

PRIMARY RESEARCH

Open Access

Identifying association model for single-nucleotide polymorphisms of *ORAI1* gene for breast cancer

Wei-Chiao Chang^{1,2,3,9†}, Yong-Yuan Fang^{4†}, Hsueh-Wei Chang^{3,5,6*}, Li-Yeh Chuang⁷, Yu-Da Lin⁸, Ming-Feng Hou^{3,9,10*} and Cheng-Hong Yang^{8*}

Abstract

Background: *ORAI1* channels play an important role for breast cancer progression and metastasis. Previous studies indicated the strong correlation between breast cancer and individual single nucleotide polymorphisms (SNPs) of *ORAI1* gene. However, the possible SNP-SNP interaction of *ORAI1* gene was not investigated.

Results: To develop the complex analyses of SNP-SNP interaction, we propose a genetic algorithm (GA) to detect the model of breast cancer association between five SNPs (rs12320939, rs12313273, rs7135617, rs6486795 and rs712853) of *ORAI1* gene. For individual SNPs, the differences between case and control groups in five SNPs of *ORAI1* gene were not significant. In contrast, GA-generated SNP models show that 2-SNP (rs12320939-GT/rs6486795-CT), 3-SNP (rs12320939-GT/rs12313273-TT/rs6486795-TC), 5-SNP (rs12320939-GG/rs12313273-TC/rs7135617-TT/rs6486795-TT/rs712853-TT) have higher risks for breast cancer in terms of odds ratio analysis (1.357, 1.689, and 13.148, respectively).

Conclusion: Taken together, the cumulative effects of SNPs of *ORAI1* gene in breast cancer association study were well demonstrated in terms of GA-generated SNP models.

Keywords: Single nucleotide polymorphism, Genetic algorithm, SNP interaction, Breast cancer

Background

Single nucleotide polymorphisms (SNPs) are one of the most common variants in human genome [1]. Currently, SNPs have been applied to the association studies for complex diseases [2-4]. Genome-wide association studies (GWAS) can identify the several SNPs predisposing to many diseases [5-8]. Although GWAS covers human genome-wide SNPs, many SNPs of non-significance are commonly ignored. Recently, the possible jointed effects of gene-gene interactions are gradually uncovered in predicting many disease risks [9-12]. However, when simultaneously evaluate the complex interactions amongst huge SNPs, these interactions are complex and it may need the help of new strategy [13] or computation [14].

Similarly, the non-GWAS association studies have the similar condition to ignore the possible gene-gene

interactions. For example, several individual SNPs of the *ORAI* calcium release-activated calcium modulator 1 (*ORAI1*) gene have reported to be involved in breast cancer susceptibility [15]. However, the possible SNP-SNP interactions of *ORAI1* gene associated with breast cancer were not addressed. Different computational analyses have been introduced to examine SNP-SNP interaction in many association studies [14,16-23]. Genetic algorithm (GA) is potential for feature selection for genome-wide scale datasets [24] and may apply to compute the difference between case and control groups to identify good models from the huge SNP combinations as well as tagSNP selection [25].

To address the possible SNP-SNP interaction in breast cancer susceptibility, five tagSNPs (rs12313273, rs6486795, rs7135617, rs12320939, and rs712853) of *ORAI1* gene were selected in this study. Therefore, we introduced the GA to optimizing the analyses of SNP-SNP interactions of *ORAI1* gene associated with breast cancer. GA is used to identify the best SNP models (SNP combinations with genotypes) with maximum frequency difference between breast cancer and control groups. Therefore, the best GA-generated SNP models of *ORAI1* gene may be useful for predicting the breast cancer risk.

* Correspondence: changhw@kmu.edu.tw; mifeho@kmu.edu.tw; chyang@cc.kuas.edu.tw

†Equal contributors

³Cancer Center, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan

⁸Department of Electronic Engineering, National Kaohsiung University of Applied Sciences, Kaohsiung, Taiwan

Full list of author information is available at the end of the article

Methods

Data set collection

The case and control subjects are 345 female breast cancer patients and 290 female normal controls where the recruitment was approved by Cancer Center of Kaohsiung Medical University Hospital. The genotype dataset of breast cancer patients of five tagSNPs (rs12313273, rs6486795, rs7135617, rs12320939, and rs712853) of *ORAI1* gene with minimum allele frequency (MAF) >10% obtained from our previous study [15]. For normal controls, samples of were collected in current study and SNP genotyping was performed as described [15].

Genetic algorithm

The GA [26] is a well-known evolutionary algorithm, and it has been applied for solving the complex problems in several fields. GA simulates the natural evolution to generate solutions of complex problems, including selection, crossover, mutation, and inheritance. The process of GA has six steps: (1) initializing population, (2) evaluate chromosome values, (3) select two parents using selection operation, (4) crossover operation, (5) mutation operation, and (6) replacement operation.

A population in first step is initialized according encoding schemes of problem. Second step aims to evaluate value of chromosomes in population using fitness function. Third step use the evaluated value of chromosomes to select the two good parents for generating two offspring (step 4). Then fifth step is probabilistic to mutate two offspring. Final step is used to improve the value of population. Thus repeat of steps 2 to 6 in several generations can effectively search the good values of chromosomes in population, and a best chromosome in population is regarded to best solution. Algorithm 1 shows the GA pseudo-code, and the below section is detailed to explain the processes of six steps.

Algorithm 1: Genetic algorithm pseudo-code.

- 01: begin
- 02: Initializing population according encoding schemes
- 03: for $g = 1$ to the number of generations
- 04: evaluate chromosomes of population using fitness function
- 05: select two parents using selection operation
- 06: generate two offspring using crossover operation
- 07: mutate two offspring using mutation operation
- 08: improve the value of population using replacement operation
- 09: next g
- 10: end

Encoding schemes

A population consists of the several possible solution of problem. The possible solution in GA is named a chromosome that is a set $C = \{c_1, \dots, c_d\}$. In this study, a

chromosome indicates a possible model of associations between SNPs. All combinations of SNPs and genotypes can be represented a set $A = S \times G = \{(s, g) \mid s \in S \text{ and } g \in G\}$ where S is a set of SNPs and G is a set of genotypes. For example, we assume an S contains two SNPs and a G contains three genotypes, i.e., $S = \{s_1, s_2\}$ and $G = \{g_1, g_2, g_3\}$. All possible subsets can be represented $A = S \times G = \{(s_1, g_1), (s_1, g_2), (s_1, g_3), (s_2, g_1), (s_2, g_2), (s_2, g_3)\}$. Each subset in A represents the selected SNP and their genotype. A chromosome is defined $C = \{c_1, \dots, c_d \mid c_i, c_j \in A, c_i \neq c_j, 1 \leq i \neq j \leq d\}$ where d is the association model size. A possible chromosome in above example can be assigned as $C = \{(s_1, g_1), (s_2, g_2)\}$; it means a model that includes the genotype "AA" of first SNP and the genotype "Aa" of second SNP.

Fitness function

A value of chromosome C can be evaluated by computing the fitness function; it facilitates GA for eliminating the worst chromosomes of population in each generation. In this study, a total number difference between case data and control data at a model is used to design a fitness function. Equation 1 is used to check a model whether a SNP is repeatedly selected or not. If a SNP is repeatedly selected in a C , the value of C is evaluated to zero. If it is not, Equation 2 is used to calculate the total number difference between cases and controls at a model. In Equation 2, the \max_P and \max_N are a total number of case data and a total number of control data, respectively. The P and N are respectively represented the set of case data and a set of control data; P_i is the i^{th} patient sample in case data and N_i is the i^{th} normal sample in control data. Equation 3 is used to evaluate whether all factors in a model are included in a set of sample. If a sample includes the model, the Equation 3 returns one value into Equation 2; whereas, it returns zero value.

$$f(C) = \begin{cases} 0 & \text{if } \bigcap_{s_i \in C} s_i \neq \emptyset \\ d(C) & \text{if } \bigcap_{s_i \in C} s_i = \emptyset \end{cases} \quad (1)$$

$$d(C) = \frac{\sum_{i=1}^{\max_P} m(C, P_i)}{\max_P} - \frac{\sum_{i=1}^{\max_N} m(C, N_i)}{\max_N} \quad (2)$$

$$m(X, Y) = \begin{cases} 0 & \text{if } X \not\subseteq Y \\ 1 & \text{if } X \subseteq Y \end{cases} \quad (3)$$

Selection operation

Selection operation aims to select the good chromosomes for generating the great offspring; the selected chromosomes name parents. Selection operation in this study uses a rank-based tournament scheme for selecting the two parents. The operation uses fitness function to evaluate all chromosomes of a population $P = \{C_1, \dots, C_i \mid i \text{ is population}$

size}, and all values in P are recorded into a set $R = \{r_1, \dots, r_i \mid i \text{ is population size}\}$. These values represent chromosome ranks. Then R is sorted from the big value to small value, i.e., $r_1 \geq r_2 \geq r_i$. Thus the r_1 and r_2 with corresponding C_s in P are two selected parents.

Crossover operation

Crossover operation is used to generate the offspring from the parents, and the operation use a uniform crossover scheme. Uniform crossover firstly generate a binary mask set $B = \{b_1, \dots, b_i \mid b_i \in [0,1], i = |C|\}$; a binary value at b is randomly generated. The one value of b_j indicates that j^{th} elements of two parents are must be exchanged; the zero value represents the no exchange. Two offspring are generated by exchanging the elements of two parents according the binary mask set, and the offspring are represented C' . For example, let a generated binary mask $B = \{1, 0, 1, 0\}$ and two parents $C_1 = \{(s_1, g_1), (s_2, g_2), (s_5, g_1), (s_3, g_3)\}$ and $C_2 = \{(s_1, g_3), (s_2, g_1), (s_4, g_2), (s_3, g_2)\}$. The generated two offspring are $C'_1 = \{(s_1, g_3), (s_2, g_2), (s_4, g_2), (s_3, g_3)\}$ and $C'_2 = \{(s_1, g_1), (s_2, g_1), (s_5, g_1), (s_3, g_2)\}$, respectively.

Mutation operation

Mutation operation can facilitate the diversity of chromosomes in population, and avoid population for trapping the local optimal. The operation uses a one-point mutation operation. A mutation point set $M = \{m_1, \dots, m_i \mid m_i \in [0,1], i = |C|\}$ is used to indicate the mutation points in the offspring C' . Each binary value in M is randomly generated according to the mutation probability. The one value of m_j represents that j^{th} element of C do the mutation; the zero value represents the no mutation. The mutation randomly generates a possible c element, where $c \in A$, to replace the original element at a mutation point. For example, let a generated mutation point $M = \{0, 0, 1, 0\}$ and offspring $C'_1 = \{(s_1, g_3), (s_2, g_2), (s_4, g_2), (s_3, g_3)\}$. Suppose the number of SNPs is five, a possible set is $E = \{(s_4, g_1), (s_4, g_2), (s_4, g_3), (s_5, g_1), (s_5, g_2), (s_5, g_3)\}$. After mutation the offspring could be $C'_1 = \{(s_1, g_3), (s_2, g_2), (s_5, g_3), (s_3, g_3)\}$.

Replacement operation

Replacement operation aims to gradually improve value of population. The generated two offspring C'_1 and C'_2 are evaluated by fitness function, and are used to compare the value to all chromosomes. When an offspring is higher value than a chromosome of population, it replaces the chromosome; otherwise, the offspring is deleted.

Parameter settings

In the GA parameters, both of the exchange probabilities in the tournament selection and uniform crossover are 1.0. The exchange probability of a one-point mutation is

0.1. The population size is 50, and the number of generations is 100.

Statistical analysis

All statistical value is computed using SPSS version 19.0 (SPSS Inc., Chicago, IL). Odds ratio (OR) with 95% confidence interval (CI) is used for measuring a single SNP and the model of association between SNPs; a P value of < 0.05 is considered statistically significant difference between the cases and controls.

Results

Data collection

The complete genotype data set is available at <http://bioinfo.kmu.edu.tw/BRCA-ORAI1-5SNPs.xlsx>. Based on these data, the GA-generated SNP models to address the possible SNP-SNP interaction in *ORAI1* gene were evaluated in terms of breast cancer association later.

Comparison of patients and normal in terms of effect of single SNP

Table 1 shows the occurrence of breast cancer for five SNPs in *ORAI1* gene. The genotype with major allele (G in rs12320939; T in rs12313273; G in rs7135617; T in rs6486795; and T in rs712853) is regarded as the reference for analyzing breast cancer risks in terms of single SNPs. Minor allele is selected according the dbSNP database of NCBI (National Center for Biotechnology Information). No significant differences between the breast cancer patients and controls in all genotypes for each single SNP were found.

Table 1 The performance of five individual SNPs for breast cancer and control groups

SNP ID	Genotype	Case (%)	Control (%)	p -value
1. rs12320939	GG	97 (28.12)	79 (27.24)	0.785
	GT	181 (52.46)	140 (48.28)	
	TT	67 (19.42)	71 (24.48)	
2. rs12313273	TT	183 (53.04)	161 (55.52)	0.189
	TC	142 (41.16)	100 (34.48)	
	CC	20 (5.80)	29 (10.00)	
3. rs7135617	GG	103 (29.86)	94 (32.41)	0.367
	GT	187 (54.20)	145 (50.00)	
	TT	55 (15.94)	51 (17.59)	
4. rs6486795	TT	137 (39.71)	121 (41.72)	0.260
	TC	173 (50.14)	126 (43.45)	
	CC	35 (10.14)	43 (14.83)	
5. rs712853	TT	154 (44.64)	128 (44.14)	0.904
	TC	158 (45.80)	134 (46.21)	
	CC	33 (9.57)	28 (9.66)	

Table 2 The best 10 models in 2-SNP order association

Combined SNP number (specific SNPs)	SNP Genotypes	Case no.	Control no.	Difference (%)*
SNP(1,4)	2-2	145	101	7.20
SNP(2,4)	2-3	22	1	6.03
SNP(3,4)	2-2	121	85	5.76
SNP(2,5)	2-3	83	54	5.44
SNP(2,4)	3-2	65	40	5.05
SNP(1,5)	2-3	81	55	4.51
SNP(4,5)	2-3	81	55	4.51
SNP(3,5)	2-3	78	53	4.33
SNP(2,3)	3-3	38	23	3.08
SNP(4,5)	2-2	77	56	3.01
SNP(1,3)	2-2	147	116	2.61

*Difference (%) = Frequency of cases (breast cancer group) - Frequency of controls.

Identification of the best model of SNPs association with maximum frequency difference between breast cancer and control groups

During GA processing, the best ten models of two SNP combinations with genotypes (2-SNP models) were demonstrated in Table 2. In these 2-SNP models, the SNPs (1, 4) with genotype 2-2, i.e., [rs12320939-GT]-[rs6486795-TC], possessed the maximum frequency difference (7.20%) between the breast cancer and control groups, namely the best 2-SNP model. Similarly, the best GA-generated SNP models involving three to five SNP were shown in left side of Table 3.

Odds ratio analysis to identify the best models of SNP associations associated with high breast cancer

Table 3 shows five best models of association involving two to five SNPs. Odds ratio analysis shows the performance (OR, 95% CI, and P value) of five types of the best models (2- to 5-SNP models) addressing the breast cancer association. The 2-SNP model, i.e., SNPs (1, 4) in genotype 2-2, indicates the borderline significance with

breast cancer (OR: 1.357, P = 0.064). The 3-SNP model, i.e., SNPs (1, 2, 4) in genotype 2-1-2, indicates three SNPs (rs12320939, rs12313273, and rs6486795) have a significant association when their genotypes are GT, TT, and TC, respectively (OR: 1.689, P = 0.028). The 5-SNP model, SNPs (1, 2, 3, 4, 5) in genotype 1-2-3-1-1, indicates all SNPs (rs12320939, rs12313273, rs7135617, rs6486795, and rs712853) have a strongly association when their genotypes are GG, TC, TT, TT, and TT, respectively (OR: 13.148, P = 0.013).

Discussion

GA is a robust non-parametric method that detects non-linear interactions amongst multiple discrete genetic factors. The advantage of GA is that the method can directly search the good models from the huge number of possible combinations without the training data set. In this study, the fitness function is designed based on the unbalanced data set to compute the difference between case data set and control data set. The function can effectively measure high-risk to search the good model in real data set.

In current study, the OR values of 2- to 3-SNP models are larger than 1 but small, suggesting that the cumulative effect of these four SNPs (rs12320939, rs7135617, rs6486795, and rs712853) are weak. When five SNPs included, the OR value is 13.148, indicating that the cumulative effect of 5-SNP model becomes strong. This unstable cumulative effect of SNP combinations in SNP models may be partly explained by the experiment design that these five SNPs were only derived from a single gene *ORAI1*. Because breast cancer is a kind of multi-gene disease [27-30], therefore, SNPs derived from more genes included in association studies may reveal the cumulative effect effectively [9,11,12,31-33]. Accordingly, the differential performance of the cumulative effects of SNPs from single gene and multigene is worth of further investigation in future.

Table 3 The odds ratio of the best SNP models associated with breast cancer

Combined SNP no. (specific SNPs)	SNP Genotypes	Case no./control no.	Cancer (%)	Difference (%)*	Odds ratio (95%CI)	P value
2-SNP (1-4)	2-2	145/101	58.94	7.20	1.357	0.064
	Others	200/189	51.41		(0.983-1.873)	
3-SNP (1-2-4)	2-1-2	58/31	65.17	6.12	1.689	0.028
	Others	287/259	52.56		(1.058-2.694)	
4-SNP (1-3-4-5)	2-2-2-1	78/53	59.54	4.33	1.306	0.180
	others	267/237	52.98		(0.884-1.930)	
5-SNP (1-2-3-4-5)	1-2-3-1-1	14/0	100.00	4.06	13.148	0.013
	Others	331/290	53.30		(1.726-100.142)	

Bold letters: The SNP models associated with breast cancer are significant (P value < 0.05).

*Difference (%) = Frequency of cases (breast cancer group) - Frequency of controls.

The computational complexity of GA is calculated by a fitness function of computation. Suppose n iterations is implemented in a test, the computational complexity of GA is $O(n)$ which represents the big-O in complexity analysis. GA in search of good association model has the below advantages: (1) GA effectively identify the high-risk models in high-order interaction, (2) the best model with statistical significant can be fast identified, and (3) it only has two parameters to need setting and is easily to fulfil for searching the good model. Further, GA is able to analyze high order SNP interactions amongst the huge number of SNPs from GWAS and pharmacogenomics studies in our experiences.

Conclusions

Although the polymorphisms of *ORAI1* gene have been reported to associate with inflammatory diseases [34–36], effects of SNP-SNP interaction to diseases are still unclear. In this study, the GA successfully identified appropriate models of SNP-SNP interactions in breast cancer association study in terms of five SNPs in *ORAI1* gene. The resulting SNP models can predict the breast cancer susceptibility more effective than the individual SNPs. This methodology can also apply to any kinds of SNP association studies, such as GWAS, pharmacogenomics and others. Therefore, the possible cumulative effect of SNP combination will be uncovered by this methodology.

Competing interests

The authors have no competing interests to declare.

Authors' contributions

W-CC and M-FH carried out the molecular genetic studies. W-CC and Y-YF drafted the manuscript. C-HY instructed Y-DL to perform SNP interaction analysis by algorithm. Y-YF, H-WC, and L-YC performed statistics analyses. H-WC, C-HY, and M-FH were involved in discussion and editing the manuscript. All authors read and approved the final manuscript.

Acknowledgements

This study was partly supported by the National Science Council of Taiwan for Grant NSC101-2221-E-214-075, NSC101-2622-E-151-027-CC3, NSC100-2221-E-151-049-MY3, NSC100-2221-E-151-051-MY2, NSC101-2320-B038-029-MY3, the National Sun Yat-sen University-KMU Joint Research Project (#NSYSU-KMU 103-p014), the Ministry of Health and Welfare, Taiwan, Republic of China (MOHW103-TD-B-111-05), and TMU101-AE1-B14 (to Wei-Chiao Chang).

Author details

¹Master Program for Clinical Pharmacogenomics and Pharmacoproteomics, School of Pharmacy, Taipei Medical University, Taipei, Taiwan. ²Department of Clinical Pharmacy, School of Pharmacy, Taipei Medical University, Taipei, Taiwan. ³Cancer Center, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan. ⁴Labor Safety and Health Office, Kaohsiung Municipal Ta-Tung Hospital, Kaohsiung, Taiwan. ⁵Department of Biomedical Science and Environmental Biology, Kaohsiung Medical University, Kaohsiung, Taiwan. ⁶Institute of Medical Science and Technology, National Sun Yat-sen University, Kaohsiung, Taiwan. ⁷Department of Chemical Engineering & Institute of Biotechnology and Chemical Engineering, I-Shou University, Kaohsiung, Taiwan. ⁸Department of Electronic Engineering, National Kaohsiung University of Applied Sciences, Kaohsiung, Taiwan. ⁹Institute of Clinical Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan. ¹⁰Kaohsiung Municipal Ta-Tung Hospital, Kaohsiung, Taiwan.

Received: 2 November 2013 Accepted: 7 March 2014

Published: 31 March 2014

References

1. Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, Belmont JW, Boudreau A, Hardenbol P, Leal SM, Pasternak S, Wheeler DA, Willis TD, Yu F, Yang H, Zeng C, Gao Y, Hu H, Hu W, Li C, Lin W, Liu S, Pan H, Tang X, Wang J, Wang W, Yu J, Zhang B, Zhang Q, Zhao H, et al: **A second generation human haplotype map of over 3.1 million SNPs.** *Nature* 2007, **449**(7164):851–861.
2. Ding H, Jing X, Ding N, Fu Z, Song Y, Zhu J: **Single nucleotide polymorphisms of CD20 gene and their relationship with clinical efficacy of R-CHOP in patients with diffuse large B cell lymphoma.** *Cancer Cell Int* 2013, **13**(1):58.
3. Flintoft L: **Complex disease: A SNP for disease prognosis.** *Nat Rev Genet* 2013, **14**(11):746.
4. Hwang CW, Lu CH, Sun SF, Sung TY, Chung HY, Huang SY, Hung HC, Chen CH, Sun YM, Lin YY, Liu WS, Wen ZH: **Comprehensive association analysis of 10 single nucleotide polymorphisms associated with osteoporosis among a Taiwanese population.** *Int J Hum Genet* 2011, **11**(4):249–257.
5. Sober S, Org E, Kepp K, Juhanson P, Eyheramendy S, Gieger C, Lichtner P, Klopp N, Veldre G, Viigimaa M, Doring A, Putku M, Kelgo P, Shaw-Hawkins S, Howard P, Onipinla A, Dobson RJ, Newhouse SJ, Brown M, Dominiczak A, Connell J, Samani N, Farrall M, Caulfield MJ, Munroe PB, Illig T, Wichmann HE, Meitinger T, Laan M: **Targeting 160 candidate genes for blood pressure regulation with a genome-wide genotyping array.** *PLoS One* 2009, **4**(6):e6034.
6. Kraft P, Haiman CA: **GWAS identifies a common breast cancer risk allele among BRCA1 carriers.** *Nat Genet* 2010, **42**(10):819–820.
7. Lewis SN, Nsoesie E, Weeks C, Qiao D, Zhang L: **Prediction of disease and phenotype associations from genome-wide association studies.** *PLoS One* 2011, **6**(11):e27175.
8. Yang C, Zhou X, Wan X, Yang Q, Xue H, Yu W: **Identifying disease-associated SNP clusters via contiguous outlier detection.** *Bioinformatics* 2011, **27**(18):2578–2585.
9. Sun J, Chang BL, Isaacs SD, Wiley KE, Wiklund F, Stattin P, Duggan D, Carpten JD, Trock BJ, Partin AW, Walsh PC, Gronberg H, Xu J, Isaacs WB, Zheng SL: **Cumulative effect of five genetic variants on prostate cancer risk in multiple study populations.** *Prostate* 2008, **68**(12):1257–1262.
10. Yen CY, Liu SY, Chen CH, Tseng HF, Chuang LY, Yang CH, Lin YC, Wen CH, Chiang WF, Ho CH, Chen HC, Wang ST, Lin CW, Chang HW: **Combinational polymorphisms of four DNA repair genes XRCC1, XRCC2, XRCC3, and XRCC4 and their association with oral cancer in Taiwan.** *J Oral Pathol Med* 2008, **37**(5):271–277.
11. Vogelsang M, Wang Y, Veber N, Mwapagha LM, Parker MI: **The cumulative effects of polymorphisms in the DNA mismatch repair genes and tobacco smoking in oesophageal cancer risk.** *PLoS One* 2012, **7**(5):e36962.
12. Yang CH, Chuang LY, Cheng YH, Lin YD, Wang CL, Wen CH, Chang HW: **Single nucleotide polymorphism barcoding to evaluate oral cancer risk using odds ratio-based genetic algorithms.** *Kaohsiung J Med Sci* 2012, **28**(7):362–368.
13. Kotti S, Bickeboller H, Clerget-Darpoux F: **Strategy for detecting susceptibility genes with weak or no marginal effect.** *Hum Hered* 2007, **63**(2):85–92.
14. Steen KV: **Travelling the world of gene-gene interactions.** *Brief Bioinform* 2012, **13**(1):1–19.
15. Chang WC, Woon PY, Hsu YW, Yang S, Chiu YC, Hou MF: **The association between single-nucleotide polymorphisms of ORAI1 gene and breast cancer in a Taiwanese population.** *ScientificWorldJournal* 2012, **2012**:916587.
16. Yang CH, Chang HW, Cheng YH, Chuang LY: **Novel generating protective single nucleotide polymorphism barcode for breast cancer using particle swarm optimization.** *Canc Epidemiol* 2009, **33**(2):147–154.
17. Chang HW, Yang CH, Ho CH, Wen CH, Chuang LY: **Generating SNP barcode to evaluate SNP-SNP interaction of disease by particle swarm optimization.** *Comput Biol Chem* 2009, **33**(1):114–119.
18. Moore JH, Asselbergs FW, Williams SM: **Bioinformatics challenges for genome-wide association studies.** *Bioinformatics* 2010, **26**(4):445–455.
19. Yang CH, Chuang LY, Chen YJ, Tseng HF, Chang HW: **Computational analysis of simulated SNP interactions between 26 growth factor-related genes in a breast cancer association study.** *OMICS* 2011, **15**(6):399–407.

20. Chuang LY, Chang HW, Lin MC, Yang CH: **Chaotic particle swarm optimization for detecting SNP-SNP interactions for CXCL12-related genes in breast cancer prevention.** *Eur J Cancer Prev* 2012, **21**(4):336–342.
21. Yang P, Ho JW, Yang YH, Zhou BB: **Gene-gene interaction filtering with ensemble of filters.** *BMC Bioinforma* 2011, **12**(Suppl 1):S10.
22. Chuang LY, Lin YD, Chang HW, Yang CH: **An improved PSO algorithm for generating protective SNP barcodes in breast cancer.** *PLoS One* 2012, **7**(5):e37018.
23. Chuang LY, Chang HW, Lin MC, Yang CH: **Improved branch and bound algorithm for detecting SNP-SNP interactions in breast cancer.** *J Clin Bioinform* 2013, **3**(1):4.
24. Zhang J, Jiang S, Zhao X, Liu L: **A genetic algorithm to filter SNPs for SNP association study.** *2012 IEEE/WIC/ACM Int Conferences Web Intelligen Intell Agent Technol* 2012, **1**:684–687.
25. Mahdevar G, Zahiri J, Sadeghi M, Nowzari-Dalini A, Ahrabian H: **Tag SNP selection via a genetic algorithm.** *J Biomed Inform* 2010, **43**(5):800–804.
26. Holland JH: *Adaptation in nature and artificial systems: An introductory analysis with applications to biology, control and artificial intelligence.* Ann Arbor, MI: University of Michigan Press; 1975.
27. Solin LJ, Gray R, Baehner FL, Butler SM, Hughes LL, Yoshizawa C, Cherbavaz DB, Shak S, Page DL, Sledge GW Jr, Davidson NE, Ingle JN, Perez EA, Wood WC, Sparano JA, Badve S: **A multigene expression assay to predict local recurrence risk for ductal carcinoma in situ of the breast.** *J Natl Canc Inst* 2013, **105**(10):701–710.
28. Duggal S, Julian TB: **A multigene expression assay to predict local recurrence risk for ductal carcinoma in situ.** *J Natl Canc Inst* 2013, **105**(10):681–683.
29. Rouzier R, Pronzato P, Chereau E, Carlson J, Hunt B, Valentine WJ: **Multigene assays and molecular markers in breast cancer: systematic review of health economic analyses.** *Breast Canc Res Treat* 2013, **139**(3):621–637.
30. Ignatiadis M: **Multigene assays for late recurrence of breast cancer.** *Lancet Oncol* 2013, **14**(11):1029–1030.
31. Lin GT, Tseng HF, Chang CK, Chuang LY, Liu CS, Yang CH, Tu CJ, Wang EC, Tan HF, Chang CC, Wen CH, Chen HC, Chang HW: **SNP combinations in chromosome-wide genes are associated with bone mineral density in Taiwanese women.** *Chin J Physiol* 2008, **51**(1):32–41.
32. Lin GT, Tseng HF, Yang CH, Hou MF, Chuang LY, Tai HT, Tai MH, Cheng YH, Wen CH, Liu CS, Huang CJ, Wang CL, Chang HW: **Combinational polymorphisms of seven CXCL12-related genes are protective against breast cancer in Taiwan.** *OMICS* 2009, **13**(2):165–172.
33. Chen JB, Chuang LY, Lin YD, Liou CW, Lin TK, Lee WC, Cheng BC, Chang HW, Yang CH: **Preventive SNP-SNP interactions in the mitochondrial displacement loop (D-loop) from chronic dialysis patients.** *Mitochondrion* 2013, **13**:698–704.
34. Chang WC, Lee CH, Hirota T, Wang LF, Doi S, Miyatake A, Enomoto T, Tomita K, Sakashita M, Yamada T, Fujieda S, Ebe K, Saeki H, Takeuchi S, Furue M, Chen WC, Chiu YC, Chang WP, Hong CH, Hsi E, Juo SH, Yu HS, Nakamura Y: **ORAI1 genetic polymorphisms associated with the susceptibility of atopic dermatitis in Japanese and Taiwanese populations.** *PLoS One* 2012, **7**(1):e29387.
35. Chou YH, Juo SH, Chiu YC, Liu ME, Chen WC, Chang CC, Chang WP, Chang JG, Chang WC: **A polymorphism of the ORAI1 gene is associated with the risk and recurrence of calcium nephrolithiasis.** *J Urol* 2011, **185**(5):1742–1746.
36. Wei JC, Yen JH, Juo SH, Chen WC, Wang YS, Chiu YC, Hsieh TJ, Guo YC, Huang CH, Wong RH, Wang HP, Tsai KL, Wu YC, Chang HW, Hsi E, Chang WP, Chang WC: **Association of ORAI1 haplotypes with the risk of HLA-B27 positive ankylosing spondylitis.** *PLoS One* 2011, **6**(6):e20426.

doi:10.1186/1475-2867-14-29

Cite this article as: Chang et al.: Identifying association model for single-nucleotide polymorphisms of *ORAI1* gene for breast cancer. *Cancer Cell International* 2014 **14**:29.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

