth suppression might enhance host protection against tumors and reduces the cancer risk upon intake of soy-based foods from the diet.

Supplementary information

Supplementary Figures S1-S6; Supplementary Tables S1-S2, Supplementary Tables S3-S5 (Excel files).

Availability of data and material

The mass spectrometry proteomics raw data files along with Proteome Discoverer search data file (including protein summary) have been deposited to the ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>) via the PRIDE partner repository^[36] with the dataset identifier PXD006178.

The proteins and peptides information from Proteome Discoverer search are available in the Supplementary Tables S3-S5 (Excel files).

Acknowledgement

This work was supported by grants from the NTU iFood (grant #: S006), the Singapore Ministry of Education (Tier 2: Grant ARC9/15), National Medical Research Council of Singapore (NMRC-OF-IRG-0003-2016) and NTU-NHG Ageing Research Grant (Grant ARG/14017).

References

[1] D. C. Knight, J. A. Eden, *Obstet Gynecol* 1996, 87, 897; R. A. Dixon, D. Ferreira, *Phytochemistry* 2002, 60, 205; A. L. Murkies, G. Wilcox, S. R. Davis, *J Clin Endocrinol Metab* 1998, 83, 297; J. L. Limer, V. Speirs, *Breast Cancer Res* 2004, 6, 119; D. M. Tham, C. D. Gardner, W. L. Haskell, *J Clin Endocrinol Metab* 1998, 83, 2223; C. Duffy, K. Perez, A. Partridge, *CA: a cancer journal for clinicians* 2007, 57, 260; D. Ingram, K. Sanders, M. Kolybaba, D. Lopez, *Lancet* (London, England) 1997, 350, 990; L. Gravitz, *Nature* 2011, 471, S5.

[2] K. Hirose, N. Imaeda, Y. Tokudome, C. Goto, K. Wakai, K. Matsuo, H. Ito, T. Toyama, H. Iwata, S. Tokudome, K. Tajima, *British Journal of Cancer* 2005, 93, 15; M. Pudenz, K. Roth, C. Gerhauser, *Nutrients* 2014, 6, 4218.

[3] B. Boyer, J. P. Thiery, *The Journal of cell biology* 1993, 120, 767; F. M. Uckun, W. E. Evans, C. J. Forsyth, K. G. Waddick, L. T. Ahlgren, L. M. Chelstrom, A. Burkhardt, J. Bolen, D. E. Myers, *Science* (New York, N.Y.) 1995, 267, 886.

[4] N. Kurahashi, M. Iwasaki, M. Inoue, S. Sasazuki, S. Tsugane, *Journal of clinical oncology* : official journal of the American Society of Clinical Oncology 2008, 26, 5923.

- 1 [5] X. Y. Huang, A. D. Morielli, E. G. Peralta, Cell 1993, 75, 1145; L. Xu, Y. Ding, W. J.
2 Catalona, X. J. Yang, W. F. Anderson, B. Jovanovic, K. Wellman, J. Killmer, X. Huang, K. A.
3 Scheidt, R. B. Montgomery, R. C. Bergan, Journal of the National Cancer Institute 2009, 101,
4 1141; J. Markovits, C. Linassier, P. Fosse, J. Couprie, J. Pierre, A. Jacquemin-Sablon, J. M.
5 Saucier, J. B. Le Pecq, A. K. Larsen, Cancer research 1989, 49, 5111.
- 6 [6] G. Bergamaschi, V. Rosti, M. Danova, L. Ponchio, C. Lucotti, M. Cazzola, Leukemia
7 1993, 7, 2012; F. Traganos, B. Ardelt, N. Halko, S. Bruno, Z. Darzynkiewicz, Cancer research
8 1992, 52, 6200.
- 9 [7] S. J. Su, T. M. Yeh, W. J. Chuang, C. L. Ho, K. L. Chang, H. L. Cheng, H. S. Liu, H. L.
10 Cheng, P. Y. Hsu, N. H. Chow, Biochemical pharmacology 2005, 69, 307.
- 11 [8] S. C. Mueller, Y. Yeh, W. T. Chen, The Journal of cell biology 1992, 119, 1309.
- 12 [9] K. Vanhees, S. Coort, E. J. Ruijters, R. W. Godschalk, F. J. van Schooten, S. Barjesteh van
13 Waalwijk van Doorn-Khosrovani, FASEB journal : official publication of the Federation of
14 American Societies for Experimental Biology 2011, 25, 797.
- 15 [10] D. C. Dolinoy, J. R. Weidman, R. A. Waterland, R. L. Jirtle, Environmental health
16 perspectives 2006, 114, 567; Q. Xie, Q. Bai, L. Y. Zou, Q. Y. Zhang, Y. Zhou, H. Chang, L. Yi,
17 J. D. Zhu, M. T. Mi, Genes, chromosomes & cancer 2014, 53, 422.
- 18 [11] Y. Zhang, Q. Li, H. Chen, Carcinogenesis 2013, 34, 1756.
- 19 [12] K. L. Greathouse, T. Bredfeldt, J. I. Everitt, K. Lin, T. Berry, K. Kannan, M. L. Mittelstadt,
20 S. M. Ho, C. L. Walker, Molecular cancer research : MCR 2012, 10, 546.
- 21 [13] N. Kikuno, H. Shiina, S. Urakami, K. Kawamoto, H. Hirata, Y. Tanaka, S. Majid, M.
22 Igawa, R. Dahiya, International journal of cancer 2008, 123, 552.
- 23 [14] M. Fang, D. Chen, C. S. Yang, The Journal of nutrition 2007, 137, 223S.

- 1 [15] Y. Li, S. M. Meeran, S. N. Patel, H. Chen, T. M. Hardy, T. O. Tollefsbol, *Molecular cancer*
2 2013, 12, 9; Y. Li, H. Chen, T. M. Hardy, T. O. Tollefsbol, *PLoS One* 2013, 8, e54369.
- 3 [16] B. Dutta, S. S. Adav, C. G. Koh, S. K. Lim, E. Meshorer, S. K. Sze, *Journal of proteomics*
4 2012, 75, 5493.
- 5 [17] B. Dutta, Y. Ren, P. Hao, K. H. Sim, E. Cheow, S. Adav, J. P. Tam, S. K. Sze, *Molecular*
6 *& cellular proteomics : MCP* 2014, 13, 2183; B. Dutta, R. Yan, S. K. Lim, J. P. Tam, S. K. Sze,
7 *Molecular & cellular proteomics : MCP* 2014, 13, 3236.
- 8 [18] M. Y. Tolstorukov, C. G. Sansam, P. Lu, E. C. Koellhoffer, K. C. Helming, B. H. Alver,
9 E. J. Tillman, J. A. Evans, B. G. Wilson, P. J. Park, C. W. Roberts, *Proceedings of the National*
10 *Academy of Sciences of the United States of America* 2013, 110, 10165.
- 11 [19] J. Xiong, Q. Du, Z. Liang, *Oncogene* 2010, 29, 4980.
- 12 [20] K. D. Setchell, L. Zimmer-Nechemias, J. Cai, J. E. Heubi, *Lancet (London, England)* 1997,
13 350, 23.
- 14 [21] P. Han, W. Li, J. Yang, C. Shang, C. H. Lin, W. Cheng, C. T. Hang, H. L. Cheng, C. H.
15 Chen, J. Wong, Y. Xiong, M. Zhao, S. G. Drakos, A. Ghetti, D. Y. Li, D. Bernstein, H. S. Chen,
16 T. Quertermous, C. P. Chang, *Biochimica et biophysica acta* 2016, 1863, 1772; D. M. Silberman,
17 G. B. Acosta, M. A. Zorrilla Zubilete, *Pharmacological research* 2016, 109, 64.
- 18 [22] S. Honda, V. T. Bicocca, J. D. Gessaman, M. R. Rountree, A. Yokoyama, E. Y. Yu, J. M.
19 Selker, E. U. Selker, *Proceedings of the National Academy of Sciences of the United States of*
20 *America* 2016, 113, E6135.
- 21 [23] Y. Cai, E. J. Geutjes, K. de Lint, P. Roepman, L. Bruurs, L. R. Yu, W. Wang, J. van
22 Blijswijk, H. Mohammad, I. de Rink, R. Bernardis, S. B. Baylin, *Oncogene* 2014, 33, 2157; M.
23 Bantscheff, C. Hopf, M. M. Savitski, A. Dittmann, P. Grandi, A. M. Michon, J. Schlegl, Y.

- 1 Abraham, I. Becher, G. Bergamini, M. Boesche, M. Delling, B. Dumpelfeld, D. Eberhard, C.
- 2 Huthmacher, T. Mathieson, D. Poeckel, V. Reader, K. Strunk, G. Sweetman, U. Kruse, G.
- 3 Neubauer, N. G. Ramsden, G. Drewes, *Nature biotechnology* 2011, 29, 255.
- 4 [24] L. Ellis, P. W. Atadja, R. W. Johnstone, *Molecular cancer therapeutics* 2009, 8, 1409.
- 5 [25] U. Singh, M. A. Malik, S. Goswami, S. Shukla, J. Kaur, *Tumour biology : the journal of*
- 6 *the International Society for Oncodevelopmental Biology and Medicine* 2016, 37, 14427.
- 7 [26] N. Dyson, *Genes & development* 1998, 12, 2245.
- 8 [27] M. S. Tsuboy, J. C. Marcarini, A. O. de Souza, N. A. de Paula, D. J. Dorta, M. S.
- 9 Mantovani, L. R. Ribeiro, *Journal of medicinal food* 2014, 17, 218.
- 10 [28] C. Y. Hsieh, R. C. Santell, S. Z. Haslam, W. G. Helferich, *Cancer research* 1998, 58, 3833.
- 11 [29] V. S. Ratnu, M. R. Emami, T. W. Bredy, *Journal of neuroscience research* 2017, 95, 301;
- 12 A. V. Whipple, L. M. Holeski, *Frontiers in genetics* 2016, 7, 189.
- 13 [30] H. P. Voon, J. R. Hughes, C. Rode, I. A. De La Rosa-Velazquez, T. Jenuwein, R. Feil, D.
- 14 R. Higgs, R. J. Gibbons, *Cell reports* 2015, 11, 405.
- 15 [31] S. J. Elsasser, K. M. Noh, N. Diaz, C. D. Allis, L. A. Banaszynski, *Nature* 2015, 522, 240;
- 16 A. D. Goldberg, L. A. Banaszynski, K. M. Noh, P. W. Lewis, S. J. Elsaesser, S. Stadler, S. Dewell,
- 17 M. Law, X. Guo, X. Li, D. Wen, A. Chapgier, R. C. DeKever, J. C. Miller, Y. L. Lee, E. A.
- 18 Boydston, M. C. Holmes, P. D. Gregory, J. M. Greally, S. Rafii, C. Yang, P. J. Scambler, D.
- 19 Garrick, R. J. Gibbons, D. R. Higgs, I. M. Cristea, F. D. Urnov, D. Zheng, C. D. Allis, *Cell* 2010,
- 20 140, 678.
- 21 [32] M. J. Gamble, W. L. Kraus, *Cell cycle (Georgetown, Tex.)* 2010, 9, 2568.
- 22 [33] R. Benetti, S. Gonzalo, I. Jaco, G. Schotta, P. Klatt, T. Jenuwein, M. A. Blasco, *The Journal*
- 23 *of cell biology* 2007, 178, 925; A. H. Peters, D. O'Carroll, H. Scherthan, K. Mechtler, S. Sauer, C.

- 1 Schofer, K. Weipoltshammer, M. Pagani, M. Lachner, A. Kohlmaier, S. Opravil, M. Doyle, M.
2 Sibilio, T. Jenuwein, *Cell* 2001, 107, 323.
- 3 [34] J. H. Nijwenning, E. J. Geutjes, R. Bernards, R. L. Beijersbergen, *PLoS One* 2011, 6,
4 e25235.
- 5 [35] Y. F. Zhang, A. R. Zhang, B. C. Zhang, Z. G. Rao, J. F. Gao, M. H. Lv, Y. Y. Wu, S. M.
6 Wang, R. Q. Wang, D. C. Fang, *Molecular biology reports* 2013, 40, 1711.
- 7 [36] J. A. Vizcaino, E. W. Deutsch, R. Wang, A. Csordas, F. Reisinger, D. Rios, J. A. Dienes,
8 Z. Sun, T. Farrah, N. Bandeira, P. A. Binz, I. Xenarios, M. Eisenacher, G. Mayer, L. Gatto, A.
9 Campos, R. J. Chalkley, H. J. Kraus, J. P. Albar, S. Martinez-Bartolome, R. Apweiler, G. S.
10 Omenn, L. Martens, A. R. Jones, H. Hermjakob, *Nature biotechnology* 2014, 32, 223.

11

Figures Section

Figure Legends

Figure 1: Schematic representation of experimental strategy.

Figure 2: a) Heat-map representing the chromatin binding topology of chromatin organizing proteins according to heterochromatin/euchromatin (H/E) association ratio. Colors indicate binding shift towards regions of heterochromatin (red) or euchromatin (black) or unchanged binding topology (green). Chromatin samples were extracted from MDA-MB-231 cells cultured under control (C), treatment (T) or genistein withdrawal (W) conditions followed by separation of proteins bound to euchromatin (supernatant fraction) or heterochromatin (undigested pellet fraction) using partial MNase digestion. Genistein exposure increase the euchromatin (b) and heterochromatin (c) binding topology of epigenetic regulatory proteins in an estrogen receptor-independent fashion. Heterochromatin/euchromatin (H/E) ratio was calculated using the TMT ratio of relative protein abundance in the pellet and supernatant fractions obtained from estrogen receptor-deficient MDA-MB-231 cells cultured under control (C), treatment (T) or genistein withdrawal (W) conditions.

Figure 3: Prolonged genistein treatment modifies higher-order chromatin structure. a) MNase finger-printing of chromatin obtained from MDA-MB-231 cells cultured under control (C), treatment (T) or genistein withdrawal (W) conditions. Chromatin samples were digested using varying concentrations of MNase (0, 2.5, 5 and 10U) and separated on 1% agarose gels. Chromatins of control MDA-MB-231 cells generated higher amount of small DNA fragments after high dose MNase treatment compare to the treated and withdrawal samples. b-c) Western blot analysis for nonchromatin-associated Histone H4 in control (C), treatment (T) or genistein

1 withdrawal (W) MDA-MB-231 (b) and relative abundance of nonchromatin-associated Histone
2 H4 in different experimental condition (c). Statistical significance was determined using
3 experimental triplicates. P-value $* < 0.05$.

4 **Figure 4:** a) differential chromatin binding of RB1, RBL1 and E2F1 in MDA-MB-231 cells
5 cultured under control (C), treatment (T) or genistein withdrawal (W) conditions b) Chromatin
6 binding topology of RB1, RBL1 and E2F1 in cells cultured under different experimental
7 conditions. c). Proportion of S-phase cells detected under different experimental conditions.
8 Proliferation of MDA-MB-231 and MCF7 cells as assessed by MTT assay (d) and cell viability
9 was measured by clonogenic assay (e-f) after culture under control, treatment, or genistein
10 withdrawal conditions. Statistical significance was determined using experimental triplicates. P-
11 value $** < 0.005$, $*** < 0.0005$, $**** < 0.00005$ and $***** < 0.000005$.

12 **Figure 5:** Genistein downregulates mRNA expression of c-MYC and c-MYC regulated E-Box-
13 related genes. Expression of target genes was detected by quantitative real-time PCR and
14 normalized with ribosomal 18S mRNA expression in MDA-MB-231 and MCF7 cells cultured
15 under control (C), treatment (T) or genistein withdrawal (W) conditions. Data are shown as mean
16 values with error bars representing $\pm$ SEM, calculated from biological triplicates

Figures

Figure 1:

1   **Figure 2:**

1 **Figure 3:**

2

3

4

Figure 4:

Figure 5: