

PRIMARY RESEARCH

Open Access

Identification of SNP barcode biomarkers for genes associated with facial emotion perception using particle swarm optimization algorithm

Li-Yeh Chuang¹, Hsien-Yuan Lane^{2,3}, Yu-Da Lin⁴, Ming-Teng Lin^{1,5}, Cheng-Hong Yang^{4*} and Hsueh-Wei Chang^{6,7,8*}

Abstract

Background: Facial emotion perception (FEP) can affect social function. We previously reported that parts of five tested single-nucleotide polymorphisms (SNPs) in the MET and AKT1 genes may individually affect FEP performance. However, the effects of SNP-SNP interactions on FEP performance remain unclear.

Methods: This study compared patients with high and low FEP performances ($n = 89$ and 93 , respectively). A particle swarm optimization (PSO) algorithm was used to identify the best SNP barcodes (i.e., the SNP combinations and genotypes that revealed the largest differences between the high and low FEP groups).

Results: The analyses of individual SNPs showed no significant differences between the high and low FEP groups. However, comparisons of multiple SNP-SNP interactions involving different combinations of two to five SNPs showed that the best PSO-generated SNP barcodes were significantly associated with high FEP score. The analyses of the joint effects of the best SNP barcodes for two to five interacting SNPs also showed that the best SNP barcodes had significantly higher odds ratios (2.119 to 3.138; $P < 0.05$) compared to other SNP barcodes. In conclusion, the proposed PSO algorithm effectively identifies the best SNP barcodes that have the strongest associations with FEP performance.

Conclusions: This study also proposes a computational methodology for analyzing complex SNP-SNP interactions in social cognition domains such as recognition of facial emotion.

Keywords: Particle swarm optimization, Single-nucleotide polymorphism, SNP interaction, Facial emotion perception, Algorithm

Introduction

Facial emotion perception (FEP) is a major indicator of social function [1-3]. In schizophrenia, FEP tends to be deficient [4], and improvements in FEP are associated with improvements in occupational functioning and independent-living capability [5]. However, potential genetic associations with FEP performance have not been addressed.

Recent reports show that the Met proto-oncogene (hepatocyte growth factor receptor) (MET) and v-akt murine thymoma viral oncogene homolog 1 (AKT1) genes have important roles in psychiatric conditions [6-8]. For

example, four single-nucleotide polymorphisms (SNPs) (rs2237717, rs41735, rs42336, and rs1858830) of the MET gene and the rs1130233 SNP of the AKT1 gene have been associated with schizophrenia [9-11] and with neurocognitive performance [9,11]. Our previous study [12] was the first to explore genetic factors in the associations of these five SNPs with FEP performance. Although the roles of individual SNPs have been investigated in FEP performance, the roles of SNP-SNP interactions remain unclear.

Accumulating evidence obtained by polygenetic models reveal the joint effects of multiple SNPs in various diseases [13-21]. Computational algorithm tools have improved the identification of functional SNPs [22], but not for the SNP-SNP interaction issue. Recently, various computational algorithms have been developed to investigate SNP-SNP interactions in numerous association studies [21,23-33].

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

⁴Department of Electronic Engineering, National Kaohsiung University of Applied Sciences, Kaohsiung 80778, Taiwan

⁶Cancer Center, Translational Research Center, Kaohsiung Medical University Hospital, Kaohsiung Medical University, Kaohsiung 80708, Taiwan
Full list of author information is available at the end of the article

This study applied the particle swarm optimization (PSO) algorithm [34] to evaluate the use of specific combinational SNP patterns and their corresponding genotypes, i.e., SNP barcodes, for discriminating between high and low FEP performances in terms of SNP-SNP interactions. The FEP performance of the PSO-generated SNP barcodes were further assessed by odds ratio (OR) analyses. In each of these two related genes, five SNP combinations were systematically assessed in terms of their potential joint effects on FEP performance.

Methods

The SNP barcodes, i.e., the unique patterns of SNP combinations with their corresponding genotypes, were generated by a PSO algorithm [34]. The PSO is characterized by its rapid convergence, which enables rapid identification of optimal solutions in a wide solution space. Here, PSO was used for rapid identification of the optimal SNP barcodes, namely, the best SNP barcodes, i.e., those that revealed the largest differences between the high and low FEP groups.

Introduction to particle swarm optimization

The robustness of PSO [34] results from its use of swarm intelligence to search for the optimal solution to a complex problem. The swarm intelligence can be described as a system that automatically evolves by simulating the social behavior of organisms, e.g., the social behavior of knowledge sharing. By sharing valuable information, the behaviors of individuals in a swarm are optimized to achieve a certain objective. In PSO, an individual is considered a particle, which is a vector in the problem space. The information for the particle includes knowledge gained from its previous experience and knowledge gained from the swarm. The value of the particle, which is estimated by the objective function, is used to update its information and to optimize the objective of the swarm. Therefore, the swarm can converge to exploit good resolution in local regions of the problem space; the common objective can also be updated when one particle finds a better objective so that the particle can lead the swarm in exploring a different region of the problem space. These superior search characteristics have made PSO the most popular evolutionary algorithm in several fields.

PSO definition

Figure 1 shows the steps of the PSO procedure: (1) initializing the particles, (2) using an objective function to evaluate the particles, (3) selecting the particles' *pbest* and *gbest*, and (4) updating the velocity and position of the particles. These procedures are repeated in successive iterations until the termination conditions are reached. The elements of the proposed PSO include (1) particle,

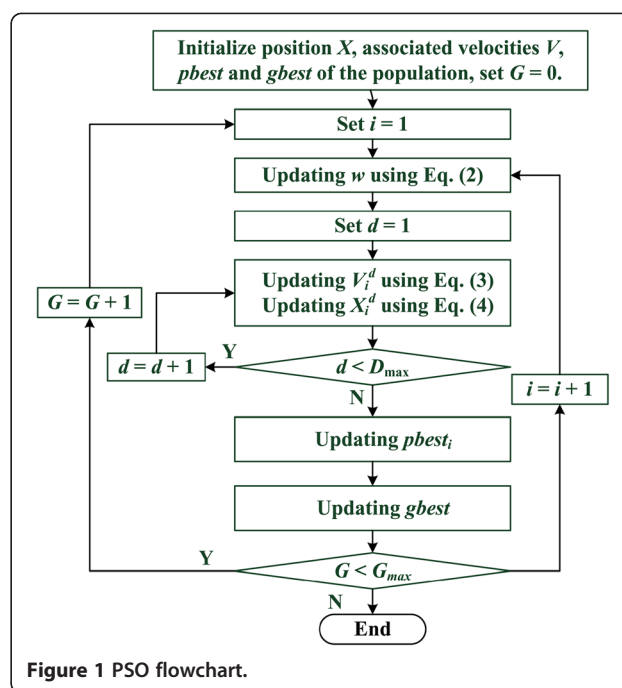


Figure 1 PSO flowchart.

(2) population, (3) particle velocity, (4) particle position, (5) inertia weight, (6) individual best value, (7) global best value, and (8) termination criteria. These elements are defined as follows.

1. Particle: a problem solution consisting of several parameters.
2. Population: the set of all particles.
3. Particle velocity (v): each particle has a velocity, and the velocity of the i th particle is $v_i = (v_{i1}, v_{i2}, \dots, v_{iD})$, where v_{iD} is the velocity of the i th particle with respect to the D th parameter.
4. Particle position (x): a vector consisting of available parameters in the practicable solution. The position of the i th particle is $x_i = (x_{i1}, x_{i2}, \dots, x_{iD})$, where x_{iD} is the position of the i th particle with respect to the D th parameter. Each parameter in the particle vector is defined by the number of selected SNPs and their corresponding genotypes for the associated SNPs.
5. Inertia weight (w): the current velocity may be affected by the inertia weight, which determines whether the previous velocity increases/decreases the current velocity. Control factor w affects the trade-off between the exploration and exploitation capabilities of the particle.
6. Individual best fitness value (*pbest*): the position of the i th particle with the highest fitness value ($pbest_i$) is considered the best current solution for the i th particle.
7. Global best fitness value (*gbest*): the best particle in the population.

8. Termination criteria: the process stops when the maximum allowed number of iterations is reached.

Encoding schemes

A vector of a particle is defined as the number of selected SNPs and their corresponding genotypes; SNPs cannot be repeatedly selected. Particle encoding can be represented as a vector as follows:

$$\text{Particle}_i = (\text{SNP}_1, \text{SNP}_2, \dots, \text{SNP}_m, \text{Geno}_1, \text{Geno}_2, \dots, \text{Geno}_m), i = 1, 2, \dots, n,$$

where SNP represents the selected SNPs, Geno represents the SNP-corresponding genotypes (AA, Aa, and aa), and n represents the population size.

Population initialization

The above encoding schemes are used for randomly generating the particle population over the problem space. SNPs are selected based on a randomly generated value among available SNPs. Three genotypes are also randomly generated between 1 and 3 to represent 'AA', 'Aa', and 'aa'. The SNPs selected between SNP_1 and SNP_m differ for each particle. If a given SNP is found to be repeated in a particle, the SNP is randomly generated among the available SNPs until it is different. For example, for a two-SNP particle with vector = (1, 4, 2, 3), SNPs 1 and 4 and genotypes 2 and 3 are chosen. The SNP barcodes (i.e., the SNPs and their corresponding genotypes) are represented as (1, 2) and (4, 3).

Objective function

An objective function was used to compute the SNP barcode for the difference between the groups with high and low FEP values; the objective function was also used to compute the fitness value. The largest difference between the high and low groups was the highest fitness value. The equation was defined as follows:

$$F(X_i) = (\text{High-FEP group} \cap X_i) - (\text{Low-FEP group} \cap X_i), \quad (1)$$

where 'high-FEP group' and 'low-FEP group' are the total number of SNP interactions in the high-FEP group and the low-FEP group, respectively. The X_i represents the i th particle. The high-FEP group $\cap X_i$ was defined as the total number of intersections between the high-FEP group and the i th particle. The low-FEP group $\cap X_i$ was defined as the total number of intersections between the low-FEP group and the i th particle. For example, $X = (1, 2, 3, 1)$, i.e., SNP 1 with 'aa' and SNP 2 with 'AA', was used to evaluate the number of matching conditions in the two groups. The high-FEP group contained 30 samples, and the low-FEP group had 18

samples. According to Equation 1, the value was calculated by subtracting 18 from 30, leaving 12.

Updating $pbest$ and $gbest$

The personal best position ($pbest$) and the global best position ($gbest$) must be recorded for moving particles. If the current particle value is better than $pbest$, then the position and fitness value of $pbest$ are updated with the current position and fitness value. If the fitness value of $pbest$ is better than that of $gbest$, $gbest$ is reset to $pbest$.

Updating particle velocity and position

The search direction of a particle is determined by the three different vectors, i.e., w , $pbest$, and $gbest$. Equation 2 shows the w updating function, which iteratively reduces the value of w from w_{max} to w_{min} , i.e., the function is a positive linear function of time that changes with the generations. The PSO algorithm applies Equation 3 to update the particle velocity. Equation 4 is the change in position (i.e., possible solution) needed to search for a better solution.

$$w_{LDW} = (w_{max} - w_{min}) \times \frac{\text{Iteration}_{max} - \text{Iteration}_i}{\text{Iteration}_{max}} + w_{min} \quad (2)$$

$$v_{id}^{new} = w_{LDW} \times v_{id}^{old} + c \times r_1 \times (pbest_{id} - x_{id}^{old}) + c \times r_2 \times (gbest_d - x_{id}^{old}) \quad (3)$$

$$x_{id}^{new} = x_{id}^{old} + v_{id}^{new}, \quad (4)$$

where w_{LDW} is the inertia weight, w_{max} is 0.9, w_{min} is 0.4, and Iteration_{max} is the maximum number of allowed iterations. The r_1 and r_2 , which are the random functions in the range [0, 1], are used to adjust the strength of the $pbest$ and $gbest$ vectors in a single generation. The c_1 and c_2 are acceleration constants that control the particle search direction ($pbest$ or $gbest$). Velocities v_{id}^{new} and v_{id}^{old} represent the new and old velocities, respectively. The x_{id}^{old} and x_{id}^{new} are the current and updated particle positions, respectively. The velocity indicates the change required for a particle to obtain its global best position at a given moment in time, i.e., the velocity of the particle flying toward the best position.

PSO parameter settings

As suggested by [34], the PSO parameters were optimized as follows. The population size was set to 50. The PSO termination condition was set to a pre-specified number of iterations (here, 100). The starting value of the inertia weight w was set to 0.9, and the final value was set to 0.4 [35]. Acceleration (learning) factors c_1 and c_2 were both set to 2 [36].

Statistical analysis

Statistical analysis was performed using SPSS version 19.0 (SPSS Inc., Chicago, IL, USA). The OR with 95% confidence interval (CI) was used for performance measurements of the best SNP barcode; a *P* value smaller than 0.05 was considered a statistically significant difference between the two groups. Finally, G*Power 3 (Heinrich-Heine-Universität Düsseldorf, Germany) [37] was used to estimate the power.

Results

Example

The dataset used in this study was obtained from our previous study of associations with FEP [12] in a Taiwan population of 182 unrelated healthy subjects (82 men and 100 women). The Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT) [38] was used to measure FEP in terms of perceiving, facilitating, understanding, and managing emotions [39]. The complete genotype data set is available at http://bioinfo.kmu.edu.tw/All_genotypes_for_low_high_FEP.xlsx.

The objective of this study was to evaluate the use of PSO for identifying the best combinations of SNPs in specific genotypes. The subjects had an average FEP score of 90.66 ± 1.49 (mean \pm SE). Based on their scores, the subjects were dichotomized into a high-FEP group (FEP score > 90.66) and a low-FEP group (FEP score ≤ 90.66). Accordingly, PSO-generating SNP barcodes were coupled with the phenotype (FEP) to analyze 89 subjects with high FEP scores and 93 subjects with low FEP scores.

Comparison of high- and low-FEP groups in terms of effects of single SNP

Table 1 shows the performance (*P* value) of each of five SNPs of two related genes (MET and AKT1). The occurrence of the genotype for each SNP did not significantly differ between the high- and low-FEP groups ($P > 0.05$).

The best SNP barcodes for the high- and low-FEP groups

Table 2 shows ten representative barcodes for combinations of two SNPs (two-SNP barcodes), and the magnitude of difference between the high- and low-FEP groups is presented from maximal to minimal. Out of all the combinations, the two-SNP barcodes, e.g., SNPs (1, 4) in genotype 3-3, [rs1130233-AG]-[rs41735-GA], showed the largest difference (12) between the high-FEP and low-FEP groups (30 vs. 18, respectively). The other two-SNP barcodes are provided on-line (http://bioinfo.kmu.edu.tw/All_2SNP_barcodes_for_FEP_scores.xlsx) along with all possible combinations of two SNPs with genotypes. The PSO was also used to obtain the SNP barcodes for the best-performing combinations of three to five SNPs

Table 1 Single SNPs in high-FEP and low-FEP groups

SNP no.	SNPs (genes)	Genotypes	High-FEP group (n)	Low-FEP group (n)	<i>P</i> value
1	rs1130233 (AKT1)	1-AA	22	33	0.087
		2-GG	21	15	
		3-AG	50	41	
2	rs1858830 (MET)	1-GG	42	37	0.797
		2-GC	44	42	
		3-CC	7	10	
3	rs2237717 (MET)	1-CC	25	28	0.370
		2-CT	18	23	
		3-TT	50	38	
4	rs41735 (MET)	1-GG	25	29	0.931
		2-AA	17	19	
		3-GA	51	41	
5	rs42336 (MET)	1-GG	17	20	0.985
		2-AA	24	28	
		3-GA	52	41	

(i.e., the combinations that obtained the largest differences between the high-FEP group and the low-FEP group).

Associations between PSO-generated SNP barcodes and FEP

The left side of Table 3 shows the best PSO-generated SNP barcodes for combinations of two to five SNPs. The PSO algorithm identified the SNP interactions with the largest differences between the high- and low-FEP groups. For example, the largest difference in two-SNP interactions was in the SNP barcode for [rs1130233-AG] and [rs41735-GA]. The largest difference in three-SNP interactions was in the SNP barcode for SNPs 1, 4, and 5 of genotypes 3-3-3, i.e., [rs1130233-AG]-[rs41735-GA]-[rs42336-GA]. The largest difference in four-SNP interactions was in the SNP barcode for SNPs 1, 3, 4, and 5 of genotypes 3-3-3-3, i.e., [rs1130233-AG]-[rs2237717-TT]-[rs41735-GA]-[rs42336-GA]. Restated, the best SNP barcodes generated by PSO for fixed numbers of SNPs consistently obtained the largest differences in SNPs between the high- and low-FEP groups.

The OR rankings indicate the SNP barcodes associated with high FEP

The right side of Table 3 shows the estimated effect (*P* value, OR, and 95% CI) of the best PSO-generated SNP barcodes in terms of FEP. The occurrence of the best SNP barcodes (combinations of two to five SNPs) significantly differed ($P < 0.05$) between the high-FEP group and the low-FEP group (range, 12 to 7 for barcodes with two to five SNPs). The OR values of the best SNP barcodes for two to five SNPs ranged from 2.119 to 3.138,

Table 2 Ten representative differences in the occurrence of SNP barcodes for two-SNP combinations

Combined SNPs ^a	Genotypes ^a	High-FEP group (n)	Low-FEP group (n)	Difference ^b
SNPs(1,4)	3-3	30	18	12
SNPs(1,3)	3-3	27	17	10
SNPs(1,2)	3-1	27	18	9
SNPs(2,3)	2-3	28	21	7
SNPs(2,4)	2-3	28	22	6
SNPs(2,3)	1-2	13	8	5
SNPs(1,2)	2-2	8	4	4
SNPs(1,2)	2-1	11	8	3
SNPs(1,4)	2-1	6	4	2
SNPs(1,2)	3-3	3	2	1

^aNumbers in parentheses indicates SNP combinations. For example, SNPs(1, 4) of genotypes 3-3 is the combination of [rs1130233-AG] and [rs41735-GA]. For detailed SNP data, see Table 1. ^bDifference should be provided with footnote, i.e., difference = High-FEP group (n) – Low-FEP group (n).

the 95% CI of OR ranged from 1.059 to 10.253, and the power ranged from 0.556 to 0.400.

Discussion

Single-SNP models are typically used to evaluate SNPs for associations with various diseases [40-44]. For some SNPs, however, effects that are potentially important but do not reach the significance level are underestimated, leading to the ‘missing heritability’ [45-47]. Recent psychiatry and neuroscience studies have investigated the importance of non-significant effects of SNPs. For example, studies of schizophrenia have revealed synergistic SNP-SNP interactions in dopamine receptor D4 (DRD4) [48] and in neuregulin 1 (NRG1)/v-erb-a erythroblastic leukemia viral oncogene homolog 4 (avian) (ERBB4)/AKT1 genes [49-51].

Our previous study [12] compared FEP scores among five SNPs of three genotypes in each of the MET and AKT1 genes. Chi-square test showed that the FEP scores significantly differed ($P < 0.05$) in only one genotype

(rs2237717) out of three MET genotypes. However, the genotypes were not analyzed in terms of associations with FEP scores or in terms of all potential SNP-SNP interactions.

In the present study, the comparison of all SNPs in the MET and AKT1 genes between the high-FEP group and the low-FEP group showed no significant differences in terms of genotype frequency (Table 1). In contrast, the SNP barcodes generated by the proposed PSO algorithm for two to five SNPs showed significantly increased OR values ($P < 0.05$) (Table 3). This suggests that gene-gene interactions involving the MET and AKT1 genes may play an important role in FEP performance. The importance of this interaction is also partly supported by evidence that MET activates AKT1 phosphorylation in an anti-apoptotic pathway [52].

By comparing the maximum difference between the high-FEP and low-FEP groups for each SNP barcode, PSO can be used to determine the relative contribution of an SNP in FEP performance. For example, the OR values between the high- and low-FEP groups for two-SNP barcodes was 2.119. Adding SNP 5 or SNP 5/SNP 3 for a three-SNP barcode and a four-SNP barcode slightly increased the OR values between the high- and low-FEP groups, which suggests that SNP 1 and SNP 4 have larger effects on FEP performance compared to SNP 5 and SNP 3. Similarly, compared to SNP2, SNP 5, and SNP 3 had larger effects on high-FEP performance. Accordingly, the relative strengths of the effects of individual SNPs on FEP performance are apparently as follows: SNP 1 (rs1130233)/SNP 4 (rs41735) > SNP 5 (rs42336)/SNP 3 (rs2237717) > SNP 2 (rs1858830).

Although this study applied the PSO algorithm in a small sample size of SNPs, the algorithm has proven effective for large datasets [53,54] such as genome-wide association studies [30,31]. Therefore, this PSO methodology is applicable for further study of SNP-SNP interactions not only in psychiatry but also in other research fields. Additionally, with the help of power calculation, the detectable minimum effect size can be

Table 3 The best SNP barcodes with the strongest associations with FEP performance

SNP ^a	Genotypes ^a	High-FEP group	Low-FEP group	Difference	P value	OR (95% CI) ^c	Power
SNPs(1,4)	3-3	30	18	12	0.028	2.119 (1.077 to 4.168)	0.528
	Other ^b	59	75				
SNPs(1,4,5)	3-3-3	27	15	12	0.023	2.265 (1.109 to 4.624)	0.556
	Other ^b	62	78				
SNPs(1,3,4,5)	3-3-3-3	25	14	11	0.032	2.204 (1.059 to 4.586)	0.500
	Other ^b	64	79				
SNPs(1,2,3,4,5)	3-1-3-3-3	11	4	7	0.048	3.138 (0.960 to 10.253)	0.400
	Other ^b	78	89				

^aFor detailed SNP data, see Table 1. ^b‘Other’ indicates reference group defined as the union of all other possible combinations of two to five SNPs. ^cDifference should be provided with footnote, i.e., difference = High-FEP group (n) – Low-FEP group (n).

predicted for suitable sample collection to improve the performance of PSO computation.

In conclusion, FEP is an essential component in social cognition and social skill not only in healthy individuals but also in patients with mental disorders such as schizophrenia [1-4]. This study confirmed that the proposed PSO algorithm can identify SNP barcodes for high-FEP performance in a healthy Taiwan population. Although further studies are needed to determine whether the findings can be extrapolated to mental disorders or to other racial populations, the FEP impact for SNPs of missing heritage was sensitively detected in this SNP-SNP interaction mode in the current study.

In conclusion, the contribution of this study is a computational methodology for analyzing complex SNP-SNP interactions in social cognition domains such as FEP performance. Although each SNP has a different impact on FEP, improved understanding of the joint effects of SNP-SNP interactions further elucidates the molecular mechanisms of facial emotion identification.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

CHY, and HWC wrote the manuscript. HYL and MTL designed the study. CHY, LYC, and YDL performed computation and conducted the statistical analysis. CHY and HWC interpreted the data. All authors read and approved the final manuscript.

Acknowledgements

This work was supported by the National Science Council, Taiwan (NSC102-2221-E-151-024-MY3, 102-2622-E-151-003-CC3, NSC101-2325-B-039-009, NSC101-2314-B-039-030-MY3, and NSC101-2221-E-214-075), National Health Research Institutes, Taiwan (NHRI-EX-101-9904NI, NHRI-101A1-PDCO-1312141, 2012-2014), Taiwan Department of Health Clinical Trial and Research Center of Excellence (DOH101-TD-B-111-004), China Medical University Hospital, Taiwan (DMR-101-082 and DMR-101-090), the Ministry of Health and Welfare, Taiwan, Republic of China (MOHW103-TD-B-111-05), and the National Sun Yat-sen University-KMU Joint Research Project (#NSYSU-KMU 103-p014).

Author details

¹Department of Chemical Engineering & Institute of Biotechnology and Chemical Engineering, I-Shou University, Kaohsiung 84001, Taiwan. ²Institute of Clinical Medical Science, China Medical University, Taichung 40402, Taiwan. ³Department of Psychiatry, China Medical University Hospital, Taichung 40402, Taiwan. ⁴Department of Electronic Engineering, National Kaohsiung University of Applied Sciences, Kaohsiung 80778, Taiwan. ⁵Department of Psychiatry, Taipei Veterans General Hospital, Hsinchu Branch, Hsinchu 31064, Taiwan. ⁶Cancer Center, Translational Research Center, Kaohsiung Medical University Hospital, Kaohsiung Medical University, Kaohsiung 80708, Taiwan. ⁷Institute of Medical Science and Technology, National Sun Yat-sen University, Kaohsiung 80424, Taiwan. ⁸Department of Biomedical Science and Environmental Biology, Kaohsiung Medical University, Kaohsiung 80708, Taiwan.

Received: 1 February 2014 Accepted: 23 April 2014

Published: 21 May 2014

References

1. Zaja RH, Rojahn J: Facial emotion recognition in intellectual disabilities. *Curr Opin Psychiatry* 2008, **21**(5):441-444.
2. Ihnen GH, Penn DL, Corrigan PW, Martin J: Social perception and social skill in schizophrenia. *Psychiatry Res* 1998, **80**(3):275-286.
3. Pinkham AE, Penn DL, Perkins DO, Graham KA, Siegel M: Emotion perception and social skill over the course of psychosis: a comparison of individuals "at-risk" for psychosis and individuals with early and chronic schizophrenia spectrum illness. *Cogn Neuropsychiatry* 2007, **12**(3):198-212.
4. Chan RC, Li H, Cheung EF, Gong QY: Impaired facial emotion perception in schizophrenia: a meta-analysis. *Psychiatry Res* 2010, **178**(2):381-390.
5. Kee KS, Green MF, Mintz J, Brekke JS: Is emotion processing a predictor of functional outcome in schizophrenia? *Schizophr Bull* 2003, **29**(3):487-497.
6. Judson MC, Eagleson KL, Levitt P: A new synaptic player leading to autism risk: Met receptor tyrosine kinase. *J Neurodev Disord* 2011, **3**(3):282-292.
7. Pantazopoulos H, Woo TU, Lim MP, Lange N, Berretta S: Extracellular matrix-glia abnormalities in the amygdala and entorhinal cortex of subjects diagnosed with schizophrenia. *Arch Gen Psychiatry* 2010, **67**(2):155-166.
8. Lai WS, Xu B, Westphal KG, Paterlini M, Olivier B, Pavlidis P, Karayiorgou M, Gogos JA: Akt1 deficiency affects neuronal morphology and predisposes to abnormalities in prefrontal cortex functioning. *Proc Natl Acad Sci U S A* 2006, **103**(45):16906-16911.
9. Burdick KE, DeRosse P, Kane JM, Lencz T, Malhotra AK: Association of genetic variation in the MET proto-oncogene with schizophrenia and general cognitive ability. *Am J Psychiatry* 2010, **167**(4):436-443.
10. Emamian ES, Hall D, Birnbaum MJ, Karayiorgou M, Gogos JA: Convergent evidence for impaired AKT1-GSK3beta signaling in schizophrenia. *Nat Genet* 2004, **36**(2):131-137.
11. Tan HY, Nicodemus KK, Chen Q, Li Z, Brooke JK, Honea R, Kolachana BS, Straub RE, Meyer-Lindenberg A, Sei Y, Mattay VS, Callicott JH, Weinberger DR: Genetic variation in AKT1 is linked to dopamine-associated prefrontal cortical structure and function in humans. *J Clin Invest* 2008, **118**(6):2200-2208.
12. Lin MT, Huang KH, Huang CL, Huang YJ, Tsai GE, Lane HY: MET and AKT genetic influence on facial emotion perception. *PLoS One* 2012, **7**(4):e36143.
13. Goodman JE, Mechanic LE, Luke BT, Ams S, Chanock S, Harris CC: Exploring SNP-SNP interactions and colon cancer risk using polymorphism interaction analysis. *Int J Cancer* 2006, **118**(7):1790-1797.
14. 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.
15. Zheng SL, Sun J, Wiklund F, Smith S, Stattin P, Li G, Adami HO, Hsu FC, Zhu Y, Balter K, Kader AK, Turner AR, Liu W, Bleecker ER, Meyers DA, Duggan D, Carpten JD, Chang BL, Isaacs WB, Xu J, Gronberg H: Cumulative association of five genetic variants with prostate cancer. *N Engl J Med* 2008, **358**(9):910-919.
16. 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.
17. 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.
18. Lavender NA, Rogers EN, Yeyeodu S, Rudd J, Hu T, Zhang J, Brock GN, Kimbro KS, Moore JH, Hein DW, Kidd LC: Interaction among apoptosis-associated sequence variants and joint effects on aggressive prostate cancer. *BMC Med Genomics* 2012, **5**(1):11.
19. Patel R, Caraviello D, Qian W: Improving LASSO performance for Grey Leaf Spot disease resistance prediction based on genotypic data by considering all possible two-way SNP interactions. *Integr Biol (Camb)* 2012, **4**(5):564-567.
20. Wang TY, Lee SY, Chen SL, Chang YH, Chen SH, Chu CH, Huang SY, Tzeng NS, Wang CL, Lee IH, Yeh TL, Yang YK, Lu RB: Interaction between serotonin transporter and serotonin receptor 1 B genes polymorphisms may be associated with antisocial alcoholism. *Behav Brain Funct* 2012, **8**:18.
21. Wu SJ, Chuang LY, Lin YD, Ho WH, Chiang FT, Yang CH, Chang HW: Particle swarm optimization algorithm for analyzing SNP-SNP interaction of renin-angiotensin system genes against hypertension. *Mol Biol Rep* 2013, **40**(7):4227-4233.
22. George Priya Doss C, Sudandiradoss C, Rajasekaran R, Choudhury P, Sinha P, Hota P, Batra UP, Rao S: Applications of computational algorithm tools to identify functional SNPs. *Funct Integr Genomics* 2008, **8**(4):309-316.

23. Chang HW, Chuang LY, Ho CH, Chang PL, Yang CH: **Odds ratio-based genetic algorithms for generating SNP barcodes of genotypes to predict disease susceptibility.** *OMICS* 2008, **12**(1):71–81.
24. 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.
25. Steen KV: **Travelling the world of gene-gene interactions.** *Brief Bioinform* 2012, **13**(1):1–19.
26. 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.
27. 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.
28. Yang CH, Chang HW, Cheng YH, Chuang LY: **Novel generating protective single nucleotide polymorphism barcode for breast cancer using particle swarm optimization.** *Cancer Epidemiol* 2009, **33**(2):147–154.
29. Moore JH, Asselbergs FW, Williams SM: **Bioinformatics challenges for genome-wide association studies.** *Bioinformatics* 2010, **26**(4):445–455.
30. 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.
31. 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.
32. Yang CH, Lin YD, Chuang LY, Chang HW: **Evaluation of breast cancer susceptibility using improved genetic algorithms to generate genotype SNP barcodes.** *IEEE/ACM Trans Comput Biol Bioinform* 2013, **10**(2):361–371.
33. Yang CH, Lin YD, Chuang LY, Chen JB, Chang HW: **MDR-ER: balancing functions for adjusting the ratio in risk classes and classification errors for imbalanced cases and controls using multifactor-dimensionality reduction.** *PLoS One* 2013, **8**(11):e79387.
34. Kennedy J, Eberhart RC: **Particle swarm optimization.** In *Proceedings of the IEEE International Conference on Neural Networks*. Edited by IEEE Neural Networks Council. Piscataway: IEEE; 1995:1942–1948.
35. Shi Y, Eberhart R: **Empirical study of particle swarm optimization.** In *Proceedings of the IEEE Congress on Evolutionary Computation*. Edited by IEEE Neural Networks Council. Piscataway: IEEE; 1999:1945–1950.
36. Ratnaweera A, Hjalgamuge SK, Watson HC: **Self-organizing hierarchical particle swarm optimizer with time-varying acceleration coefficients.** *IEEE T Evolut Comput* 2004, **8**(3):240–255.
37. Faul F, Erdfelder E, Buchner A, Lang AG: **Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses.** *Behav Res Methods* 2009, **41**(4):1149–1160.
38. Mayer JD, Salovey P, Caruso DR, Sitarenios G: **Measuring emotional intelligence with the MSCEIT V2.0.** *Emotion* 2003, **3**(1):97–105.
39. Mayer JD, Salovey PS, Caruso DR: *Mayer-Salovey-Caruso Emotional Intelligence Test User's Manual (MESCEIT) Item Booklet*. Toronto: MHS;
40. Wang K, Guo H, Hu H, Xiong G, Guan X, Li J, Xu X, Yang K, Bai Y: **A functional variation in pre-microRNA-196a is associated with susceptibility of esophageal squamous cell carcinoma risk in Chinese Han.** *Biomarkers* 2010, **15**(7):614–618.
41. Chaudhary AK, Pandya S, Mehrotra R, Bharti AC, Jain S, Singh M: **Functional polymorphism of the MMP-1 promoter (-1607 1G/2G) in potentially malignant and malignant head and neck lesions in an Indian population.** *Biomarkers* 2010, **15**(8):684–692.
42. Yang D, Zhou F, Wang X, Gao H, Li G, Xue M: **Association analysis between MDR1 gene polymorphisms and risk of hepatocellular carcinoma in Chinese population.** *Biomarkers* 2013, **18**(3):236–241.
43. Cheng P, Chen H, Liu SR, Pu XY: **A ZC: SNPs in KIT and KITLG genes may be associated with oligospermia in Chinese population.** *Biomarkers* 2013, **18**(8):650–654.
44. Chang WC, Woon PY, Hsu YW, Yang S, Chiu YC, Hou MF: **The association between single-nucleotide polymorphisms of ORA1 gene and breast cancer in a Taiwanese population.** *Sci World J* 2012, **2012**:916587.
45. Marian AJ: **Elements of 'missing heritability'.** *Curr Opin Cardiol* 2012, **27**(3):197–201.
46. Ehret GB, Lamparter D, Hoggart CJ, Whittaker JC, Beckmann JS, Kutalik Z: **A multi-SNP locus-association method reveals a substantial fraction of the missing heritability.** *Am J Hum Genet* 2012, **91**(5):863–871.
47. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorf LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttmacher AE, Kong A, Kruglyak L, Mardis E, Rotimi CN, Slatkin M, Valle D, Whittemore AS, Boehnke M, Clark AG, Eichler EE, Gibson G, Haines JL, Mackay TF, McCarrroll SA, Visscher PM: **Finding the missing heritability of complex diseases.** *Nature* 2009, **461**(7265):747–753.
48. Nakajima M, Hattori E, Yamada K, Iwayama Y, Toyota T, Iwata Y, Tsuchiya KJ, Sugihara G, Hashimoto K, Watanabe H, Iyo M, Hoshika A, Yoshikawa T: **Association and synergistic interaction between promoter variants of the DRD4 gene in Japanese schizophrenics.** *J Hum Genet* 2007, **52**(1):86–91.
49. Norton N, Moskvina V, Morris DW, Bray NJ, Zammit S, Williams NM, Williams HJ, Preece AC, Dwyer S, Wilkinson JC, Spurlock G, Kirov G, Buckland P, Waddington JL, Gill M, Corvin AP, Owen MJ, O'Donovan MC: **Evidence that interaction between neuregulin 1 and its receptor erbB4 increases susceptibility to schizophrenia.** *Am J Med Genet B Neuropsychiatr Genet* 2006, **141B**(1):96–101.
50. Shiota S, Tochigi M, Shimada H, Ohashi J, Kasai K, Kato N, Tokunaga K, Sasaki T: **Association and interaction analyses of NRG1 and ERBB4 genes with schizophrenia in a Japanese population.** *J Hum Genet* 2008, **53**(10):929–935.
51. Nicodemus KK, Law AJ, Radulescu E, Luna A, Kolachana B, Vakkalanka R, Rujescu D, Giegling I, Straub RE, McGee K, Gold B, Dean M, Muglia P, Callicott JH, Tan HY, Weinberger DR: **Biological validation of increased schizophrenia risk with NRG1, ERBB4, and AKT1 epistasis via functional neuroimaging in healthy controls.** *Arch Gen Psychiatry* 2010, **67**(10):991–1001.
52. Xiao GH, Jeffers M, Bellacosa A, Mitsuuchi Y, Vande Woude GF, Testa JR: **Anti-apoptotic signaling by hepatocyte growth factor/Met via the phosphatidylinositol 3-kinase/Akt and mitogen-activated protein kinase pathways.** *Proc Natl Acad Sci U S A* 2001, **98**(1):247–252.
53. Chuang LY, Yang CH, Li JC: **A hybrid BPSO-CGA approach for gene selection and classification of microarray data.** *J Comput Biol* 2012, **19**(1):68–82.
54. Mohamad MS, Omatu S, Deris S, Yoshioka M: **A modified binary particle swarm optimization for selecting the small subset of informative genes from gene expression data.** *IEEE Trans Inf Technol Biomed* 2011, **15**(6):813–822.

doi:10.1186/1744-859X-13-15

Cite this article as: Chuang et al.: Identification of SNP barcode biomarkers for genes associated with facial emotion perception using particle swarm optimization algorithm. *Annals of General Psychiatry* 2014 **13**:15.

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

