Gene 580 (2016) 159–168

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Research paper

Artificial neural network-based exploration of gene-nutrient interactions in folate and xenobiotic metabolic pathways that modulate susceptibility to breast cancer☆

Shaik Mohammad Naushad [a,*], M. Janaki Ramaiah [a], Manickam Pavithrakumari [a], Jaganathan Jayapriya [a], Tajamul Hussain [b], Salman A. Alrokayan [c], Suryanarayana Raju Gottumukkala [d], Raghunadharao Digumarti [e], Vijay Kumar Kutala [f]

[a] School of Chemical & Biotechnology, SASTRA University, Tirumalaisamudram, Thanjavur 613401, India
[b] Center of Excellence in Biotechnology Research, King Saud University, PO Box 2455, Riyadh 11451, Saudi Arabia
[c] Department of Biochemistry, College of Science, King Saud University, PO Box 2455, Riyadh 11451, Saudi Arabia
[d] Department of Surgical Oncology, Nizam's Institute of Medical Sciences, Hyderabad -500082, India
[e] Department of Medical Oncology, Nizam's Institute of Medical Sciences, Hyderabad 500082, India
[f] Department of Clinical Pharmacology and Therapeutics, Nizam's Institute of Medical Sciences, Hyderabad 500082, India

ARTICLE INFO

Article history:
Received 27 November 2015
Received in revised form 1 January 2016
Accepted 12 January 2016
Available online 16 January 2016

Keywords:
Breast cancer
Folate pathway
Xenobiotic pathway
Artificial neural network
Methylome

ABSTRACT

In the current study, an artificial neural network (ANN)-based breast cancer prediction model was developed from the data of folate and xenobiotic pathway genetic polymorphisms along with the nutritional and demographic variables to investigate how micronutrients modulate susceptibility to breast cancer. The developed ANN model explained 94.2% variability in breast cancer prediction. Fixed effect models of folate (400 μg/day) and B_{12} (6 μg/day) showed 33.3% and 11.3% risk reduction, respectively. Multifactor dimensionality reduction analysis showed the following interactions in responders to folate: RFC1 G80A × MTHFR C677T (primary), COMT H108L × CYP1A1 m2 (secondary), MTR A2756G (tertiary). The interactions among responders to B_{12} were RFC1G80A × cSHMT C1420T and CYP1A1 m2 × CYP1A1 m4. ANN simulations revealed that increased folate might restore ER and PR expression and reduce the promoter CpG island methylation of extra cellular superoxide dismutase and BRCA1. Dietary intake of folate appears to confer protection against breast cancer through its modulating effects on ER and PR expression and methylation of EC-SOD and BRCA1.

The etiology of breast cancer is complex involving several physiological, genetic, environmental, and epigenetic factors (Petrakis, 1977; McPherson et al., 2000). Given the high incidence of breast cancer in women, specific emphasis was given on the estrogen metabolism as catechol estrogens were reported to form adducts with DNA, which may contribute to induce mutagenicity (Cavalieri and Rogan, 2011). On the other hand, methoxy estrogens were found to be protective (Dawling et al., 2003). The phase I enzymes belong to cytochrome P450 super family and convert estrogens to catechol estrogens (Hachey et al., 2003). Certain genetic variants of CYP1A1, i.e. CYP1A1 m1, CYP1A1 m4 were reported to induce high expression of CYP1A1 leading to increased catechol estrogen production (Naushad et al., 2011a). The phase II enzymes detoxify catechol estrogens by the following mechanisms: i) O-methylation of catechol estrogens catalyzed by catechol-o-amine methyl transferase (COMT); ii) conjugation of semiquinones/quinones (formed by catechol estrogens) with glutathione in the presence of GSTs (Chen et al., 2004). The COMT H108L and GSTT1/GSTM1 null variants hamper this detoxification process thus exerting the breast cancer risk (Naushad et al., 2011b).

The methylation of catechol estrogens further depends on the bioavailability of S-adenosylmethionine (SAM) (JE1 et al., 2002), a byproduct of one-carbon metabolism (Inoue-Choi et al., 2012). The dietary folate in the form of folylpolyglutamate enters the intestine and undergoes hydrolysis to form folylmonoglutamate by the action of glutamate carboxypeptidase II (GCPII) and thus gets absorbed by the intestine. Folate reductase catalyzes the two-step reduction of folate to form dihydrofolate (DHF) and tetrahydrofolate (THF). The THF from the plasma is transported to RBC with the help of reduced folate

Abbreviations: GCPII, Glutamate carboxypeptidase II; RFC1, Reduced folate carrie 1; cSHMT, Cytosolic serine hydroxymethyl transferase; TYMS, Thymidylate synthase; MTHFR, 5,10-methylene tetrahydrofolate reductase; MTR, Methionine synthase; MTRR, Methionine synthase reductase; SAM, S-adenosyl methionine; SAH, S-adenosyl homocysteine; CYP, Cytochrome P450; COMT, Catechol-o-amine methyl transferase; GST, Glutathione-S-transferase; RASSF1, Ras association (RalGDS/AF-6) domain family member 1; BRCA1, Breast cancer 1,early onset; RARB1, Retinoic acid receptor, beta; EC-SOD, Extracellular superoxide dismutase; ANN, Artificial neural network; MDR, Multifactor dimensionality reduction analysis.

☆ Conflicts of interest statement: All the authors hereby declare no conflicts of interest.

* Corresponding author at: School of Chemical and Biotechnology, SASTRA University, Tirumalaisamudram, Thanjavur 613401, India.
E-mail address: naushadsm@gmail.com (S.M. Naushad).

http://dx.doi.org/10.1016/j.gene.2016.01.023
0378-1119/

carrier 1 (RFC1). THF accepts methylene moiety from serine and forms 5,10-methylene THF in the presence of cytosolic serine hydroxymethyltransferase (cSHMT). 5,10-methylene THF is the substrate for two enzymes, i.e. thymidylate synthase (TYMS) and methylene tetrahydrofolatereductase (MTHFR), which catalyze the conversion of dUMP to dTMP, and FAD-dependent reduction of 5,10-methylene THF to 5-methyl THF. 5-methyl THF remethylates homocysteine to methionine in the presence of methionine synthase/methionine synthase reductase (MTR/MTRR) holoenzyme complex. Methionine is the precursor for the synthesis of SAM, which is a universal methyl donor that donates methyl group to DNA, catecholamines and proteins. Upon donating methyl group, SAM is converted to S-adenosyl homocysteine (SAH), which gives back homocysteine through hydrolysis (Naushad et al., 2011a).

Thiamine and cofactors of folate pathway, namely folate, riboflavin, and vitamin B6, were reported to confer protection against breast cancer (Cancarini et al., 2015). Several functional polymorphisms in folate pathway have been investigated by various researchers for their possible association with breast cancer (Stevens et al., 2007). In our earlier study, we reported positive association of RFC1 G80A and MTRR A66G and inverse association of cSHMT C1420T with breast cancer (Mohammad et al., 2011). We have also showed the cross-talk between the folate and xenobiotic metabolic pathways modulating the breast cancer risk (Naushad et al., 2011b). The folate pathway also plays a pivotal role in DNA methylation and thus any deregulation might induce hypermethylation of tumor suppressors and hypomethylation of proto-oncogenes, which are hallmarks of cancer (Naushad et al., 2012a).

Dietary folate and cobalamin intake were shown to exhibit inverse association with methylation of breast cancer 1,early onset (BRCA1), and retinoic acid receptor, beta (RARB1) (Pirouzpanah et al., 2015). Breast cancer patients with plasma folate levels in the highest tertile were reported to have less risk for mortality compared to those with plasma folate in the lowest tertile (McEligot et al., 2015). Mediterranean diet rich in cofactors of folate pathway was shown to reduce breast cancer risk in subjects with MTHFR 677 T and MTR 2756 A variant alleles (Kakkoura et al., 2015). Higher dietary folate intake was shown to reduce risk for the ER-negative breast cancer in pre-menopausal women (de Batlle et al., 2014). BRCA1 methylation in all types of breast cancers and Ras association (RalGDS/AF-6) domain family member 1 (RASSF1) methylation in the ER/PR-negative breast cancers was reported to correlate positively with total plasma homocysteine (Naushad et al., 2014).

In the current study, we have aimed to develop a risk prediction model for breast cancer by incorporating demographic data, family history, dietary intake of folate, B_2, B_6, B_{12} along with data on the fourteen polymorphisms of folate and xenobiotic metabolic pathways. The rationale of this risk prediction model was to assess the influence of life style modulation in bringing down breast cancer risk, specifically by modulating micronutrient intake. In parallel, we have studied folate-mediated changes in methylome at RASSF1, BRCA1, (BCL2/adenovirus E1B 19 kDa interacting protein 3 (BNIP3), extracellular superoxide dismutase (EC-SOD) loci.

1. Materials and methods

1.1. Sample size calculation

Based on our previous study, the difference in means of plasma folate levels between controls vs. basal-like breast cancer was 0.62, standard deviation was 1.2, the ratio of controls to basal-like breast cancer was 5.0. Hence, 36 basal-like breast cancers and 240 healthy controls were required to obtain 80% power with type I α error of 0.05. The incidence of basal-like breast cancer in India was reported to vary from 12.5% to 29.8%. Hence, assuming the incidence to be around 12.5%, the required number of total breast cancer samples was calculated as 288.

1.2. Recruitment of subjects

We have conducted a case–control study by recruiting 342 breast cancer patients and 253 normal healthy controls in the Departments of Medical and Surgical Oncology, Nizam's Institute of Medical Sciences, Hyderabad, India, during the period of June 2009 to June 2012. The diagnosis of breast cancer was based on the mammogram and histopathological examination of the biopsy. The inclusion criteria were i) patients aged between 18 and 70 yr. with confirmed diagnosis of breast cancer based on mammography and histopathological examination; ii) controls matched with cases in terms age, ethnicity, and geographical location with no history of any benign or malignant breast disease; and iii) subjects willing to give informed consent. The exclusion criteria were i) patients with any co-morbid disorder or malignancy; ii) cases already under radiation and chemotherapy will be excluded from methylation studies; and iii) patients whose medical records are not accessible. The Institutional Ethical Committee of Nizam's Institute of Medical science (NIMS), Hyderabad India, has approved the study protocol (EC/NIMS/767/2007, dated 05.09.2008). The informed consent was obtained from all the subjects.

1.3. Measurements

From all the subjects, the demographic characteristics such as age (yr), body mass index (BMI, kg/m^2), age of menarche (yr), parity, and the menopausal status (pre-/post-menopausal) were recorded during personal interviews conducted by a team of trained researchers. To calculate body mass index, height and weight were recorded to the nearest measurement of 0.1 cm and 0.1 kg, respectively. The estrogen exposure time was calculated based on the age of menarche and age of menopause (post-menopausal)/age at the time of sample collection (pre-menopausal).

1.4. Dietary assessment

All the participants were asked to complete a dietary record of all the food items consumed and their quantity for 4-day period from Thursday to Sunday as this method is validated and being followed by National Diet and Nutrition Survey in the United Kingdom. The data obtained were segregated into white vegetables, green vegetables, leafy vegetables, fruits, milk products, and non-vegetarian foods. The dietary intake of folate, vitamins B_2, B_6, and B_{12} was assessed based on these diaries. The nutritive value of Indian foods (Gopalan et al., 1989), McCance and Widdowson's the composition of foods (Krebs, 2002), and the United States Department of Agriculture's National Nutrient Database for Standard Reference release 18 (USDA, Washington, DC, USA) (U.S. Department of Agriculture and Agricultural Research Service., 2006) were referred to calculate micronutrient quantity per food item. Average daily nutrient intakes were calculated as grams of food multiplied by the amount of each micronutrient in the food and the frequency of consumption, summing over all the foods consumed. None of the subjects enrolled in this study were on any vitamin supplements.

1.5. Immunohistochemistry

Immunohistochemistry for ER, PR, HER2/Neu was performed on the serial sections of paraffin-embedded breast cancer tissues using the standard streptavidin–biotin complex method with 3, 3′-diaminobenzidine as the chromogen. ER antibody (Clone SP1, Lab Vision) was used at 1:250 dilution in 10 mM citrate buffer (pH 6.0) with an 8-min microwave antigen retrieval. PR antibody (Clone 1E2, Ventana) was used as per the Ventana automated stainer standard CC1 protocol. HER2 antibody (Clone SP3, Lab Vision) was used at 1:100 dilution in 0.05 M Tris buffer (pH 10.0) with heat-induced antigen retrieval at 95 °C for 30 min.

Biomarker expression from immunohistochemistry assays was scored by two pathologists, who were blinded to the clinicopathological

characteristics. If immunostaining was observed in more than 1% of tumor nuclei, the tumor was considered positive for ER or PR. If immunostaining was scored as 3+ according to HercepTest criteria, the tumor was considered positive for HER2.

1.6. Genetic analysis

Whole blood in EDTA vacutainer and tumor biopsies were collected from the breast cancer patients and only blood samples were collected from the controls. DNA was isolated using proteinase K digestion followed by phenol–chloroform extraction from these samples. PCR-RFLP approach was used for the analysis of GCPII C1561T, RFC1 G80A, cSHMT C1420T, TYMS 3′-UTR ins6/del6, MTHFR C677T, MTR A2756G, MTRR A66G, CYP1A1m1, CYP1A1m2, CYP1A1m4, and COMT H108L polymorphisms. PCR-AFLP method was used for the analysis of TYMS 5′-UTR 28 bp tandem repeat. Multiplex PCR was used to detect GSTT1 and GSTM1 null variants. The detailed protocols of all this analysis were described elsewhere (Mohammad et al., 2011; Naushad et al., 2014). In order to ensure quality in genotyping, all restriction digestion experiments were accompanied by positive and negative controls that ensure quality of digestion. The genotyping was repeated in 70 random samples and 100% concordance in genotyping was obtained. (Supplementary Table 1).

1.7. Methylome analysis

The linear sequence of candidate genes (including the promoter region) was retrieved from Fasta (NCBI database). Using CpG island finder software (http://dbcat.cgm.ntu.edu.tw/), CpG motifs were identified in target genes (Fig 1). By using methylation primer designing software (http://www.urogene.org/cgi-bin/methprimer/methprimer.cgi), specific primers were designed for melting curve methylation analysis and bisulfite-sequencing. The genomic DNA isolated from tissues was denatured for 10 min in 2 M NaOH at 37 °C before the addition of 30 μl of 10 mM hydroquinone (SIGMA) and 520 μl of 3 M sodium bisulphite (pH 5.0). The mixture was incubated for 16 h at 50 °C. The resultant modified DNA was purified using a Wizard DNA Purification System (Promega, Madison, WI, USA), after which it was again treated with NaOH and precipitated. Finally, the DNA precipitate was resuspended in 20 μl of millipore water and stored at −20 °C until used. Methylation-specific PCR (MSP) was performed to study promoter methylation of RASSF1, BRCA1, and EC-SOD. Combined bisulfite restriction analysis (COBRA) was performed to study promoter methylation of BNIP3 (Naushad et al., 2011b; Naushad et al., 2014; Naushad et al., 2012b). (Supplementary Table 2)

1.8. Estradiol estimation

For the estimation of estradiol in serum, the commercially available ELISA KIT (DRG Diagnostics, Marburg, Germany) was used. The assay was performed as per the manufacturer's instructions.

1.9. Development of artificial neural network (ANN)

The ANN model was developed using the computation tool "MATLAB (R2013a)." To develop this model, 26 input variables and 1 target variable were used. The input variables were plasma estradiol (pg/ml), estrogen exposure time (yr), age (yr), BMI (kg/m^2), age of menarche (yr), parity, menopausal status (pre/post), family history of breast cancer (yes/no), GCPII C1561T, RFC1 G80A, cSHMT C1420T, TYMS 5′UTR 28 bp tandem repeat, TYMS 3′UTR ins6/del6, MTHFR C677T, MTR A2756G, MTRR A66G, COMT H108L, CYP1A1m1, CYP1A1 m2, CYP1A1 m4, GSTT1, GSTM1, folate, B_2, B_6, B_{12}. The target variable represented the presence or absence of the disease.

Neural network pattern recognition tool of MATLAB was used to generate a two-layered feed forward network with sigmoid hidden neurons and linear output neurons. The number of hidden layers optimized to be 10. For training the ANN, 75% of the data were used and for validation and testing, 15% of the data each was used. Training was based on the conjugate gradient back propagation. The receiver operating characteristic (ROC) curves and confusion matrix plots were generated indicating the performance of the ANN model.

Neural network clustering tool was used to ascertain the potential of four gene methylome in classifying the breast cancer cases. The data were arranged based on immunohistochemical phenotypes and self-organizing map (SOM) batch algorithm was used to train the network. The data were presented in the form of SOM weight planes.

1.10. Multifactor dimensionality reduction (MDR) analysis

All the genetic polymorphisms were considered as input variables and computed as ×1, ×2, ×3, and the presence or absence of the disease was used as a class. MDR analysis was performed using computational platform www.multifactordimensionalityreduction.org. This platform strategically identified the most important genetic variables and explored bivariate and trivariate interactions. Furthermore, Frutcherman–Rheingold plots were obtained that identify epistatic and hypostatic alleles based on entropy values. The strength of interaction between the variables measured in terms of color scale and entropy, wherein blue represents redundancy and gradation upto red (BGYOR) showing increasing strength of interaction.

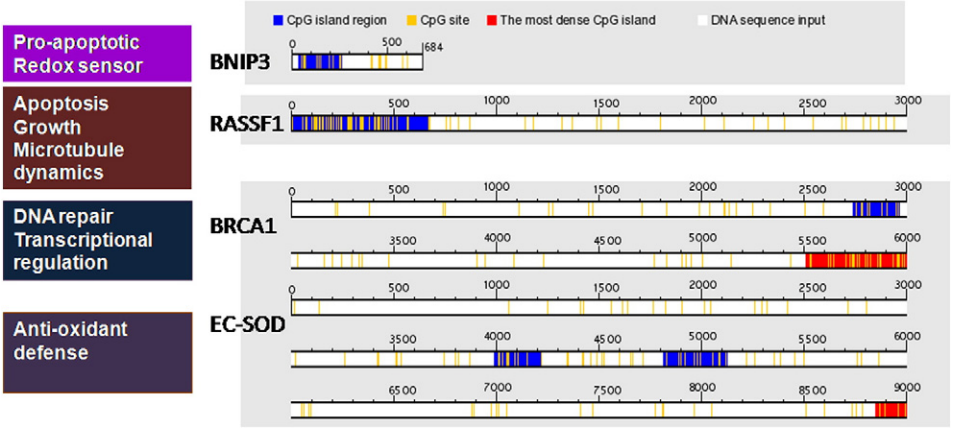


Fig. 1. *CpG island methylome mapping.* The promoter sequence of RASSF1A, BRCA1, BNIP3, and EC-SOD was computed using "http://dbcat.cgm.ntu.edu.tw/" to identify dense CpG islands. The most dense CpG islands were represented in red. Other CpG islands were represented in blue while the CpG sites were represented in yellow.

Table 1
Distribution of continuous variables in breast cancer cases and controls.

Input variable	Mean ± SD		P value
	Cases	Controls	
Age (yr)	51.34 ± 13.20	49.92 ± 12.04	0.17
Body mass index (kg/m^2)	26.25 ± 6.23	25.38 ± 4.81	0.06
Age of menarche (yr)	13.24 ± 1.26	13.18 ± 1.37	0.58
Parity	2.57 ± 1.51	2.65 ± 1.58	0.53
Estradiol (pg/ml)	124.96 ± 32.52	118.25 ± 21.72	<**0.005**
Estrogen exposure time (yr)	38.10 ± 13.29	35.73 ± 12.10	**0.03**
Folate intake (μg/day)	341.56 ± 110.45	377.25 ± 140.97	**0.0006**
B2 intake (μg/day)	1.27 ± 0.42	1.22 ± 0.46	0.17
B6 intake (μg/day)	1.51 ± 0.36	1.47 ± 0.40	0.20
B_{12} intake (μg/day)	5.96 ± 1.51	5.94 ± 1.26	0.86

SD: standard deviation; $p < 0.05$: statistically significant. Bold entries represent statistical significance.

1.11. Statistical analysis

Fisher exact test was performed for the univariate analysis by computing the genetic data in 2×2 contingency table based on the presence or absence of variable in cases and controls. Odds ratio (OR) and 95% confidence interval (CI) and phi coefficient were obtained by this analysis. In order to minimize the confounding effects of other risk factors, multiple logistic regression analysis was performed, which demonstrated independent genetic effects in terms of adjusted odds ratios and 95% CI. Student t-test was performed to analyze the distribution of continuous variables in cases and controls. For all the analyses, only data available were used without considering any missing values.

1.12. Results

As shown in the *flowchart 1*, out of 342 cases and 253 controls, DNA samples extracted from whole blood samples were available from all the subjects and were used for genetic analysis. The DNA extracted from tissues was adequate for methylome analysis in 300 samples. Only 258 cases and 245 controls responded with 4-day food frequency dietary records. Paraffin-embedded tissue blocks for immunohistochemistry were available only in 262 breast cancer cases.

1.13. Univariate analysis

The mean age of breast cancer cases and controls were 51.34 ± 13.20 yr. and 49.92 ± 12.04 yr., respectively. The body mass index of breast cancer cases was higher than controls, but not statistically significant. Breast cancer cases exhibited higher estradiol levels and increased exposure time to estrogens than the controls. Breast cancer cases had lower dietary intake of folate as compared to healthy controls. (Table 1).

Based on the ER, PR, and HER2/Neu, breast cancer cases were classified as luminal A (ER and PR positive and HER2 negative): 92, luminal B (ER/PR weakly positive and HER2 negative): 41, luminA-HH (ER, PR, and HER2 positive): 47, luminB-HH (ER/PR weakly positive and HER2 positive): 21, HER-enriched (ER/PR negative and HER2 positive): 21, and basal-like (ER, PR and HER2 negative): 40.

Table 2 represents the distribution of 14 genetic polymorphisms in breast cancer cases and controls. All the cases and controls showed genotype distribution in accordance with Hardy–Weinberg equilibrium except for MTR A2756G and MTRR A66G polymorphisms in both the groups; and RFC1 G80A and cSHMT C1420T in controls. As shown in Fig 1, out of fourteen polymorphisms, three were protective, seven were associated with risk, and four showed null association with breast cancer. RFC1 G80A, TYMS 5′-UTR 28 bp tandem repeat, MTHFR C677T, COMT H108L, CYP1A1 m1, CYP1A1 m4, GSTT1 null variant were shown to increase risk for breast cancer. GCPII C1561T, cSHMT C1420T, CYP1A1 m2 were found to confer protection. Multiple logistic regression analysis showed no independent protective role of GCPII C1561T and cSHMT C1420T after adjusting for the confounding factors. Null association was observed with TYMS 3′-UTR ins6/del6, MTR A2756G, MTRR A66G, and GSTM1 null variant (Fig 2).

Table 2
Distribution of folate and xenobiotic polymorphisms in breast cancer cases and controls.

Polymorphism	Study group	WW	WM	MM	P_{HWE}	Adjusted OR	95% CI	P value
GCPII C1561T	Cases	297	45	0	0.58	0.64	0.40–1.04	0.07
	Controls	203	50	0	0.31			
RFC1 G80A	Cases	112	162	68	0.80	1.60	1.22–2.10	**0.0006**
	Controls	99	137	17	0.003			
SHMT C1420T	Cases	64	181	97	0.42	0.81	0.61–1.07	0.14
	Controls	24	154	75	0.002			
TYMS 5′ 3R/2R	Cases	136	152	54	0.60	1.35	1.04–1.76	**0.03**
	Controls	111	116	26	0.85			
TYMS 3′ ins6/del6	Cases	87	165	90	0.81	1.02	0.79–1.32	0.87
	Controls	59	136	58	0.45			
MTHFR C677T	Cases	269	69	4	1.00	2.00	1.25–3.22	**0.004**
	Controls	222	30	1	1.00			
MTR A2756G	Cases	165	160	17	0.02	1.19	0.88–1.62	0.26
	Controls	126	117	10	0.03			
MTRR A66G	Cases	93	233	16	0.0001	1.32	0.94–1.85	0.11
	Controls	80	161	12	0.0001			
COMT H108L	Cases	122	154	66	0.41	1.52	1.17–1.96	**0.002**
	Controls	120	107	26	0.96			
CYP1A1 m1	Cases	168	125	49	0.01	1.48	1.13–1.94	**0.005**
	Controls	140	98	15	0.96			
CYP1A1 m2	Cases	221	108	13	1.00	0.63	0.46–0.87	**0.005**
	Controls	144	99	10	0.38			
CYP1A1 m4	Cases	274	68	0	0.16	4.99	2.60–9.56	<**0.0001**
	Controls	240	13	0	1.00			
GSTT1	Cases	255	87	-	0.53	1.64	1.06–2.54	**0.03**
	Controls	205	48	-	0.74			
GSTM1	Cases	228	114	-	0.24	0.86	0.59–1.26	0.44
	Controls	168	85	-	0.34			

WW: wild; WM: heterozygous; MM: homozygous mutant; OR: odds ratio; CI: confidence interval; $p < 0.05$: statistically significant. Bold entries represent statistical significance.

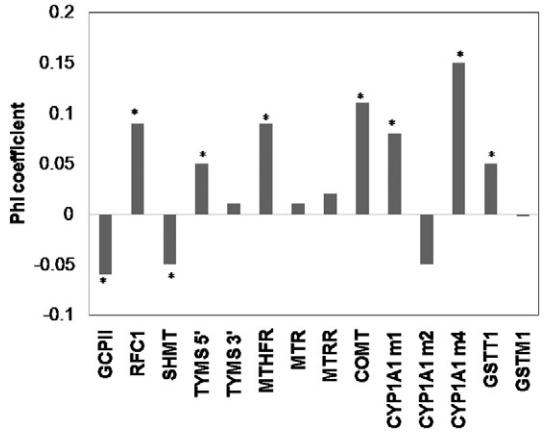


Fig. 2. *Association of genetic polymorphisms with breast cancer risk.* This illustrates the association of fourteen genetic polymorphisms with breast cancer risk in terms of phi coefficients. The "+" and "−"symbols of the phi coefficient are suggestive of positive and inverse association. The digital value is an index of strength of association with breast cancer. RFC1 G80A,TYMS5′ UTR 28 bp tandem repeat, MTHFR C677T, COMT H108L, CYP1A1m1, CYP1A1m4, and GSTT1 null polymorphisms showed positive association while GCPII C1561T and cSHMT C1420T showed inverse association with breast cancer.

1.14. ANN model

The training, validation, and testing datasets showed 100%, 84.6%, and 76.9% accuracy in predicting the breast cancer risk. The whole dataset showed 94.2% accuracy (Figs 3 & 4). In order to assess the beneficial effects of micronutrients, we have developed fixed effect models of folate and B_{12}. Dietary intake of 400 μg folate was shown to reduce risk for breast cancer by 33.3%. The fixed effect model of folate showed benefit in reducing breast cancer risk in the following genotype combinations: RFC1 G80A × MTHFR C677T, COMT H108L × CYP1A1 m2, and MTR A2756G (Fig 5).

The fixed effect model of B_{12} showed 11% risk reduction with 6 μg/day intake. The MDR analysis revealed the possible benefit with B_{12} in following genotype combinations: RFC1 G80A × cSHMT C1420T and CYP1A1 m2 × CYP1A1 m4 (Fig 6).

In order to ensure nutri-genomics potential of this ANN model, we have performed simulations on the data of a 40 yr. old premenopausal woman whose BMI was 28.4 kg/m^2 and was affected with triple negative breast cancer. This woman has multiple risk alleles as shown below: RFC1 80 AA, cSHMT 1420 CT, TYMS 3′-UTR del6/del6, MTR 2756 AG, MTRR 66 AG, COMT 108 HL, CYP1A1 m1/m1, GSTM1 null. Her folate intake at the time of diagnosis was 224 μg/day. The ANN simulations showed restoration of ER and PR expression with increased

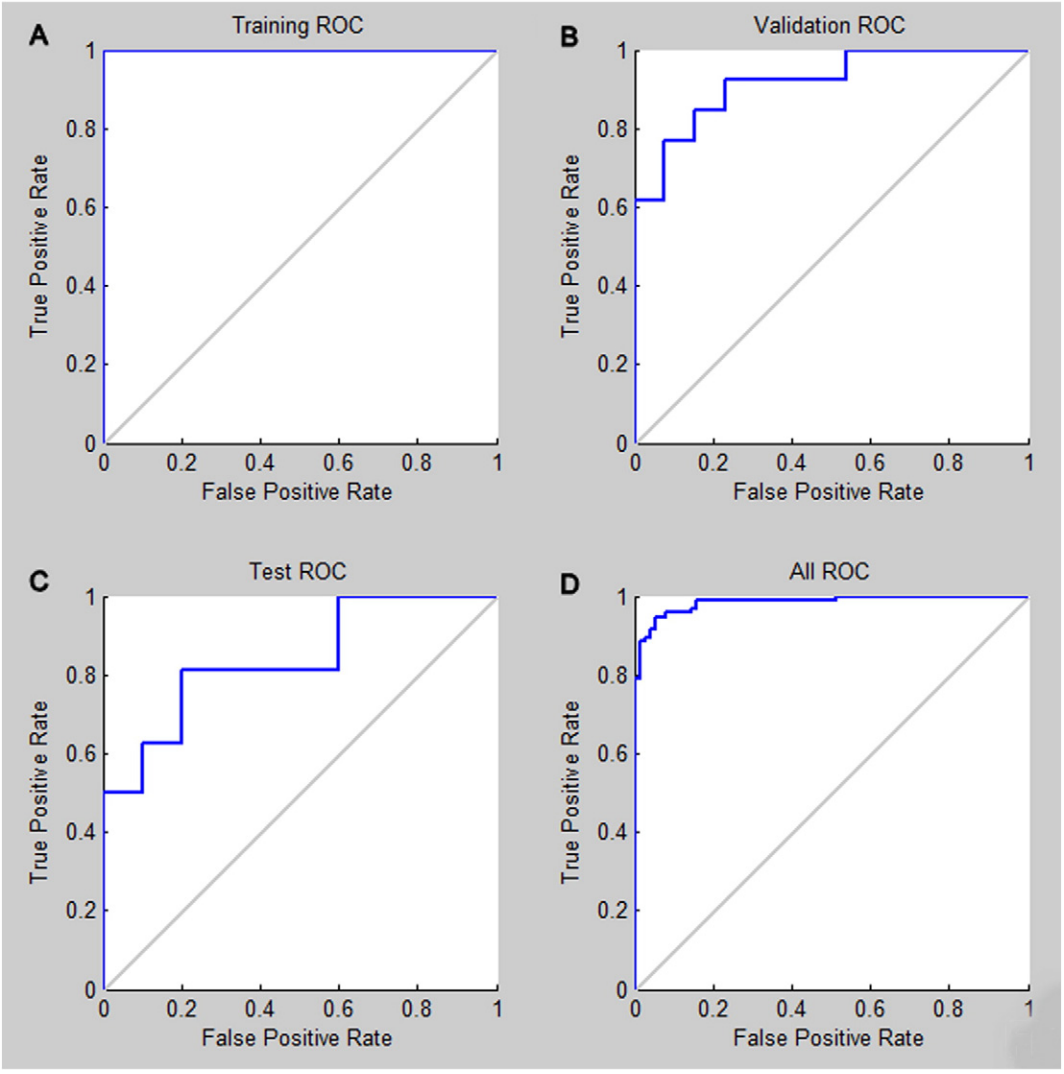


Fig. 3. *Receiver operating characteristic curves of artificial neural network model.* This illustrates the performance of ANN model in terms of true positive vs false positive rates in (A) training, (B) validation, (C) testing, and (D) whole datasets.

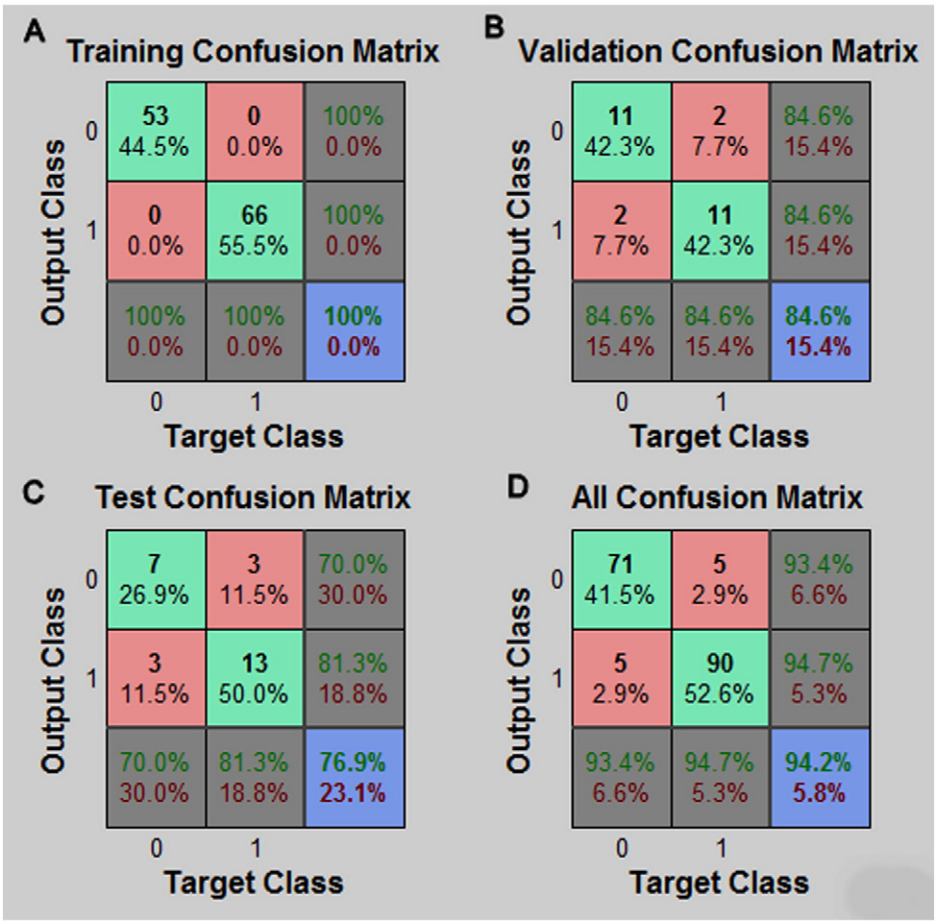


Fig. 4. *Confusion matrix ANN model.* This illustrates confusion matrix of ANN model, which suggests 100%, 84.6%, 76.9%, 94.2% accuracy in (A) training, (B) validation, (C) testing, and (D) whole datasets.

folate. However, no improvement was observed in HER2 expression (Fig 7).

This ANN model was also capable of predicting methylation at four loci, i.e. extracellular superoxide dismutase (EC-SOD), RASSF1A, BRCA1, and BNIP3 with an accuracy of 98.7%, 99.6%, 99.8%, and 99.9%, respectively. The distribution of these four markers was found to be distinct from one another and together they are forming four clusters (Fig 8). Triple negative breast cancer cases, which formed the top nodes of the clustering model, exhibited hypermethylation of RASSF1A and BRCA1 in comparison to other immunohistochemical phenotypes. The pattern of input 1 (RASSF1) showed inverse association with input 4 (BNIP3) methylation. EC-SOD methylation showed no specific association with any immunohistochemical breast cancer phenotype. The ANN simulations showed reversal of CpG methylation in EC-SOD and BRCA1 with increased folate intake (Fig 9).

2. Discussion

In the current study, we have attempted to explore the individual genetic effects and the gene–gene interactions that are likely to be modulated through intake of folate and B_{12} thus contributing to reduced risk of breast cancer. The univariate analysis results justify the rationale of the study by demonstrating higher plasma estradiol, increased exposure time to estrogens, and folate deficiency in the breast cancer patients. Due to complexity of interaction among the physiological, nutritional, and genetic factors, we have employed artificial neural network model to simulate these interactions. The risk prediction model exhibited 94.2% accuracy in the prediction of breast cancer risk. Earlier studies have used the additive genetic model (Kraft et al., 2009; Naushad et al., 2015), MDR model (Naushad et al., 2011b), and recursive partitioning analysis (Xu et al., 2007) to explore gene–gene interactions.

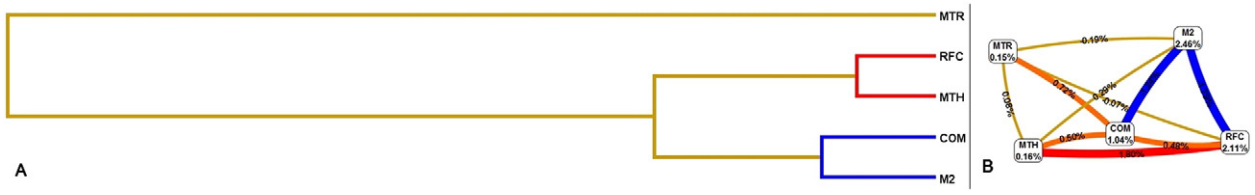


Fig. 5. *Fixed folate model.* The ANN model was used to simulate breast cancer risk pattern when dietary folate intake was considered to be 400 μg/day. Multifactor dimensionality reduction analysis was performed to identify genotype combinations that are likely to be benefited with folate supplementation. (A) The dendrogram revealed first-order interactions between RFC1 and MTHFR, second-order interactions between COMT and CYP1A1 m2, and third-order interaction with MTR. (B) The Frutcherman–Rheingold plot showed RFC1 CYP1A1 m2 as the most important polymorphisms that are benefited by folate. The interactions among RFC1 MTHFR, COMT, and MTR were found to be stronger.

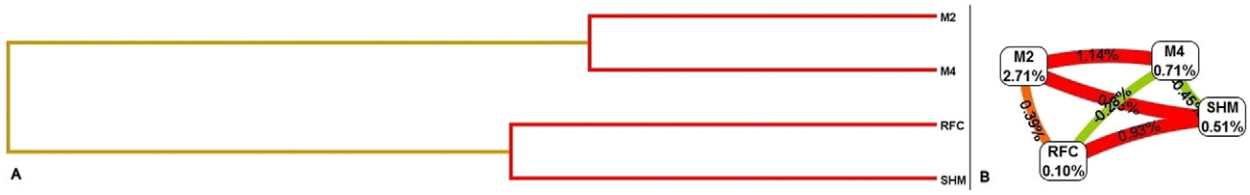


Fig. 6. *Fixed B_{12} model.* The ANN model was simulated to assess breast cancer risk pattern when dietary B_{12} intake fixed to be 6 μg/day. (A) The dendrogram showed that RFC1 × cSHMT and CYP1A1 m2 × CYP1A1 m4 genotype combinations likely to be benefited with B_{12}. (B) The Frutcherman–Rheingold plot showed CYP1A1 m2 as the important predictor with interactions of RFC1 × cSHMT and CYP1A1 m2 × CYP1A1 m4 being stronger.

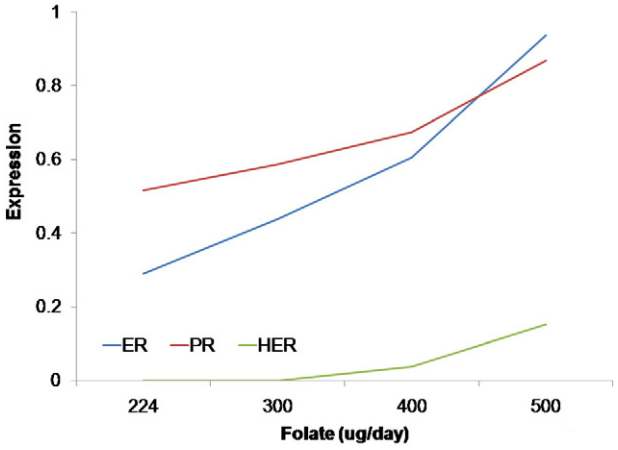


Fig. 7. *Impact of folate intake on ER, PR, HER2/Neu expression.* The ANN model was simulated based on a triple negative breast cancer case whose dietary intake of folate was 224 μg/day. Increased folate was shown to restore ER and PR expression. However, no significant change was observed in HER2/Neu expression.

Earlier, we have reported lower plasma folate levels in the breast cancer patients compared to healthy controls (Naushad et al., 2012a), which substantiated the current association between folate deficiency and breast cancer. Higher intake of folate, vitamin B_6, and vitamin B_{12} was found to nullify the risk for breast cancer associated with MTHFR 677 T and MTR 2756G variant alleles in Chinese population (Jiang-Hua et al., 2014). Although, no direct protective role of plasma vitamin B_{12} was observed in our previous study (Naushad et al., 2014), we have shown inverse association of plasma vitamin B_{12} with CpG methylation of RASSF1 and BRCA1 loci (Naushad et al., 2014).

An ANN model for the breast cancer prediction using demographic risk factors such as age, BMI, age of menarche, estrogen exposure duration, age of first full-term pregnancy, number of children, menopausal status, and family history of breast cancer as input variables showed moderate predictability (AUC = 0.64) (Lee et al., 2015). The current study also highlighted higher plasma estradiol levels, increased exposure to estrogens, and low folate intake as the important risk factors for breast cancer. Another ANN model showed moderate predictability for breast cancer with genetic (AUC = 0.603) or imaging parameters (AUC = 0.693) while the combined model showed improvement in predictability (AUC = 0.731) (Liu et al., 2013). In our current model, the inclusion of nutritional and genetic factors along with the demographic factors enhanced the predictability (AUC = 0.99). Furthermore, an attempt was made to predict ER, PR, and HER2/Neu expressions with the demographic, nutritional, and genetic polymorphisms as the predictors. This model performed well with excellent predictability (AUC = 0.99).

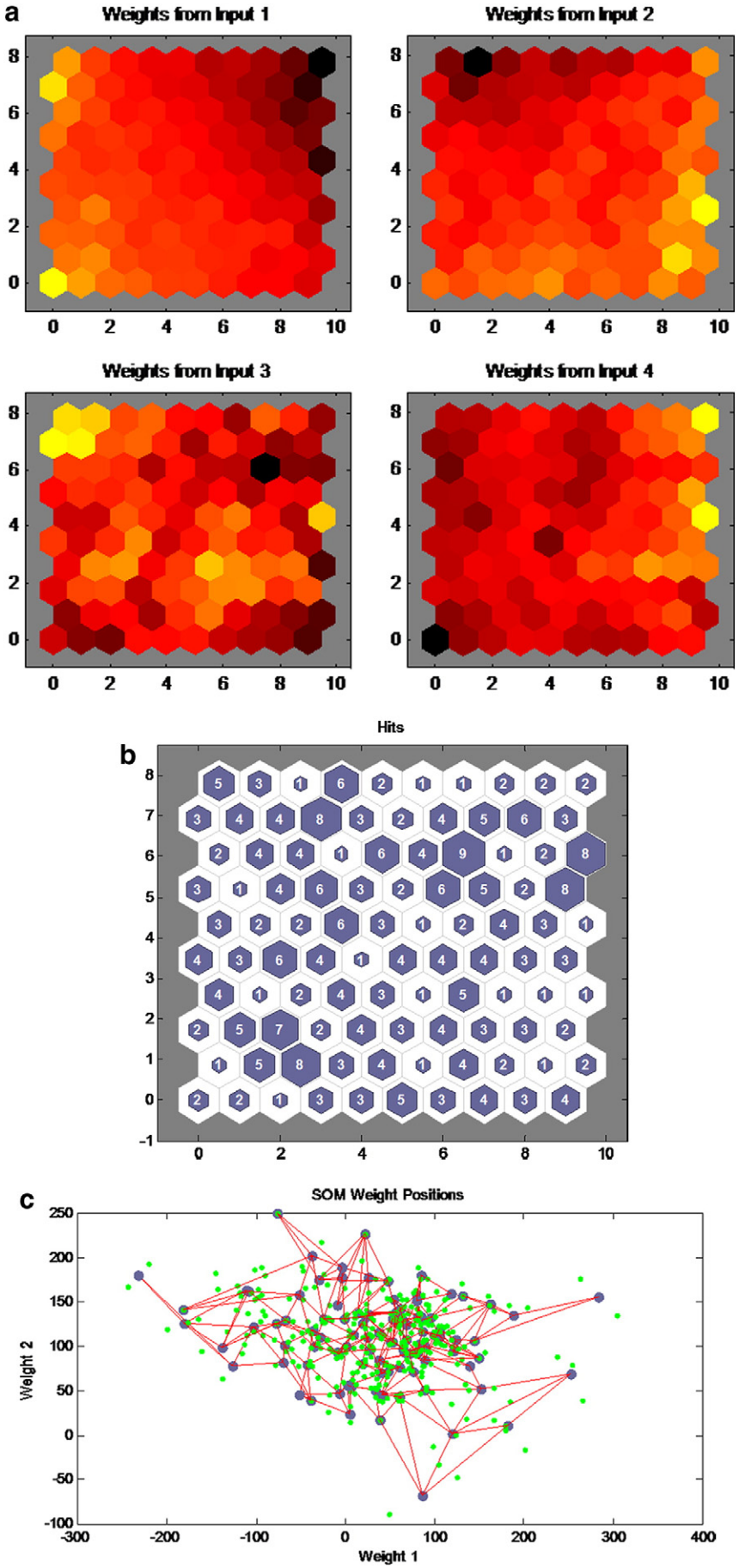
The ANN simulations of the current study demonstrated the beneficial effects of folate in reversing the methylation of BRCA1, which is in agreement with a previous study that showed increased BRCA1 and mRNA expression in HepG2, Huh-7D12, Hs578T, and JURKAT cell lines following treatment with folic acid (Price et al., 2015). The inverse association of folate intake and SOD methylation corroborated with another study that showed beneficial effects of folate and B_{12} in the up-regulation of anti-oxidant defense enzymes (Majumdar et al., 2012). High consumption of fruits and vegetables was reported to confer protection against the breast cancer risk (Karimi et al., 2015). The positive association of folate with ER and PR expression is in agreement with a previous study that showed increased methylation and more ER negativity in subjects having low dietary folate intake (Christensen et al., 2010). Earlier, we have reported lower plasma folate levels in basal-like breast cancer and higher folate levels in Luminal A-HH, which supports the association of ER and PR expression with folate (Naushad et al., 2012a). The current study showing BRCA1 hypermethylation in basal-like breast cancer denotes the link between the folate deficiency and BRCA1 methylation.

The RFC1 G80A polymorphism was reported to be a risk factor for gastro-esophageal cancer (Wang et al., 2006), cleft lip (Vieira et al., 2008), pre-term birth (Wang et al., 2015), and colorectal adenoma (Levine et al., 2011). Shorter telomere length was observed in subjects with RFC1 80 AA-genotype (Milne et al., 2015).

Our results showing association of CYP1A1 m1 with breast cancer risk are in agreement with Li *et al.* (Li et al., 2005). However, this association was not consistent across the different ethnic groups and population (Singh et al., 2007; Okobia et al., 2005; Li et al., 2004; Zhang et al., 2004). Our study explains these discrepancies by highlighting the importance of the gene–gene and gene-nutrient interactions rather than the individual genetic effects.

The beneficial effects of folate in subjects with RFC1 G80A and MTHFR C677T can be explained based on the biological mechanisms. RFC1 is a folate carrier that facilitates transport of folate into RBC. RFC1 G80A polymorphism was reported to be associated with lower circulating folate levels (Naushad et al., 2011c). Under conditions of severe folate deprivation, RFC mRNA levels were reported to decrease 2.5-folds (Ifergan et al., 2008). These two studies support that higher dietary folate may negate RFC1 impairment by inducing RFC1 expression. MTHFR is a rate-limiting enzyme of folate pathway whose activity dictates the direction of flux of folate and when folate availability is adequate, both processes, i.e. thymidine synthesis and remethylation of homocysteine, will not be affected. COMT H108L induces heat labile variant of COMT thus impairing the conversion of catechol estrogen to methoxyestrogen. When folate is adequate, it increases SAM synthesis and thus facilitates more methoxyestrogen production. The beneficial effects of B_{12} in RFC1 G80A and cSHMT C1420T could be due to possible stabilization of MTR/MTRR holoenzyme complex through adequate availability of co-factor methyl cobalamin that increases rate of remethylation thus maintaining adequate SAM.

The major strength of the current study is the utilization of artificial neural network platform to delineate complex gene–gene and gene–nutrient interactions. The current study provides rationale for folate supplementation programs by highlighting the genotypes that are likely to benefit by such programs. The utilization of MDR analysis along with ANN helped in achieving this objective. However, this retrospective study should be confirmed in the future through the prospective studies. Inter-observer variability in gathering the information of the subjects was minimized by using standardized protocols for the data


a
Weights from Input 1
Weights from Input 2
Weights from Input 3
Weights from Input 4
b
Hits
c
SOM Weight Positions
Weight 1
Weight 2

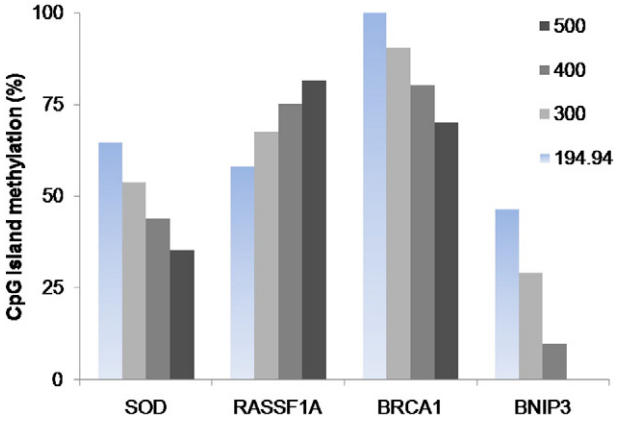


Fig. 9. *Impact of dietary folate intake on CpG island methylation of the promoter.* The ANN model of methylome predicted inverse association of folate with SOD, BRAC1, and BNIP3 methylation. However, RASSF1A methylation showed positive association with folate.

collection and training the personnel. The personnel involved in the analytical aspects of the study were blinded regarding the case/control status of the sample. In order to detect any selection bias, Hardy–Weinberg equilibrium statistics was employed. Multiple logistic regression analysis was performed to adjust for the confounding effects of other variables. The current ANN model can be upgraded by including ethnicity as a variable so as to improve the generalizability of the model. Nevertheless, the previous studies (Pirouzpanah et al., 2015; McEligot et al., 2015; Kakkoura et al., 2015; de Batlle et al., 2014; Naushad et al., 2014) corroborate with our findings by demonstrating folate as an effect modifier, specifically when the risk is mediated through aberrations in folate metabolism.

To conclude, higher plasma estradiol, higher exposure time to estrogen, folate deficiency, and perturbations in folate and xenobiotic pathway may increase the breast cancer risk. Fixed dose effects of folate and B_{12} suggest that increased intake of these two micronutrients might reduce the breast cancer risk by correcting the intracellular folate levels. Folate can influence breast cancer phenotype by modulating ER and PR and also by altering the methylome of SOD and BRCA1.

Acknowledgments

This work was supported by the extra mural research grant of Indian Council of Medical Research (ICMR) (ref no. 5/13/32/2007) bestowed upon Dr. Vijay Kumar Kutala.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.gene.2016.01.023.

References

Cancarini, I., Krogh, V., Agnoli, C., Grioni, S., Matullo, G., Pala, V., Pedraglio, S., Contiero, P., Riva, C., Muti, P., Sieri, S., 2015. Micronutrients involved in one-carbon metabolism and risk of breast cancer subtypes. PLoS One 10 (9), e0138318.

Cavalieri, E.L., Rogan, E.G., 2011. Unbalanced metabolism of endogenous estrogens in the etiology and prevention of human cancer. J. Steroid Biochem. Mol. Biol. 125 (3–5), 169–180.

Chen, Z.H., Hurh, Y.J., Na, H.K., Kim, J.H., Chun, Y.J., Kim, D.H., Kang, K.S., Cho, M.H., Surh, Y.J., 2004. Resveratrol inhibits TCDD-induced expression of CYP1A1 and CYP1B1 and catechol estrogen-mediated oxidative DNA damage in cultured human mammary epithelial cells. Carcinogenesis 25 (10), 2005–2013 (Oct).

Christensen, B.C., Kelsey, K.T., Zheng, S., Houseman, E.A., Marsit, C.J., Wrensch, M.R., Wiemels, J.L., Nelson, H.H., Karagas, M.R., Kushi, L.H., Kwan, M.L., Wiencke, J.K., 2010. Breast cancer DNA methylation profiles are associated with tumor size and alcohol and folate intake. PLoS Genet. 6 (7), e1001043.

Dawling, S., Roodi, N., Parl, F.F., 2003. Methoxyestrogens exert feedback inhibition on cytochrome P450 1 A1 and 1B1. Cancer Res. 63 (12), 3127–3132.

de Batlle, J., Ferrari, P., Chajes, V., JY, P., Slimani, N., McKenzie, F., Overvad, K., Roswall, N., Tjønneland, A., MC, B.-R., Clavel-Chapelon, F., Fagherazzi, G., Katzke, V., Kaaks, R., MM, B., Trichopoulou, A., Lagiou, P., Trichopoulos, D., Palli, D., Sieri, S., Panico, S., Tumino, R., Vineis, P., HB, B.-d.-M., PH, P., Hjartåker, A., Engeset, D., Weiderpass, E., Sánchez, S., Travier, N., MJ, S., Amiano, P., MD, C., Barricarte Gurrea, A., KT, K., TJ, K., KE, B., Ericson, U., Sonestedt, E., Van Guelpen, B., Schneede, J., Riboli, E., I., R., 2014. Dietary folate intake and breast cancer risk: European prospective investigation into cancer and nutrition. J. Natl. Cancer Inst. 107 (1), 367 (Dec 12).

Gopalan, C., Rama Sastri, B.V., Balasubramanian, S.C., 1989. Nutritive Value of Indian Foods. National Institute of Nutrition, Indian Council of Medical Research.

Hachey, D.L., Dawling, S., Roodi, N., Parl, F.F., 2003. Sequential action of phase I and II enzymes cytochrome P450 1B1 and glutathione S-transferase P1 in mammary estrogen metabolism. Cancer Res. 63 (23), 8492–8499.

Ifergan, I1., Jansen, G., YG, A., 2008. The reduced folate carrier (RFC) is cytotoxic to cells under conditions of severe folate deprivation. RFC as a double edged sword in folate homeostasis. J. Biol. Chem. 283 (30), 20687–20695 (Jul 25).

Inoue-Choi, M., Nelson, H.H., Robien, K., Arning, E., Bottiglieri, T., Koh, W.P., Yuan, J.M., 2012. One-carbon metabolism nutrient status and plasma S-adenosylmethionine concentrations in middle-aged and older Chinese in Singapore. Int. J. Mol. Epidemiol. Genet. 3 (2), 160–173.

JE1, G., LT, J., P, H., JD, Y., 2002. Characterization of human soluble high and low activity catechol-O-methyltransferase catalyzed catechol estrogen methylation. Pharmacogenetics 12 (7), 517–528.

Jiang-Hua, Q., De-Chuang, J., Zhen-Duo, L., Shu-de, C., Zhenzhen, L., 2014 Dec. Association of methylenetetrahydrofolate reductase and methionine synthase polymorphisms with breast cancer risk and interaction with folate, vitamin B6, and vitamin B 12 intakes. Tumour Biol. 35 (12), 11895–11901.

Kakkoura, M.G., Demetriou, C.A., Loizidou, M.A., Loucaides, G., Neophytou, I., Marcou, Y., Hadjisavvas, A., Kyriacou, K., 2015. Single-nucleotide polymorphisms in one-carbon metabolism genes, Mediterranean diet and breast cancer risk: a case–control study in the Greek-Cypriot female population. Genes Nutr. 10 (2), 453.

Karimi, Z., Bahadoran, Z., Abedini, S., Houshyar-Rad, A., Rashidkhani, B., 2015. Dietary total antioxidant capacity and the risk of breast cancer: a case–control study. East Mediterr. Health J. 21 (8), 564–571.

Kraft, P., Wacholder, S., Cornelis, M.C., Hu, F.B., Hayes, R.B., Thomas, G., Hoover, R., Hunter, D.J., Chanock, S., 2009. Beyond odds ratios–communicating disease risk based on genetic profiles. Nat. Rev. Genet. 10 (4), 264–269.

Krebs, J., 2002. McCance and Widdowson's the Composition of Foods: Summary Edition, 6th Summary ed. The Royal Society of Chemistry/Food Standards Agency, Cambridge/London.

Lee, C., Lee, J.C., Park, B., Bae, J., Lim, M.H., Kang, D., Yoo, K.Y., Park, S.K., Kim, Y., Kim, S., 2015. Computational discrimination of breast cancer for Korean women based on epidemiologic data only. J. Korean Med. Sci. 30 (8), 1025–1034.

Levine, A.J., Lee, W., Figueiredo, J.C., Conti, D.V., Vandenberg, D.J., Davis, B.D., Edlund, C.K., Henning, S.M., Heber, D., Stern, M.C., Haile, R.W., 2011. Variation in folate pathway genes and distal colorectal adenoma risk: a sigmoidoscopy-based case–control study. Cancer Causes Control 22 (4), 541–552.

Li, Y., Millikan, R.C., Bell, D.A., Cui, L., Tse, C.K., Newman, B., Conway, K., 2004. Cigarette smoking, cytochrome P4501A1 polymorphisms, and breast cancer among African-American and white women. Breast Cancer Res. 6 (4), R460–R473.

Li, Y., Millikan, R.C., Bell, D.A., Cui, L., Tse, C.K., Newman, B., Conway, K., 2005. Polychlorinated biphenyls, cytochrome P450 1 A1 (CYP1A1) polymorphisms, and breast cancer risk among African American women and white women in North Carolina: a population-based case–control study. Breast Cancer Res. 7 (1), R12–R18.

Liu, J., Page, D., Nassif, H., Shavlik, J., Peissig, P., McCarty, C., Onitilo, A.A., Burnside, E., 2013. Genetic variants improve breast cancer risk prediction on mammograms, AMIA Annu Symp Proc. pp. 876–885.

Majumdar, S., Maiti, A., Karmakar, S., Das, A.S., Mukherjee, S., Das, D., Mitra, C., 2012. Antiapoptotic efficacy of folic acid and vitamin B_{12} against arsenic-induced toxicity. Environ. Toxicol. 27 (6), 351–363.

McEligot, A.J., Ziogas, A., Pfeiffer, C.M., Fazili, Z., Anton-Culver, H., 2015. The association between circulating total folate and folate vitamers with overall survival after postmenopausal breast cancer diagnosis. Nutr. Cancer 67 (3), 442–448.

McPherson, K., Steel, C.M., Dixon, J.M., 2000. Breast cancer epidemiology, risk factors, and genetics. BMJ 321 (7261), 624–628.

Milne, E., O'Callaghan, N., Ramankutty, P., de Klerk, N.H., Greenop, K.R., Armstrong, B.K., Miller, M., Fenech, M., 2015. Plasma micronutrient levels and telomere length in children. Nutrition 31 (2), 331–336.

Fig. 8. *Application of neural network clustering tool to identify IHC classes based on methylome.* The methylome data of RASSF1A, BRCA1, EC-SOD, and BNIP3 on different immunohistochemical phenotypes of breast cancer, namely luminal A, luminal B, lumin A-HH, lumin B-HH, HER-enriched, and basal like was subjected to neural network clustering to identify utility of these markers in molecular phenotyping. (A) Self-organizing map (SOM) weight planes; (B) SOM sample hits; (C) SOM weight positions. SOM weight planes show the distribution of RASSF1A (input 1), BRCA1 (input 2), EC-SOD (input 3), and BNIP3 (input 4) in different immunohistochemical phenotypes of breast cancer. These markers although showed clear distinction from one another. SOM sample hits and SOM weight positions suggest four distinct classes in the breast cancer data instead of six groups as expected due to overlapping between certain breast cancer phenotypes.

Mohammad, N.S., Yedluri, R., Addepalli, P., Gottumukkala, S.R., Digumarti, R.R., Kutala, V.K., 2011. Aberrations in one-carbon metabolism induce oxidative DNA damage in sporadic breast cancer. Mol. Cell. Biochem. 349, 159–167.

Naushad, S.M., Pavani, A., Rupasree, Y., Sripurna, D., Gottumukkala, S.R., Digumarti, R.R., Kutala, V.K., 2011a. Modulatory effect of plasma folate and polymorphisms in one-carbon metabolism on catecholamine methyltransferase (COMT) H108L associated oxidative DNA damage and breast cancer risk. Indian J. Biochem. Biophys. 48 (4), 283–289.

Naushad, S.M., Reddy, C.A., Rupasree, Y., Pavani, A., Digumarti, R.R., Gottumukkala, S.R., Kuppusamy, P., Kutala, V.K., 2011b. Cross-talk between one-carbon metabolism and xenobiotic metabolism: implications on oxidative DNA damage and susceptibility to breast cancer. Cell Biochem. Biophys. 61, 715–723.

Naushad, S.M., Pavani, A., Digumarti, R.R., Gottumukkala, S.R., Kutala, V.K., 2011c. Epistatic interactions between loci of one-carbon metabolism modulate susceptibility to breast cancer. Mol. Biol. Rep. 38 (8), 4893–4901 (Nov).

Naushad, S.M., Pavani, A., Rupasree, Y., Divyya, S., Deepti, S., Digumarti, R.R., Gottumukkala, S.R., Prayaga, A., Kutala, V.K., 2012a. Association of aberrations in one-carbon metabolism with molecular phenotype and grade of breast cancer. Mol. Carcinog. 51 (Suppl. 1), E32–E41 (Oct).

Naushad, S.M., Prayaga, A., Digumarti, R.R., Gottumukkala, S.R., Kutala, V.K., 2012b. Bcl-2/adenovirus E1B 19 kDa-interacting protein 3 (BNIP3) expression is epigenetically regulated by one-carbon metabolism in invasive duct cell carcinoma of breast. Mol. Cell. Biochem. 361 (1–2), 189–195 (Feb).

Naushad, S.M., Reddy, C.A., Kumaraswami, K., Divyya, S., Kotamraju, S., Gottumukkala, S.R., Digumarti, R.R., Kutala, V.K., 2014. Impact of hyperhomocysteinemia on breast cancer initiation and progression: epigenetic perspective. Cell Biochem. Biophys. 68 (2), 397–406.

Naushad, S.M., Hussain, T., Al-Attas, O.S., Prayaga, A., Digumarti, R.R., Gottumukkala, S.R., Kutala, V.K., 2015. Molecular insights into the association of obesity with breast cancer risk: relevance to xenobiotic metabolism and CpG island methylation of tumor suppressor genes. Mol. Cell. Biochem. 392 (1–2), 273–280.

Okobia, M., Bunker, C., Zmuda, J., Kammerer, C., Vogel, V., Uche, E., Anyanwu, S., Ezeome, E., Ferrell, R., Kuller, L., 2005. Cytochrome P4501A1 genetic polymorphisms and breast cancer risk in Nigerian women. Breast Cancer Res. Treat. 94 (3), 285–293.

Petrakis, N.L., 1977. Genetic factors in the etiology of breast cancer. Cancer 39, 2709–2715.

Pirouzpanah, S., Taleban, F.A., Mehdipour, P., Atri, M., 2015. Association of folate and other one-carbon related nutrients with hypermethylation status and expression of RARB, BRCA1, and RASSF1A genes in breast cancer patients. J. Mol. Med. (Berl) 93 (8), 917–934.

Price, R.J., Lillycrop, K.A., Burdge, G.C., 2015. Folic acid supplementation in vitro induces cell type-specific changes in BRCA1 and BRCA 2 mRNA expression, but does not alter DNA methylation of their promoters or DNA repair. Nutr. Res. 35 (6), 532–544.

Singh, N., Mitra, A.K., Garg, V.K., Agarwal, A., Sharma, M., Chaturvedi, R., Rath, S.K., 2007. Association of CYP1A1 polymorphisms with breast cancer in North Indian women. Oncol. Res. 16 (12), 587–597.

Stevens, V.L., McCullough, M.L., Pavluck, A.L., Talbot, J.T., Feigelson, H.S., Thun, M.J., Calle, E.E., 2007. Association of polymorphisms in one-carbon metabolism genes and postmenopausal breast cancer incidence. Cancer Epidemiol. Biomark. Prev. 16 (6), 1140–1147 (Jun).

U.S. Department of Agriculture, Agricultural Research Service., 2006v. USDA National Nutrient Database for Standard Reference, Release 18. Nutrient Data.

Vieira, A.R., Cooper, M.E., Marazita, M.L., Castilla, E.E., Orioli, I.M., 2008. Reduced folate carrier 1 (RFC1) is associated with cleft of the lip only. Braz. J. Med. Biol. Res. 41 (8), 689–693.

Wang, L., Chen, W., Wang, J., Tan, Y., Zhou, Y., Ding, W., Hua, Z., Shen, J., Xu, Y., Shen, H., 2006. Reduced folate carrier gene G80A polymorphism is associated with an increased risk of gastroesophageal cancers in a Chinese population. Eur. J. Cancer 42 (18), 3206–3211.

Wang, B.J., Liu, M.J., Wang, Y., Dai, J.R., Tao, J.Y., Wang, S.N., Zhong, N., Chen, Y., 2015. Association between SNPs in genes involved in folate metabolism and preterm birth risk. Genet. Mol. Res. 14 (1), 850–859.

Xu, X., Gammon, M.D., Zhang, H., Wetmur, J.G., Rao, M., Teitelbaum, S.L., Britton, J.A., Neugut, A.I., Santella, R.M., Chen, J., 2007. Polymorphisms of one-carbon-metabolizing genes and risk of breast cancer in a population-based study. Carcinogenesis 28 (7), 1504–1509.

Zhang, Y., Wise, J.P., Holford, T.R., Xie, H., Boyle, P., Zahm, S.H., Rusiecki, J., Zou, K., Zhang, B., Zhu, Y., Owens, P.H., Zheng, T., 2004. Serum polychlorinated biphenyls, cytochrome P-450 1 A1 polymorphisms, and risk of breast cancer in Connecticut women. Am. J. Epidemiol. 160 (12), 1177–1183.