BMC Genetics

METHODOLOGY ARTICLE

Open Access

CrossMark

Aggregate blood pressure responses to serial dietary sodium and potassium intervention: defining responses using independent component analysis

Gengsheng Chen[1], Lisa de las Fuentes[1], Chi C. Gu[1], Jiang He[2], Dongfeng Gu[3], Tanika Kelly[2], James Hixson[4], Cashell Jacquish[5], D. C. Rao[1] and Treva K. Rice[1]*

Abstract

Background: Hypertension is a complex trait that often co-occurs with other conditions such as obesity and is affected by genetic and environmental factors. Aggregate indices such as principal components among these variables and their responses to environmental interventions may represent novel information that is potentially useful for genetic studies.

Results: In this study of families participating in the Genetic Epidemiology Network of Salt Sensitivity (GenSalt) Study, blood pressure (BP) responses to dietary sodium interventions are explored. Independent component analysis (ICA) was applied to 20 variables indexing obesity and BP measured at baseline and during low sodium, high sodium and high sodium plus potassium dietary intervention periods. A "heat map" protocol that classifies subjects based on risk for hypertension is used to interpret the extracted components. ICA and heat map suggest four components best describe the data: (1) systolic hypertension, (2) general hypertension, (3) response to sodium intervention and (4) obesity. The largest heritabilities are for the systolic (64 %) and general hypertension (56 %) components. There is a pattern of higher heritability for the component response to intervention (40–42 %) as compared to those for the traditional intervention responses computed as delta scores (24 %–40 %).

Conclusions: In summary, the present study provides intermediate phenotypes that are heritable. Using these derived components may prove useful in gene discovery applications.

Keywords: Blood pressure, Cardiovascular diseases, Hypertension, Independent component analysis

Background

Hypertension is a major risk factor for many cardiovascular diseases. [1, 2] Moreover, there is evidence for substantial heritability as well as environmental causes underlying these diseases. [3–5] For example, an environmental intervention of reduced sodium and/or increased potassium intake can lead to reduced systolic and diastolic blood pressure (BP) [6–11], and familial factors in part underlie these BP responses to sodium intervention. [11–16] Moreover, significant inter-correlations among this set of variables suggest a complex underlying network involving multiple genes, multiple environments and gene-by-environment interactions [17]. However, the structure and mechanisms underlying such complex networks are unclear. Recent studies investigating these issues focus on endophenotypes, defined here as intermediate traits that are derived from observed measures [18, 19]. Such endophenotypes may provide clues to the mechanisms and genetic architecture underlying clinical diseases, and factor analysis is a tool that may be used to extract these hidden components.

Traditionally, principal component analysis (PCA) [20] is used to identify uncorrelated factors. PCA assumes mutivariate normality among the variables. However, if multivariate normality is violated the resulting components may not be independent and thus not uniquely

* Correspondence: treva@wubios.wustl.edu
[1]Washington University in St. Louis School of Medicine, St. Louis, MO, USA
Full list of author information is available at the end of the article

Chen *et al. BMC Genetics* (2015) 16:64

Page 2 of 10

interpretable. The related method of independent component analysis (ICA) [21] produces uncorrelated and independent component, even when multivariate normality is violated [22, 23]. Consequently, in the investigation reported here, ICA is used to construct factors that may serve as endophenotypes to help identify the underlying genetic architecture.

The purpose of this report is to construct endophenotypes using ICA analysis of several obesity and BP measures during a sodium intervention test in the GenSalt study, and assess their utility using heritability analysis. GenSalt is a dietary sodium/potassium intervention study on blood pressure levels in rural Chinese families [24]. After baseline, there were 3 intervention periods (low and high sodium and a potassium supplement). ICA analysis extracted 4 endophenotypes (component factors) which were clinically interpreted using a heat map protocol. The utilities of the endophenotypes for future genetic applications were evaluated by heritability analysis.

Methods
Study subjects
The GenSalt study was designed to identify genes related to variation in BP responses to dietary sodium [24]. The study was conducted in six rural Chinese provinces around Beijing using probands (i.e. high BP) and their family members. Eligible probands (see [24]) were 18–60 years of age, had high BP (SBP ≥ 130 mm Hg and/or DBP ≥ 85 mm Hg), and had at least 1 parent and at least 1 sibling who could participate, although additional family members (spouses and offspring) also could participate. Eligible siblings, spouses and offspring were aged 18–60 years (≥16 for offspring) and lived in the same village with the proband. Exclusion criteria for all participants are listed elsewhere [24] but generally exclude those with stage 2 hypertension or on antihypertensive medications, having a history of cardiovascular disease, diabetes, kidney disease, or liver disease, who were pregnant or heavy alcohol drinkers, or on a low-sodium diet. Only probands and their siblings, spouses and offspring (i.e. not parents) participated in the dietary intervention. A total of 3,150 individuals from 658 families were included in the study at baseline, but only 1,906 individuals participated in the dietary sodium intervention. The reduced sample size (1,906 versus 3,150) primarily is due to the parents of the probands (older generation) not participating in the intervention by design. The rational for this study design (omitting probands parents from the intervention) was that the siblings of the probands would provide the most useful information [24]. Isolated missing data occurred for 64 individuals, resulting in 1,842 individuals with complete data available for analyses.

This study protocol was approved by the Institutional Review Boards at all participating institutions, including Washington University in St. Louis and Tulane University, and all participants gave their informed consent.

Dietary sodium intervention
In the first 3-day baseline observation phase participants consumed their usual diet. Then, participants received a low sodium diet (51.3 mmol of sodium per day) for 1 week, followed by a high sodium diet (307.8 mmol of sodium per day) for 1 week, and finally a high sodium diet plus a potassium supplement (60 mmol potassium) for 1 week, see [24] for further details about the intervention. BP was measured (see below) during baseline (B) and during each of the low sodium (L), high sodium (H), and potassium supplement (K) intervention periods.

Phenotypes
A variety of demographic, pedigree, medical history, lifestyle risk factor, anthropometric, blood and urine data was collected on all study participants. For the current analysis, a total of 20 phenotypes (plus age) are used (see Table 1). There are 2 anthropomorphic measures (body mass index (BMI) and waist circumference (WST]) and 12 BP measures for systolic (SBP), diastolic (DBP) and pulse pressure (PP = SBP-DBP). BP was measured by trained and certified observers according to a common protocol adapted from the American Heart Association [25]. Participants were seated and BP was measured after a 5-min rest using a random-zero sphygmomanometer with an appropriate cuff size. Participants were advised to avoid alcohol, cigarettes, coffee/tea and exercise for at least 30 min prior to their BP assessment. At baseline (B), BP was measured during the morning and averaged across the 3 baseline days (B_SBP, B_DBP, and B_PP). During each of the intervention weeks, BP was measured on the mornings of days 2, 5, 6, and 7. However, the analysis variable was the average across only days 5, 6, and 7 (L_SBP, L_DBP and L_PP for low-sodium week, H_SBP, H_DBP and H_PP for high sodium week, and K_SBP, K_DBP and K_PP for potassium supplement week).

Statistical analysis
Age adjustments and distributional properties The anthropometric and hemodynamic variables are adjusted for a polynomial in age (age, age^2, age^3) separately by gender groups after exploring distributional properties. First, sparse outliers (see definition in supplemental material) were temporarily set aside. A given phenotype was regressed on a polynomial in age retaining only significant terms. Age effects in the residual variance (heteroscedasticity) were also examined. A given analysis phenotype was the standardized residual from this regression. Skewness and kurtosis of these standardized variables was also explored to ensure each variable was normally distributed.

Table 1 Variable characteristics

Variable	Males (N = 974)			Females (N = 868)		
	Mean	SD	SE	Mean	SD	SE
Demographic						
AGE	39.2	9.5	0.3	38.2	9.3	0.3
Anthropometric						
B2_WST	81.9	9.9	0.3	78.5	9.5	0.3
B_BMI	23.2	3.1	0.1	23.5	3.2	0.1
Hemodynamic						
B_SBP	118.6	12.5	0.4	115.0	15.5	0.5
B_DBP	75.6	9.8	0.3	71.8	10.5	0.3
L_SBP	113.3	11.1	0.4	109.3	13.1	0.4
L_DBP	73.2	9.3	0.3	68.6	9.6	0.3
H_SBP	117.8	12.1	0.4	114.5	15.0	0.5
H_DBP	74.5	9.9	0.3	71.1	10.3	0.4
K_SBP	114.1	11.6	0.4	111.1	14.3	0.5
K_DBP	73.4	9.5	0.3	69.4	9.9	0.3
B_PP	43.0	9.0	0.3	43.2[a]	9.8	0.3
L_PP	40.1	8.9	0.3	40.7[a]	9.4	0.3
H_PP	43.2	8.8	0.3	43.4[a]	9.5	0.3
K_PP	40.8	8.9	0.3	41.7	9.6	0.3
D_LH_SBP	4.4	5.9	0.2	5.3	6.1	0.2
D_LH_DBP	1.3	5.5	0.2	2.5	5.3	0.2
D_HK_SBP	−3.6	5.2	0.2	−3.4	5.8	0.2
D_HK_DBP	−1.2	4.7	0.2	−1.7	4.4	0.2
D_BL_SBP	−5.3	6.9	0.2	−5.7[a]	7.1	0.2
D_BL_DBP	−2.4	5.7	0.2	−3.2	5.3	0.2

Abbreviations: SD = standard deviation; SE = standard error; AGE = age in years; SBP = Systolic Blood Pressure (mm Hg); DBP = Diastolic Blood Pressure (mm Hg); PP = Pulse Pressure = SBP-DBP; BMI = Body Mass Index (kg/m^2); WST = waist circumference (cm); B_ = Baseline; L_ = Low sodium intervention; H_ = High sodium intervention; K_ = Potassium supplement intervention; D_BL_ = response (delta), Low minus Baseline; D_LH_ = response (delta), High minus Low sodium; D_HK_ = response (delta), Potassium minus High sodium
[a]All mean values are significantly different across genders at 0.05 except as noted by asterisk using standard error comparison

ICA and PCA Independent components analysis (ICA) is used to identify independent components using correlated variables. The 20 correlated variables include: SBP, DBP and PP, each measured at baseline and during each of the 3 intervention periods (N = 12 BP variables), BMI and waist circumference (N = 2 variables), and the SBP and DBP responses to adjacent intervention weeks (i.e. baseline to low, low to high, and high to potassium supplement, N = 6). The software R (fastICA) [21] is used to extract the independent components.

The optimal number of components will be determined using the screen criteria. A preliminary principal components analysis (PCA, using SAS Proc Factor) [20, 26] is performed for the purpose of generating the scree plot [27] (see supplemental material). The scree method plots the relationship between the eigenvalues and the number of factors. The optimal number of factors is where the line stops its precipitous descent and levels off. Each ICA is subscripted by a first digit to denote the predefined number of components (2, 3 or 4), and a second digit to label the individual component (i.e. ICA31 represents the first component of a 3-component solution).

Identification of component hypertension traits Several methods are used to define the component traits. First, loadings of each component on the original phenotypes are used to identify enriched representation of any particular phenotypic groups. Second, clinical characteristics of each component are classified using a heat map (Excel, Microsoft Office, version 2010, Redmond, WA). The heat map depicts the means of the original 20 phenotypes separately for two groups, with high risk values shown in red and low risk values in green. The brightness
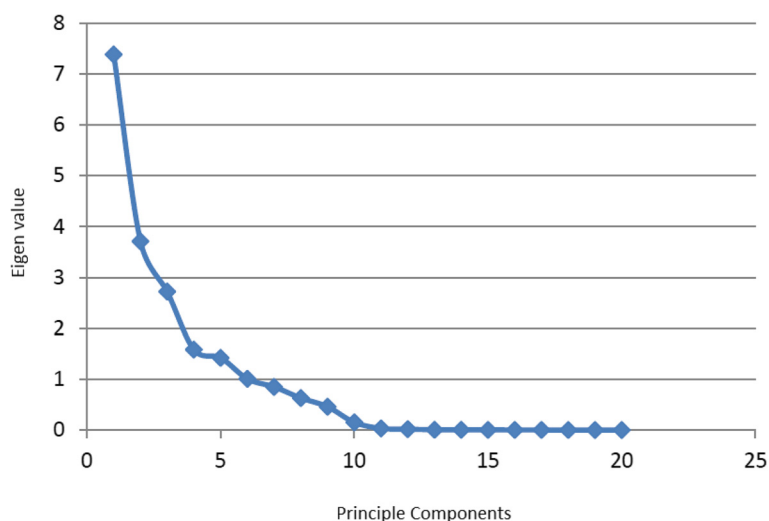
Chen *et al. BMC Genetics* (2015) 16:64

Page 4 of 10

of the colors is determined by within-variable comparisons across groups where more intense colors represent the extremes of the variable range.

Analysis of heritability Heritability of the adjusted and standardized variables is computed using maximum likelihood methods under a variance components model as parameterized in the computer program QTDT [28]. In this model the residual phenotypic variance is stratified into an additive familial effect (heritability) and the remaining non-familial effects. Heritability is defined as the percentage of the total variance that is due to familial factors. QTDT allows for complex family structures (i.e. extended pedigrees). GenSalt family structures ranged from 2- to 3-generations with the total number of individuals in each pedigree ranging from 1 through 22. Covariates in the heritability model include gender, generation, field center and average room temperature (i.e. these effects were partitioned out prior to computing heritability).

Results

The descriptive statistics for the variables are presented in Table 1 (separately by males and females) and in Additional file 1: Table S2 of the supplemental materials. As shown, there are 974 males and 868 females. In general, females have a larger BMI, are younger and have lower BP levels than males. Consequently, age adjustments are performed separately for males and females. Age is a significant covariate for most of the variables (see Additional file 1: Table S1 in supplemental materials) accounting for up to 19 % of the phenotypic variation. Also, the mean BP responses for probands are larger than those for the other groups (spouses, siblings and offspring), as expected since probands were selected for high blood pressure. Additionally, there were few deviations from normality based on the skewness and kurtosis of the individual analysis variables (see Additional file 1: Table S4 in supplemental material).

Factor interpretation

Table 2 shows the correlation coefficients among the ICA components and Fig. 1 shows the scree plot. The scree plot clearly suggests 4 factors best represent the data (i.e. the eigen values begin to level off after 4 components). However, we present results for 2-, 3-, and 4-component solutions in order to evaluate the stability of the preferred 4-component solution. Fig. 2 depicts the graphic loadings for each of the 2-, 3-, and 4-component solutions, and Fig. 3 shows the heat map for each of the components and solutions. In Table 2, within each 2- or 3- or 4-component solution the off-diagonal elements are zero as expected since the components within each solution are derived to be independent. Informative correlations occur between solutions (e.g. between 2- and 3-components). High inter-correlations suggest that the same underlying constructs replicate across different solutions. The first pattern is among ICA21, ICA31 and ICA42; the second is among ICA22, ICA33 and ICA43; and the third is between ICA32 and ICA41. ICA44 is found only in the 4-factor solution. Thus, the "same" constructs reoccur whether there are 2-, 3- or 4-components, and suggest that the 4-component solution is stable. This "stability" is further demonstrated below where the constructs underlying the factors is discussed.

Systolic hypertension (ICA21, ICA31, ICA42) A comparison of the loadings for this factor (whether under the 2-, 3-, or 4-component solution, Fig. 2, top row) consistently shows more extreme loadings (either positive or negative) for all of the systolic and pulse pressure variables. Note that the positive or negative directions of the extreme loadings and the component factor values are arbitrary in ICA analysis. The heat map (Fig. 3) provides corroborating information regarding interpretation of the factor. That is, there is a strong red-green contrast between the lower and upper component factor median groups for all of the SBP and PP traits. Moreover, this pattern is consistent across the three

Table 2 Inter-correlation coefficients for different ICA components

	ICA21	ICA22	ICA31	ICA32	ICA33	ICA41	ICA42	ICA43	ICA44
ICA21	1	0	−0.980	0.141	0.102	0.120	−0.970	0.072	0.197
ICA22		1	−0.037	0.400	−0.915	0.452	−0.090	−0.773	−0.434
ICA31			1	0	0	0.014	0.990	0.010	−0.130
ICA32				1	0	0.993	0.000	0.002	0.109
ICA33					1	−0.060	0.057	0.845	0.527
ICA41						1	0	0	0
ICA42							1	0	0
ICA43								1	0
ICA44									1

Fig. 1 Scree plot of Eigen values. Although there are 6 eigenvalues greater than 1.0, the graphical Scree plot indicates 4 components may best represent the data

solutions, suggesting the construct is robust regardless of the number of components extracted. Given the systolic and pulse pressure variables loading on this factor and the heat map, it is labeled a "systolic hypertension" construct.

Generalized hypertension (ICA22, ICA33, ICA43) Although these inter-correlations are on average somewhat lower than those as for the first construct, the basic pattern is consistent across different solutions. The factor loadings and heat map across the 3 solutions consistently show high (positive or negative) loadings on the eight SBP and DBP measures. However, there are additional moderate contributions from other variables for two of the factors. ICA22 has additional moderate contributions from the adiposity variables and delta BP responses, ICA33 shows moderate additional contributions from the adiposity variables, and ICA43 has little additional contributions. Thus, as the number of extracted factors increases from 2 through 4, there is a pattern of increasing homogeneity among the variables. This component represents "generalized hypertension" whose precise interpretation depends on the number of factors extracted.

Response to sodium intervention (ICA32, ICA41) This factor was detected only when 3 and 4 components were extracted. Both the factor loadings and the heat map suggest this factor primarily reflects the six delta response variables. It is interesting that the contrasting directions of the loadings for the response variables in this component are consistent with their effects on blood pressure. That is, the loadings for variables that tend to increase BP (LH_) are in one direction while those that tend to lower BP (BL_ and HK_) are in the other. Both factor loadings and heat map show additional moderate contributions from the

high-sodium BP variables (B_ and H_). The group with the more extreme responses also has somewhat higher baseline and high-sodium BP levels. Thus, this construct represents a "response to sodium intervention".

Adiposity Finally, one factor (ICA44) was detected only when 4 components were extracted. The factor loadings (Fig. 2) and heat map (Fig. 3) suggest this factor primarily represents an "adiposity" construct. When fewer components were extracted, the adiposity variables loaded moderately on the "generalized hypertension" factor.

Heritability
Heritabilities of the individual BP traits and ICA components are shown in Tables 3 and 4. In general, the heritabilities are higher in the intervention conditions (low sodium, high sodium and high sodium + potassium supplement) than in the baseline condition (Table 3). And, heritabilities for PP pressure tend to be higher than those for the corresponding SBP and DBP (Table 3). In general, heritabilities tend to be larger for the components as compared to the individual variables. For example, heritability for the systolic hypertension (64 %) and the generalized hypertension components (as high as 56 %) (Table 4) tend to be higher than those for SBP and DBP (23–53 %) (Table 3), although they are strictly significantly higher (based on standard error comparison) only for the delta SBP responses to low and high sodium conditions. Similarly, the heritabilities for the response to intervention component (40–42 %) tend to be higher than those for the delta variables (24–40 %). Another interesting pattern is seen for the generalized hypertension component. Here, the heritabilities increase (from 34 % to 44 % to 56 %) as the number of extracted components increases from 2 to 3 to 4.
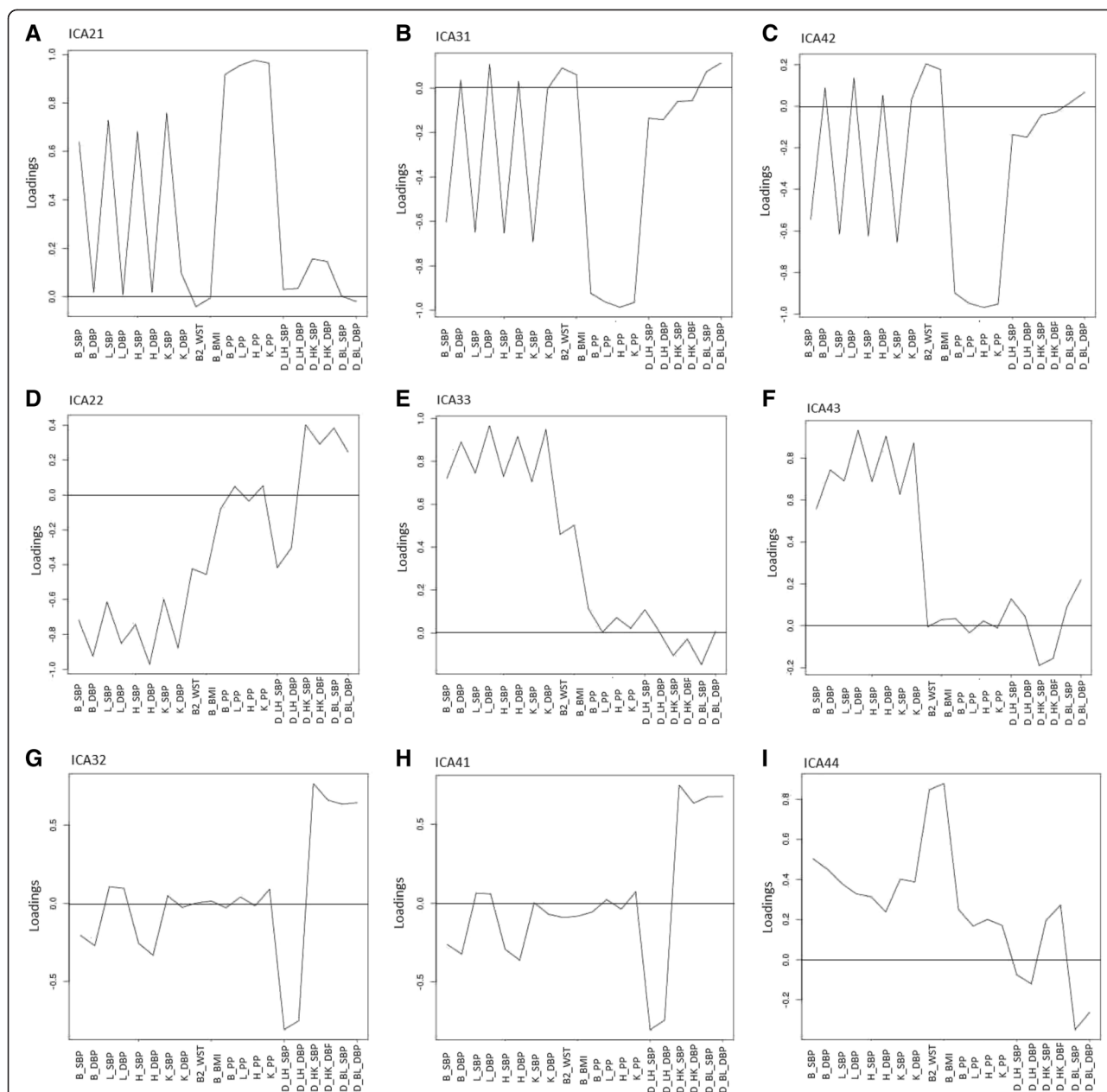
Chen *et al. BMC Genetics* (2015) 16:64

Page 6 of 10

Fig. 2 Graphic loadings for the 2-, 3- and 4-component solutions for ICA. The top row (panels **a**, **b**, and **c**) represent solutions for the "Systolic Hypertension" component (A = ICA21 = 1st factor of the 2-factor solution, B = ICA31 = 1st factor of the 3-factor solution, and C = ICA42 = 2nd factor of the 4-factor solution). Similarly, the middle row (panels **d**, **e**, and **f**) represent solutions for the "Generalized Hypertension" component (D = 2nd factor of the 2-factor solution, E = 3rd factor of the 3-factor solution, and F = 3rd factor of the 4-factor solution). The first two panels of the bottom row (**g** and **h**) represent the "Response to Intervention" component (G = 2nd factor of the 3-factor solution and H = 1st factor of the 4-factor solution). Finally, panel **i** represented the "Adiposity" component (4th factor of the 4-factor solution)

Discussion

In this study, ICA was used to explore the factor structure underlying the BP responses to a longitudinal sodium intervention study. Factors representing four basic underlying constructs were derived and were relatively stable across 2-, 3- and 4-factor solutions. Heritability analysis revealed that familial factors accounted for 40 % to 60 % of the variance in the component factors. Thus, these new

endophenotypes have the potential to provide novel information about genetic architecture underlying sodium-dependent variability in hypertension, and may enhance gene discovery efforts.

ICA was chosen over PCA to extract factors since any departures from multivariate normality would result in factors that, although uncorrelated, may not be independent. This would yield factors that would be difficult to interpret
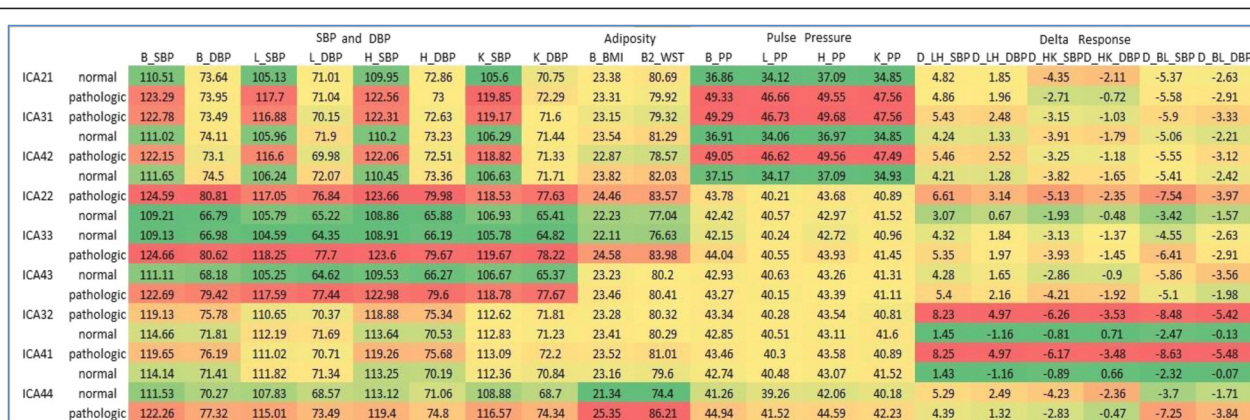
Chen et al. BMC Genetics (2015) 16:64

Page 7 of 10

		SBP and DBP								Adiposity		Pulse Pressure				Delta Response					
		B_SBP	B_DBP	L_SBP	L_DBP	H_SBP	H_DBP	K_SBP	K_DBP	B_BMI	B2_WST	B_PP	L_PP	H_PP	K_PP	D_LH_SBP	D_LH_DBP	D_HK_SBP	D_HK_DBP	D_BL_SBP	D_BL_DBP
ICA21	normal	110.51	73.64	105.13	71.01	109.95	72.86	105.6	70.75	23.38	80.69	36.86	34.12	37.09	34.85	4.82	1.85	-4.35	-2.11	-5.37	-2.63
	pathologic	123.29	73.95	117.7	71.04	122.56	73	119.85	72.29	23.31	79.92	49.33	46.66	49.55	47.56	4.86	1.96	-2.71	-0.72	-5.58	-2.91
ICA31	pathologic	122.78	73.49	116.88	70.15	122.31	72.63	119.17	71.6	23.15	79.32	49.29	46.73	49.68	47.56	5.43	2.48	-3.15	-1.03	-5.9	-3.33
	normal	111.02	74.11	105.96	71.9	110.2	73.23	106.29	71.44	23.82	81.29	36.91	34.06	36.97	34.85	4.24	1.33	-3.91	-1.79	-5.06	-2.21
ICA42	pathologic	122.15	73.1	116.6	69.98	122.06	72.51	118.82	71.33	22.87	78.57	49.05	46.62	49.56	47.49	5.46	2.52	-3.25	-1.18	-5.55	-3.12
	normal	111.65	74.5	106.24	72.07	110.45	73.36	106.63	71.71	23.82	82.03	37.15	34.17	37.09	34.93	4.21	1.28	-3.82	-1.65	-5.41	-2.42
ICA22	pathologic	124.59	80.81	117.05	76.84	123.66	79.98	118.53	77.63	24.46	83.57	43.78	40.21	43.68	40.89	6.61	3.14	-5.13	-2.35	-7.54	-3.97
	normal	109.21	66.79	105.79	65.22	108.86	65.88	106.93	65.41	22.23	77.04	42.42	40.57	42.97	41.52	3.07	0.67	-1.93	-0.48	-3.42	-1.57
ICA33	normal	109.13	66.98	104.59	64.35	108.91	66.19	105.78	64.82	22.11	76.63	42.15	40.24	42.72	40.96	4.32	1.84	-3.13	-1.37	-4.55	-2.63
	pathologic	124.66	80.62	118.25	77.7	123.6	79.67	119.67	78.22	24.58	83.98	44.04	40.55	43.93	41.45	5.35	1.97	-3.93	-1.45	-6.41	-3.56
ICA43	normal	111.11	68.18	105.25	64.62	109.53	66.27	106.67	65.37	23.23	80.2	42.93	40.63	43.26	41.31	4.28	1.65	-2.86	-0.9	-5.86	-3.56
	pathologic	122.69	79.42	117.59	77.44	122.98	79.6	118.78	77.67	23.46	80.41	43.27	40.15	43.39	41.11	5.4	2.16	-4.21	-1.92	-5.1	-1.98
ICA32	pathologic	119.13	75.78	110.65	70.37	118.88	75.34	112.62	71.81	23.28	80.32	43.34	40.28	43.54	40.81	8.23	4.97	-6.26	-3.53	-8.48	-5.42
	normal	114.66	71.81	112.19	71.69	113.64	70.53	112.83	71.23	23.41	80.29	42.85	40.51	43.11	41.6	1.45	-1.16	-0.81	0.71	-2.47	-0.13
ICA41	pathologic	119.65	76.19	111.02	70.71	119.26	75.68	113.09	72.2	23.52	81.01	43.46	40.3	43.58	40.89	8.25	4.97	-6.17	-3.48	-8.63	-5.48
	normal	114.14	71.41	111.82	71.34	113.25	70.19	112.36	70.84	23.16	79.6	42.74	40.48	43.07	41.52	1.43	-1.16	-0.89	0.66	-2.32	-0.07
ICA44	normal	111.53	70.27	107.83	68.57	113.12	71.06	108.88	68.7	21.34	74.4	41.26	39.26	42.06	40.18	5.29	2.49	-4.23	-2.36	-3.7	-1.71
	pathologic	122.26	77.32	115.01	73.49	119.4	74.8	116.57	74.34	25.35	86.21	44.94	41.52	44.59	42.23	4.39	1.32	-2.83	-0.47	-7.25	-3.84

Fig. 3 Heat map for 2-, 3- and 4-component solutions. These maps are grouped by pattern for ease of interpretation. For each extracted component, the cohort was sorted by ICA scores and split into two equal subsamples by the median; thus, the two bars for a given component represents the lowest and highest ICA scores, labeled low risk and high risk in the figure. For each median group, the average scores for each of the 20 variables are provided. High risk means for the individual variables (e.g. higher BP values) are depicted as red and low risk means are shown in green. Brightness of color denotes more extreme (high or low) scores

clinically and less likely to represent endophenotypes that would be useful in genetic applications. The component constructs arising from this study appeared to be clinically relevant and were consistent across the alternative 2-, 3-, and 4-factor solutions. Clinical interpretation of each factor was aided by the results from the heat map which discriminated among individuals based on the means of the extracted factor value.

The first derived factor was labeled systolic hypertension and provided the strongest contrasts for SBP and PP means, across all background sodium levels. Clinically, systolic hypertension is characterized by elevated SBP (i.e. > 140), and is further classified as isolated systolic hypertension if DBP is normal (i.e. < 90) [29]. As reviewed by Kaplan [30], a larger pulse pressure reflects a greater systolic to diastolic difference and this is a major cardiovascular risk factor in the elderly caused in part by reduced contractility of major blood vessels. The component factor identified in the GenSalt study does not reach the recommended levels for a clinical definition because the probands in this sample, although having high BP, were not clinically hypertensive. However, it does distinguish among individuals with isolated higher systolic and pulse pressures. Further, the current study suggests that there is a very strong genetic component for this component factor that accounts for nearly 65 % of the total phenotypic variance.

The second construct was generalized hypertension (ICA22, ICA33 and ICA43). The heat map shows that this factor primarily discriminates among individuals based on SBP and DBP means, although additional variables also had moderate loadings depending on the number of factors extracted. For example, when only 2 factors were extracted (ICA22) there were additional moderate loadings for both adiposity and delta responses to the intervention. When 3 factors were extracted (ICA33), the delta response variables were removed from the hypertension factor and formed a third factor. Similarly, the addition of a 4th factor (ICA43) removed the adiposity variables to a new factor, leaving the hypertension factor a more homogeneous construct reflecting only SBP and DBP traits.

As reviewed by Delles et al. [31], while hypertension is influenced by genetic factors, the environment can have a powerful modulating effect which influences the magnitude of the genetic effect. This was demonstrated in the current study for the generalized hypertension factor. As the number of extracted factors increased, the number of variables that loaded on the factor decreased with an accompanying increase in the magnitude of the heritability and presumed homogeneity of the factor. The heritability for the most homogeneous solution (ICA43, 4-components) was significantly higher at 56 % than that for the most heterogeneous factor (ICA22, 2-components) at 34 %. This change was a

Table 3 Heritabilities (+/- Standard Error) of Systolic, Diastolic and Pulse Pressure under Different Interventions

Intervention	Systolic (SBP) Heritability	Diastolic (DBP) Heritability	Pulse Pressure (PP) Heritability
Baseline (B)	0.23 (0.10)	0.25 (0.10)	0.43 (0.08)
Low Sodium (L)	0.40 (0.08)	0.50 (0.07)	0.56 (0.07)
High Sodium (H)	0.34 (0.08)	0.43 (0.08)	0.52 (0.07)
High Sodium + Potassium Supplement (K)	0.53 (0.07)	0.48 (0.07)	0.63 (0.07)

Chen *et al. BMC Genetics* (2015) 16:64

Page 8 of 10

Table 4 Heritabilities (± Standard Error) of Responses and Component Traits

Construct	Variable	Heritability	V_e	V_g
Responses to Intervention (Deltas)				
Response to Low Sodium Intervention	Δ (L-B) SBP	0.26 (0.10)	0.84	0.30
	Δ(L-B) DBP	0.31 (0.09)	0.76	0.35
Response to High Sodium Intervention	Δ (H-L) SBP	0.24 (0.11)	0.86	0.27
	Δ (H-L) DBP	0.40 (0.08)	0.64	0.43
Response to Potassium Supplement Intervention	Δ (K-H) SBP	0.39 (0.08)	0.73	0.47
	Δ (K-H) DBP	0.34 (0.09)	0.72	0.38
Adiposity				
Body Mass Index	B_BMI	0.60 (0.07)	0.43	0.65
Waist Circumference	B_WST	0.52 (0.07)	0.51	0.55
Component Traits (ICA)				
Systolic Hypertension	ICA21	0.64 (0.06)	0.36	0.63
	ICA31	0.63 (0.06)	0.37	0.63
	ICA42	0.64 (0.06)	0.36	0.64
Generalized Hypertension + anthropometric & delta	ICA22	0.34 (0.09)	0.66	0.34
+ anthropometric	ICA33	0.44 (0.08)	0.55	0.44
	ICA43	0.56 (0.06)	0.43	0.55
Response to Intervention	ICA32	0.42 (0.07)	0.58	0.42
	ICA41	0.40 (0.08)	0.60	0.40
Adiposity	ICA44	0.49 (0.07)	0.51	0.48

Heritability (h^2) is computed as (V_g) / ($V_e + V_g$), where V_e is environmental variance and V_g is genetic variance

function of both decreasing environmental variance (0.66 to 0.43) and increasing genetic variance (0.34 to 0.56, respectively) as the number of extracted factors increased. The utility of this new endophenotype as a predictor of future cardiovascular risk with a substantial genetic component should be further investigated.

The third construct is a response to the sodium intervention. This factor was stable across both the 3- and 4-factor solutions but did not appear in the 2-factor solution. The positive versus negative means in the heat map and factor loadings is consistent with the direction of effects on BP in the intervention. That is, loadings and means for the responses that lower sodium (HK_ and BL_) are opposite to those that increase sodium levels (LH_). Moreover, the group with the larger (positive or negative) responses also has somewhat higher SBP and DBP means. This suggests that individuals who are more sensitive to the sodium intervention also have higher pressures on high sodium diets, as may be expected. Genetic factors underlying this endophenotype account for over 42 % of the variance, suggesting it should be investigated to isolate the causal genes.

The final construct in this study was labeled adiposity. Both the body mass index and the waist circumference measures loaded heavily on this factor. However, the factor also includes smaller loadings for most of the BP traits and the delta responses. The heritability for this construct is 49 %, as compared to heritabilities of 60 % for BMI and

52 % for the waist circumference. As recently reviewed [32], the heritability for BMI ranges from 30–50 % in nuclear families to as high as 80 % in twin studies [33].

Heritabilities for the baseline BP traits in GenSalt (23 % to 25 %) are within range of those reported in the literature. For example, previous family studies from various ethnic backgrounds [3, 34–37], including other Chinese family studies [38] suggest between 20 % to over 50 % of the variance in SBP and DBP is due to familial factors. The estimates reported here also are very comparable to previous GenSalt reports (23 % to 32 %) in which the entire baseline sample of over 3,000 individuals were used [5] as compared to only ~1,800 individuals who completed the intervention protocol in the current report.

It is also noteworthy that the BP heritabilities, standardized for the background levels of sodium (i.e. L, H, K interventions), were nearly doubled as compared to those for the unstandardized BPs (at baseline). This pattern was strongest for SBP (increased from 23 % to as high as 53 %) and DBP (from 25 % to 48 %), but was also noted for PP (43 % to over 60 %). As previously indicated [31], the environment can have a powerful modulating effect, which can in turn influence the magnitude of the genetic effect (i.e. gene by environment interaction) as was dramatized here.

There are limitations in this study. First, while PCA analysis always extracts as the first factor the one that accounts for the most variance among the inter-correlated traits,

Chen *et al. BMC Genetics* (2015) 16:64

Page 9 of 10

ICA does not order the extracted components. The ICA also does not provide statistics for determining the optimal number of components to extract. However, a priori clinical knowledge about the specific disease under study can assist in making a meaningful selection of traits and in providing information about the optimal number of factors. The preliminary use of exploratory PCA also yields information about the optimal number of factors. In general, ICA may be preferred when there are uncertainties about departures from multivariate normality and to ensure appropriate clinical interpretation of independent factors.

Second, although GenSalt probands have high blood pressure, they did not meet the definition of hypertensive. Thus, some of the conclusions relating to hypertension remain to be tested in samples of hypertensive individuals. However, the fact that none of the individuals in the GenSalt were being treated for hypertension obviates other problems such as those related to the effects of medication use on blood pressure levels.

Third, in general, when using nuclear family data (i.e. parents and offspring only), the heritability estimate includes both shared genes and shared environmental (lifestyle) factors provided the latter exist. Thus, the "heritability" computed using a nuclear family design may be an overestimate of the genetic effect to the extent there are shared environments or lifestyle factors that affect BP levels. However, in the current study, our family structures are extended pedigrees consisting of up to three generations (probands and their siblings as well as the spouse and offspring of the proband). Using this extended pedigree design the heritability may be less biased as compared to a more simple design of nuclear families alone [39].

Finally, the specific results regarding the components arising from the ICA analysis may not be easily applied to other populations since our study was based on a very specific design. However, the applicability of this study is that we have found evidence of an underlying biological genetic cause for a blood pressure response to sodium. Further, this represents a "generalized" response to sodium since it includes concurrent responses to high and low sodium levels as well as to a potassium supplement. Thus, it is likely to capture genes that are common to each of these mechanisms. The identification of the genetic variants leading to this response will be investigated in follow-up studies. As reviewed in the introduction, blood pressure levels are affected by sodium intake in many populations. Whether the genetic variants identified in this Chinese sample are applicable to the general population will be investigated once the variants are identified.

Conclusion

In summary, the present study used ICA as a robust method to extract several novel components which may enhance our ability to find the underlying genes for sodium-dependent hypertension. Certainly, as compared to the individual variables that contributed to the factors, heritabilities are higher for components indexing systolic hypertension, generalized hypertension, and BP responses to the sodium intervention. These results validate the use of these factors as novel endophenotypes in future genetic association studies.

Availability of supporting data

The data are available from the authors upon request to and approval from the GenSalt steering committee.

Additional file

Additional file 1: Additional material. 1. Optimal Number of Components. 2. Age adjustments and distributional properties. 3. Inter-correlations among variables. 4. Non-independence among family members and ICA. 5. Table S1. Significant age covariates and percents of variance. 6. Table S2. Variable Characteristics. 7. Table S3. Correlation between individual variables. 8. Table S4. Kurtosis and Skewness of residuals. 9. Table S5. Correlation between ICA and family ID.

Author details
[1]Washington University in St. Louis School of Medicine, St. Louis, MO, USA. [2]Tulane University Health Sciences Center, New Orleans, LA, USA. [3]Chinese Academy of Medical Sciences, Beijing, China. [4]University of Texas Health Sciences Center at Houston, Houston, TX, USA. [5]National Heart, Lung, and Blood Institute, Bethesda, MD, USA.

References

1. Lewington S, Clarke R, Qizilbash N, Peto R, Collins R. Prospective Studies Collaboration. Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies. Lancet. 2002;9349:1903–13.

Chen *et al. BMC Genetics* (2015) 16:64

Page 10 of 10

2. Singer DR, Kite A. Management of hypertension in peripheral arterial disease: does the choice of drugs matter? Eur J Vasc Endovasc Surg. 2008;35:701–8.

3. Barlassina C, Lanzani C, Manunta P, Bianchi G. Genetics of essential hypertension: from families to genes. J Am Soc Nephrol. 2002;13:S155–64.

4. Svetkey LP, McKeown SP, Wilson AF. Heritability of salt sensitivity in Black Americans. Hypertension. 1996;287:854–8.

5. Gu D, Rice T, Wang S, Yang W, Gu C, Chen CS, et al. Heritability of Blood Pressure Responses to Dietary Sodium and Potassium Intake in Chinese Population. Hypertension. 2007;50:116–22.

6. Elliott P, Stamler J, Nichols R, Dyer AR, Stamler R, Kesteloot H, et al. Intersalt revisited: further analyses of 24 hour sodium excretion and blood pressure within and across populations. BMJ. 1996;312:1249–53.

7. He J, Tell GS, Tang YC, Mo PS, He GQ. Relation of electrolytes to blood pressure in men. Hypertension. 1991;17:378–85.

8. Chobanian AV, Hill M. National Heart, Lung, and Blood Institute Workshop on Sodium and Blood Pressure A critical review of current scientific evidence. Hypertension. 2000;35:858–63.

9. Cutler JA, Follmann D, Allender PS. Randomized trials of sodium reduction: an overview. Am J Clin Nutr. 1997;65(suppl):643S–51.

10. Midgley JP, Matthew AG, Greenwood CM, Logan AG. Effect of reduced dietary sodium on blood pressure: a meta-analysis of randomized controlled trials. JAMA. 1996;275:1590–7.

11. Whelton PK, He J, Cutler JA, Brancati FL, Appel LJ, Follmann D, et al. Effects of oral potassium on blood pressure. Meta-analysis of randomized controlled clinical trials. JAMA. 1997;277:1624–32.

12. Weinberger MH. Salt sensitivity of blood pressure in humans. Hypertension. 1996;27:481–90.

13. Obarzanek K, Proschan MA, Vollmer WM, Moore TJ, Sacks FM, Appel LJ, et al. Individual blood pressure responses to changes in salt intake: results from the DASH-sodium trial. Hypertension. 2003;42:459–67.

14. Vollmer WM, Sacks FM, Ard J, Appel LJ, Bray GA, Simons-Morton DG, et al. Effects of diet and sodium intake on blood pressure: subgroup analysis of the DASH-Sodium trial. Ann Intern Med. 2001;135:1019–28.

15. Wright Jr JT, Rahman M, Scarpa A, Fatholahi M, Griffin V, Jean-Baptiste R, et al. Determinants of salt sensitivity in black and white normotensive and hypertensive women. Hypertension. 2003;42:1087–92.

16. Beeks E, Kessels AG, Kroon AA, van der Klauw MM, de Leeuw PW. Genetic predisposition to salt-sensitivity: a systematic review. J Hypertens. 2004;22:1243–9.

17. Gu CC, Flores HR, de las Fuentes L, Dávila-Román VG. Enhanced detection of genetic association of hypertensive heart disease by analysis of latent phenotypes. Genet Epidemiol. 2008;32:528–38.

18. Almasy L, Blangero J. Endophenotypes as quantitative risk factors for psychiatric disease: rationale and study design. Am J Med Genet. 2001;105:42–4.

19. Gottesman II, Gould TD. The endophenotype concept in psychiatry: etymology and strategic intentions. Am J Psychiatry. 2003;160:636–45.

20. Abdi H, Williams LJ. Principal component analysis. Wiley Interdiscip Rev Comput Stat. 2010;2:433–59.

21. Hyvarinen A, Oja E. Independent component analysis: algorithms and applications. Neural Netw. 2000;13:411–30.

22. Kim D, Kim S-K. Comparing patterns of component loadings: Principal Component Analysis (PCA) versus Independent Component Analysis (ICA) in analyzing multivariate non-normal data. Behav Res. 2012;44:1239–43.

23. Bugli C, Lambert P. Comparison between principal component analysis and independent component analysis in electroencephalograms modelling. Biom J. 2007;49:312–27.

24. Gensalt Collaborative Research Group. Genetic epidemiology network of salt sensitivity (GenSalt): Rationale, design, methods, and baseline characteristics of study participants. J Hum Hypertens. 2007;21:639–46.

25. Perloff D, Grim C, Flack J, Frohlich ED, Hill M, McDonald M, et al. Human blood pressure determination by sphygmomanometer. Circulation. 1993;88:2460–70.

26. The output/code/data analysis for this paper was generated using SAS software, Version 9.1 of the SAS System for UNIX. Copyright © 2002–2003 SAS Institute Inc. SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc., Cary, NC, USA.

27. Peres-Neto PR, Jackson DA, Somers KM. How many principal components? Stopping rules for determining the number of non-trivial axes revisited. Comput Statist Data Anal. 2005;49:974–97.

28. Abecasis GR, Cardon LR, Cookson WO. A General Test of Association for Quantitative Traits in Nuclear Families. Am J Hum Genet. 2000;66:279–92.

29. Smulyan H, Safar ME. The diastolic blood pressure in systolic hypertension. Ann Intern Med. 2000;132:233–7.

30. Kaplan NM. New Issues in the Treatment of Isolated Systolic Hypertension. Circulation. 2000;102:1079–81.

31. Delles C, McBride MW, Graham D, Padmanabhan S, Dominiczak AF. Genetics of hypertension: from experimental animals to humans. Biochem Biophys Acta. 1802;2010:1299–308.

32. Herrera BM, Lindgren CM. The genetics of obesity. Curr Diab Rep. 2010;10:498–505.

33. Bouchard C, Pérusse L, Leblanc C, Tremblay A, Thériault G. Inheritance of the amount and distribution of human body fat. Int J Obes. 1988;12:205–15.

34. Rotimi CN, Cooper RS, Cao G, Ogunbiyi O, Ladipo M, Owoaje E, et al. Maximum-likelihood generalized heritability estimate for blood pressure in Nigerian families. Hypertension. 1999;33:874–8.

35. Saavedra JM. Studies on genes and hypertension: a daunting task. J Hypertens. 2005;23:929–32.

36. Mitchell GF, DeStefano AL, Larson MG, Benjamin EJ, Chen MH, Vasan RS, et al. Heritability and a genome-wide linkage scan for arterial stiffness, wave re-flection, and mean arterial pressure: the Framingham Heart Study. Circulation. 2005;112:194–9.

37. Puppala S, Coletta DK, Schneider J, Hu SL, Farook VS, Dyer TD, et al. Genome-wide linkage screen for systolic blood pressure in the Veterans Administration Genetic Epidemiology Study (VAGES) of Mexican-Americans and confirmation of a major susceptibility locus on chromosome 6q14.1. Hum Hered. 2011;71:1–10.

38. Luo BF, Du L, Li JX, Pan BY, Xu JM, Chen J, et al. Heritability of metabolic syndrome traits among healthy younger adults: a population based study in China. J Med Genet. 2010;47:415–20.

39. Bochud M. Estimating heritability from nuclear family and pedigree data. In Elston, Satagopan and Sun (Eds), Statistical Human Genetics: Methods and Protocols. Humana Press, NY, 2012, pp 171–186.