Personality and Individual Differences xxx (xxxx) xxxx

Contents lists available at ScienceDirect

Personality and Individual Differences

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

Big Five personality and CTRA gene expression: Lack of association in a midlife sample of US adults (MIDUS-Refresher)

Kelsey A. Hobbs[a,*], Frank D. Mann[a], Steven W. Cole[b], Robert F. Krueger[a]

[a] University of Minnesota, Twin Cities, Department of Psychology, N631 Elliott Hall, 75 E River Rd, Minneapolis, MN 55455, United states
[b] University of California, Los Angeles, Department of Psychiatry & Biobehavioral Sciences and Medicine, 11-934 Factor Building, Los Angeles, CA 90095, United States

ARTICLE INFO

Keywords:
Big Five
Gene expression
Inflammation
Antiviral response
CTRA

ABSTRACT

A previous study found that Extraversion and Conscientiousness were associated with increased and decreased expression of a pre-specified set of pro-inflammatory indicator genes. The present study aimed to replicate these findings in a sample of adults from the Refresher Cohort of the Study of Midlife in the United States (MIDUS). Analysis of gene expression composite scores and generalized linear models that took into account the heterogeneity and non-independence of RNA expression across different genes found no significant associations between the pro-inflammatory indicator gene set and the Big Five domains of personality. In addition, there was no significant association between a pre-specified antiviral indicator gene set and the Big Five domains. These findings suggest that relations between Big Five personality and expression of these two immune response indicator gene sets do not consistently appear across samples and may be context-dependent in ways that remain to be elucidated.

Human personality is often framed in terms of the Big Five or Five-Factor Model, which describes normal-range variation in terms of five broad domains: Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism (John, Naumann, & Soto, 2008). Previous studies have found evidence that individual differences in the Big Five domains are associated with health-related outcomes, including longevity and risk for disease (Chapman, Roberts, & Duberstein, 2011). For example, some studies have found that high levels of conscientiousness are associated with increased longevity (Kern & Friedman, 2008; Roberts, Kuncel, Shiner, Caspi, & Goldberg, 2007). Other studies have found that high levels of neuroticism are related to deleterious outcomes, including physical conditions like ulcers and coronary heart disease (Charles, Gatz, Kato, & Pedersen, 2008) and decreased longevity (Roberts et al., 2007). However, after accounting for demographic factors and the other Big Five domains, conscientiousness may be the only robust correlate of mortality (Bogg & Roberts, 2004; Chapman et al., 2011; Jokela et al., 2013a). In comparison, associations between health-related outcomes and the other Big Five domains, specifically openness, agreeableness, and extraversion, tend to vary across studies with confidence intervals that often include zero (Kern & Friedman, 2011).

Although these associations are well-documented, the biological mechanisms that explain why individual differences in personality are associated with physical health remain difficult to discern (Smith, Williams, & Segerstrom, 2015). A study by Vedhara et al. (2015) found that extraversion and conscientiousness were associated with increased and decreased expression of a pre-specified set of 19 pro-inflammatory indicator genes. Interpreted within the framework of behavioral immune response theory and the conserved transcriptional response to adversity (CTRA: Cole, 2013; Powell et al., 2013), these associations highlight one potential mechanism that may explain why individual differences in personality are linked to physical health.

The CTRA captures a pattern of immune-related gene expression that involves the up-regulation of genes involved in inflammation and the down-regulation of genes involved in type-I interferon response and antibody production (Cole, 2013). This immune response may have evolved to help the body respond to threatening conditions and associated risk of wounding injuries (i.e., to combat bacterial infection and promote healing). However, discrepancies between the contemporary environment and the environments in which our immune systems evolved are hypothesized to result in over-activation of pro-inflammatory genes in response to stressors that are not physically injurious but result in subjective distress and psychological impairment. Examples of these stressors include low socioeconomic status (SES) and social isolation characterized by feelings of loneliness, which have both been previously related to inflammation and the CTRA (Cole et al., 2015; Turcotte & Verne, 2016). Thus, in modern contexts the expression of pro-inflammatory genes may be up-regulated in response to

* Corresponding author.
E-mail address: hobbs145@umn.edu (K.A. Hobbs).

https://doi.org/10.1016/j.paid.2020.109908
Received 28 November 2019; Received in revised form 8 February 2020; Accepted 8 February 2020
0191-8869/ © 2020 Elsevier Ltd. All rights reserved.

ARTICLE IN PRESS

economic, social, and relational factors that are not life threatening but induce stress, providing a potential mechanism by which these social factors become biologically embedded to impact physical and mental health.

In short, if individual differences in personality are related to the differential expression of immune-related genes, then this may help explain why individual differences in personality are related to physical health outcomes. The associations reported by Vedhara et al. (2015) provide a starting point for such inquiry, but these estimated associations were based on a community sample of young adults from the United Kingdom ($N = 121$) and have yet to be replicated. Additionally, Brown and Nickerson (2016) reanalyzed the data reported by Vedhara et al. (2015) and concluded that associations between pro-inflammatory gene expression and the Big Five domains may be fragile. Thus, a replication attempt would serve to provide a clearer picture of how personality relates to the CTRA. In the present study we examined the same associations as Vedhara et al. (2015) in a community sample of middle-aged adults ($N = 539$) participating in the Study of Midlife in the United States (MIDUS). We sought to determine whether the pro-inflammatory gene set used by Vedhara et al. (2015) is associated with any of the Big Five domains of personality. We also examined in parallel the antiviral immune response gene set that Vedhara et al. (2015) found was not significantly associated with any of the Big Five domains.

1. Methods

1.1. Sample

Data for the current study came from a sample of adults ($n = 539$) who participated in the biomarker project of the Study of Midlife in the United States (MIDUS; Ryff & Krueger, 2018), consented to provide genetic data, passed RNA quality control metrics, and completed Big Five adjectival scales (i.e., the MIDI). The average age of participants was 52 years (median = 52, min. = 25, max. = 76, SD = 13.29). Approximately 50% of the sample was female ($n = 271$ females; $n = 268$ males). Approximately 73% of the sample identified as White/European American ($n = 392$), ~17% Black/African American ($n = 91$), ~2% Asian American ($n = 8$), ~2% Native American ($n = 11$), and ~6% "Other" race/ethnicity ($n = 34$). Detailed information regarding participant recruitment and data collection can be found elsewhere (Brim, Ryff, & Kessler, 2004). Calculated using G*Power 3.1 (Erdfelder, Faul, & Buchner, 1996), an *a priori* power analysis for a multiple linear regression with 25 covariates indicates that sample sizes of $N = 507$, $N = 250$, and $N = 82$ participants are needed to be adequately powered (80%) to detect the incremental effects of five independent variables, assuming small ($\Delta R^2 = .025$), small-to-moderate ($\Delta R^2 = .05$) and moderate effect sizes ($\Delta R^2 = .15$), respectively. Thus, the current study was adequately powered to detect expected associations.

2. Measures

2.1. Gene expression

Blood samples were provided by participants in the MIDUS biomarker project during a laboratory visit and shipped to a central biospecimen repository at the University of Wisconsin. Peripheral blood mononuclear cells (PBMCs) were then isolated and frozen at −70 C for several years prior. Frozen PBMC samples were subsequently thawed and RNA was extracted using standard protocols (Qiagen RNeasy). Extracted RNA was quality control checked for yield and integrity, and subject to transcriptome profiling using an mRNA-targeted sequencing approach (using Lexogen QuantSeq 3′ FWD cDNA library synthesis with sequencing on an Illumina HiSeq 4000 instrument) targeting >10 million 65-bp single stranded sequencing reads per sample. RNA sequencing was conducted in batches comprised of a 96-well plate, which was indicated by a nominal variable and included as a covariate in inferential analyses. Two sets of genes were selected *a priori* based on their previous use in CTRA research: 19 pro-inflammatory genes (*IL1A, IL1B, IL6, IL8, TNF, PTGS1, PTGS2, FOS, FOSB, FOSL1, FOSL2, JUN, JUNB, JUND, NFKB1, NFKB2, REL, RELA,* & *RELB*) and 32 genes involved in Type I interferon responses and antibody production (*GBP1, IFI16, IFI27, IFI27L1- 2, IFI30, IFI35, IFI44, IFI44L, IFI6, IFIH1, IFIT1-3, IFIT5, IFIT1L, IFITM1-3, IFITM4P, IFITM5, IFNB1, IRF2, IRF7-8, MX1- 2, OAS1-3, OASL, IGJ, IGLL1,* & *IGLL3*).

2.2. Midlife development inventory (MIDI)

The MIDI (Lachman & Weaver, 1997) is a 26-item questionnaire that measures the Big Five domains using between four and seven adjectives for each domain drawn from existing inventories (Goldberg, 1992; John, 1990; Trapnell & Wiggins, 1990). Adjective groupings are as follows: organized, responsible, hardworking, careless, thorough (for Conscientiousness); helpful, warm, caring, softhearted, sympathetic (for Agreeableness); moody, worrying, nervous, calm (for Neuroticism); creative, imaginative, intelligent, curious, broad-minded, sophisticated, adventurous (for Openness); outgoing, friendly, lively, active, talkative (for Extraversion). Participants were asked to rate how well each adjective described them on a Likert scale from 1 (a lot) to 4 (not at all). Responses were then reverse coded (and specific adjective scores reversed as necessary) so that a higher number indicated higher levels of the personality domain. Calculated using the "psych" package (Revelle, 2019), Omega index for the MIDI scales evinced moderate-to-high internal consistencies (range of $\omega = .71$ to .80). Providing evidence for concurrent validity, a previous study found that four of the five domains of the MIDI (i.e. openness, conscientiousness, extraversion, and neuroticism) are highly correlated with the NEO scales with r ranging from .63 to .81 (Lachman, 2005). In this study, the agreeableness scale from the MIDI was moderately correlated (.42) with its NEO counterpart, due to the MIDI covering only two NEO facets of agreeableness (i.e. altruism and trust).

2.3. Covariates

A number of demographic variables were included as covariates of pro-inflammatory and antiviral indictor gene expression: self-reported age (reported in years and mean-centered), biological sex (0 = Male, 1 = Female), Black/African-American (0 = No, 1 = Yes), Native American (0 = No, 1 = Yes), Asian American (0 = No, 1 = Yes), and "Other" race/ethnicity (0 = No, 1 = Yes). Variables that capture individual differences in physical and behavioral health were also included as covariates: body mass index (BMI) was calculated based on participants` height and weight (BMI = weight in kilograms divided by height in meters squared); History of alcohol consumption was measured on an ordinal scale by asking participants "During the past month, how often did you drink any alcoholic beverages, on the average? Would you say every day (5), 5 or 6 days a week (4), 3 or 4 days a week (3), 1 or 2 days a week (2), less often than 1 day a week (1), or never (0)?"; Smoking status was measured by asking participants whether they smoked cigarettes regularly (0 = No, 1 = Yes); Number of Physical Conditions was measured as a continuous variable by taking the sum of 39 self-reported chronic and acute conditions that were experienced by participants in the last twelve months (e.g., asthma, tuberculosis, sciatica, thyroid disease, hay fever, urinary/bladder problems, AIDS/HIV, lupus, hypertension, autoimmune disorders, anxiety and depression, diabetes, neurological disorders, ulcer, hernia, hemorrhoids, etc.).

In addition to a nominal variable representing plate batch, continuous variables that captured the prevalence of transcripts marking T lymphocytes subsets (*CD3D, CD3E, CD4, CD8A*), B lymphocytes (*CD19*), NK cells (CD16/*FCGR3A*, CD56/*NCAM1*), and monocytes (*CD14*) were also included as covariates of pro-inflammatory and antiviral gene expression. Inclusion of physical condition measures and leukocyte subset

prevalence may constitute over-control (because the latter constitutes the outcome that the gene expression dynamics are hypothesized to mediate and may thus share variance with it, whereas the former represents one mechanism underlying CTRA expression, i.e., changes in monocyte prevalence/activity). We included these variables in primary analyses to follow the analytic strategy reported by Vedhara et al. (2015), but we also conducted sensitivity analyses removing those potentially over-controlling covariates.

2.4. Data analytic procedures

Gene expression values were log2 transformed and mean-centered within-gene (i.e., log2 expression values for each gene were subtracted from the mean expression value for that gene) before averaging across pro-inflammatory and antiviral indicator gene sets to create continuous gene expression composite scores. The difference between the pro-inflammatory and antiviral composite scores were also calculated to create continuous CTRA scores. Descriptive statistics and zero-order Pearson's product-moment correlations between continuous measures of the Big Five domains and gene expression composite scores are reported in Table S1 in the supplemental materials.

The relations between the Big Five domains with pro-inflammatory and antiviral gene expression composite scores were then estimated using multiple linear regressions, controlling for demographic variables (i.e., age, sex, and race/ethnicity), physical and behavioral health (i.e., BMI, history of alcohol consumption, smoking status, and number of physical conditions) and technical covariates (i.e., plate batch and the prevalence of transcripts marking T lymphocytes subsets in PBMC pools). First, multiple linear regressions were estimated that regressed pro-inflammatory and antiviral gene expression composite scores on all study covariates, including demographic variables, physical and behavioral health variables, and technical covariates. These models served as a baseline for latter comparison. Using information criteria (AIC & BIC) and percent of variance explained (R^2) to guide model selection, baseline models that included demographic variables, physical and behavioral health variables, and technical covariates as independent variables were compared to models that included the same set of independent variables plus continuous measures of each of the Big Five transformed to a standardized metric ($M = 0$, SD = 1). In all models, independent variables were specified as simultaneous predictors. Consequently, multiple regression coefficients predicting pro-inflammatory and antiviral gene expression composite scores are adjusted for the presence of the other predictors in the model.

Analysis of composite scores using multiple linear regression provides a simple and straightforward approach to analyzing the differential expression of pro-inflammatory and antiviral gene sets. However, because RNA expression can be heterogenous and correlated across different genes, analyzing composite scores that fail to account for these patterned nuances can lead to type-II errors. Of course, this would be a critical limitation to any replication attempt. Therefore, to circumvent this potential problem, additional analyses were conducted that treated each indicator gene as a repeated RNA value for each participant and tested the average multiple regression weight across all indicator genes for each of the Big Five domains, while also estimating a variance-covariance matrix of residuals. The freely estimated diagonal elements of the residual matrix accounts for the potential heterogeneity of RNA expression across different genes, and the non-redundant off-diagonal elements account for the non-independence of RNA expression across different genes. In these models, all associations were adjusted for potential confounding by the same set of covariates as the multiple linear regression analysis of gene expression composite scores: age, sex, race/ethnicity, BMI, alcohol consumption, smoking, number of physical conditions, and 8 RNA transcripts indicting the relative prevalence of T lymphocytes, B lymphocytes, NK cells, and monocytes (again omitting physical conditions and leukocyte subset measures in subsequent sensitivity analyses).

Prior to estimating repeated-measure mixed models, quantile-normalized gene expression values were log2-transformed and standardized within gene and then sign-inverted for antiviral indicator genes. Because RNA sequencing can yield an excess of zeros, the distributions of gene transcripts were screened, and descriptive statistics were calculated to exclude data from any gene with an average expression level $< .5$ log2 transcripts per million mapped reads (TPM- i.e., the native value of the normalized expression data). This screening processes resulted in the deletion of data from 9 genes (*CXCL8, IL1A, FOSL1, IL6, IFI27, IFITM4P, IFITM5, IFNB1, IGLL1*). RNA expression values from the remaining pro-inflammatory and antiviral indicator genes were specified as a repeated-measure and analyzed using the SAS PROC MIXED command with maximum likelihood estimation.[1] In addition to predicting RNA expression in pro-inflammatory and antiviral (sign-inverted) indicator genes, we also conducted analyses using the same models predicting RNA expression for pro-inflammatory and antiviral indicator gene sets independently of each other.

Ancillary Pearson's product-moment correlations between the Big Five domains and body mass index (BMI) were also calculated as a straightforward check on the convergent validity of the Big Five measures used herein, as these measures differed from those used by Vedhara et al. (2015). Past studies have found that BMI tends to correlate positively with neuroticism and negatively with conscientiousness (Brummett et al., 2006; Sutin & Terracciano, 2016), perhaps the more robust association being with conscientiousness (Chapman, Fiscella, Duberstein, & Kawachi, 2009; Jokela et al., 2013b) and especially in older adults (Mõttus et al., 2013). Correlations between BMI and the other Big Five domains, on the other hand, tend to vary across studies with confidence intervals that often include zero.

3. Results

Characteristics of the analyzed sample ($n = 539$), including descriptive statistics for focal study variables, are reported in the supplemental materials (see Table S1). Table 1 reports the results of the multiple linear regressions relating pro-inflammatory and antiviral gene expression composite scores to the Big Five domains. Using information theoretic criteria to guide model selection, models that did not include the Big Five as predictors of pro-inflammatory gene expression ($AIC = -132.794$, $BIC = -18.097$) were preferred over models that included the Big Five as predictors ($AIC = -126.501$, $BIC = 9.437$). Moreover, after accounting for the effects of study covariates, the incremental percent of variance in pro-inflammatory gene expression collectively explained by the Big Five domains approached zero ($\Delta R^2 = .001$, $\Delta R^2_{adjusted} = .000$, $F(5, 491) = 0.699$, $p = .624$). The association between pro-inflammatory gene expression and extraversion was in the predicted direction ($b = .002$, $SE = .011$, $p = .888$) but not significantly different from zero. The association between conscientiousness and pro-inflammatory gene expression was in the opposite direction as predicted and not significantly different than zero

[1] At the request of a reviewer, additional sensitivity analyses were performed using the SAS "EMPIRICAL" option, whereby the same models were fit to the data, except the precision and statistical significance of fixed effects were evaluated using robust standard errors computed using a Huber-White sandwich estimator. These models demonstrated that none of the Big Five domains were significantly associated with pro-inflammatory expression, before and after excluding potentially over-controlling covariates. However, Openness was negatively associated with antiviral gene expression ($p = .022$) when controlling for all covariates. After omitting potentially over-controlling covariates, the association was no longer significant ($p = .489$). Additionally, neuroticism was negatively associated with CTRA gene expression (both pro-inflammatory and antiviral components; $p = .043$), when including all covariates. This relationship was in the opposite direction as expected. Again, after removing potentially over-controlling covariates this relationship was no longer significant ($p = .184$).

ARTICLE IN PRESS

Table 1
Results of Multiple Linear Regressions Testing Whether the Big Five Domains are Associated with the Expression of Pre-Specified Sets of Pro-Inflammatory and Antiviral Indicator Genes.

n = 539 DVs (row)	Pro-Inflammatory Indicator Gene Expression Composite Score						Antiviral Indicator Gene Expression Composite Score						
IVs (column)	*b*	*SE*	*p*	*b*	*SE*	*p*	*b*	*SE*	*p*	*b*	*SE*	*p*	*VIF*
(Intercept)	.016	.029	.596	.016	.030	.584	−.272	.055	<.001	−.267	.056	<.001	–
CD3E	−.003	.021	.881	−.005	.021	.828	−.056	.040	.160	−.053	.040	.184	2.32
CD3D	.017	.016	.288	.017	.016	.280	.088	.030	.003	.085	.030	.004	4.23
CD4	.060	.019	.001	.060	.019	.002	.101	.035	.003	.105	.035	.003	4.32
CD8A	.023	.011	.030	.025	.011	.025	.050	.020	.015	.051	.021	.013	2.25
CD14	.152	.013	<.001	.151	.014	<.001	.111	.025	<.001	.113	.025	<.001	4.21
CD19	.004	.009	.665	.004	.009	.672	−.019	.017	.264	−.019	.017	.258	1.49
FCGR3A	.034	.010	.001	.034	.010	<.001	.100	.019	<.001	.099	.019	<.001	2.80
NCAM1	.015	.009	.112	.014	.009	.130	−.024	.017	.163	−.024	.017	.174	1.44
Batch Plate (2)	−.158	.073	.031	−.151	.073	.040	−.176	.136	.198	−.185	.137	.176	3.99
(3)	−.154	.038	<.001	−.152	.038	<.001	.342	.070	<.001	.345	.070	<.001	–
(4)	−.134	.037	<.001	−.132	.038	<.001	.290	.070	<.001	.276	.070	<.001	–
(5)	.046	.038	.220	.049	.038	.204	.369	.070	<.001	.363	.071	<.001	–
(6)	−.073	.040	.069	−.075	.040	.066	.380	.075	<.001	.377	.075	<.001	–
(7)	−.068	.039	.084	−.069	.039	.078	−.001	.073	.984	.000	.073	.996	–
(8)	.258	.039	<.001	.258	.039	<.001	.252	.073	<.001	.250	.073	<.001	–
Age	.003	.001	.003	.002	.001	.006	.001	.002	.396	.002	.002	.293	1.74
Gender	−.009	.019	.642	−.012	.020	.548	−.006	.036	.879	−.007	.038	.853	1.25
Black	.030	.028	.273	.036	.028	.200	.219	.052	<.001	.221	.052	<.001	1.33
Native	.071	.069	.302	.062	.069	.374	−.065	.128	.613	−.051	.129	.691	1.10
Asian	−.012	.081	.881	−.027	.082	.742	.028	.151	.852	.030	.153	.846	1.08
Other	.010	.039	.795	.011	.039	.787	.093	.073	.204	.082	.073	.267	1.08
BMI	−.005	.010	.635	−.004	.010	.687	−.003	.019	.889	−.002	.019	.916	1.22
Smoking	.031	.032	.340	.029	.033	.373	.051	.060	.402	.040	.061	.511	1.22
Alcohol Use	−.007	.010	.488	−.005	.010	.603	.008	.019	.657	.005	.019	.776	1.28
# of Conditions	.004	.010	.676	.008	.011	.448	−.007	.019	.698	−.015	.020	.436	1.37
Conscientious	–	–	–	.009	.011	.389	–	–	–	.015	.020	.439	1.36
Agreeableness	–	–	–	−.012	.012	.293	–	–	–	−.006	.022	.773	1.62
Neuroticism	–	–	–	−.010	.011	.368	–	–	–	.032	.020	.108	1.35
Openness	–	–	–	−.011	.011	.327	–	–	–	.023	.021	.263	1.45
Extraversion	–	–	–	.002	.011	.888	–	–	–	−.025	.021	.248	1.58
	Effect Sizes												
	R^2		$R^2_{adjusted}$		R^2	$R^2_{adjusted}$		R^2		$R^2_{adjusted}$	R^2	$R^2_{adjusted}$	
	.808		.798		.809	.797		.618		.599	.622	.599	

Notes. n = sample size. DV = dependent variable. IV = independent variable. b = multiple linear regression coefficient. *SE* = standard error. p = probability of the data if the null hypothesis is true (i.e., $b = 0$). *VIF* = variance inflation factor for multiple regression including the Big Five. R^2 = percent of variation in gene expression scores explained by multiple regressions. $R^2_{adjusted} = [1 - (1 - R^2)] \times [(n - k)/(n - k - 1)]$, where k = number of independent variables.

($b = 0.009$, $SE = .011$, $p = .389$).

Similarly, a multiple linear regression model that did not include the Big Five as predictors of antiviral gene expression (*AIC* = 512.328, *BIC* = 627.025) was preferred over a model that included the Big Five as predictors of antiviral gene expression (*AIC* = 516.662, *BIC* = 652.600), and the percent of incremental variance in antiviral gene expression collectively explained by the Big Five domains approached zero ($\Delta R^2 = .004$, $\Delta R^2_{adjusted} < .001$, $F(5, 491) = 1.07$, $p = .375$). Beta weights predicting change in gene expression composite scores given a standard deviation increase in a Big Five domain varied in direction and magnitude (range of $\beta = -.025$ to .032), but none of the associations between the Big Five domains and gene expression composite scores met conventional standards for "statistical significance" (*ps* > .05), even before adjusting alpha-level for multiple comparisons. Moreover, results remained largely unchanged after excluding potentially over-controlling covariates. Zero-order correlations between the Big Five domains and gene expression composite scores are reported in Table S1 in the supplemental materials. The results of sensitivity analyses are reported in Table S2.

To quantify the degree of multicollinearity, variance inflation factors (*VIFs*) were calculated for models that included that Big Five domains as predictors of gene expression composite scores. The square root of a *VIF* indicates how much multicollinearity has decreased the precision of the estimated effect by increasing its standard errors, that is, compared to what the standard error of the estimated effect would be if the variable were uncorrelated with the other independent variables in the model. As a general rule of thumb, a *VIF* greater than or equal to 10 indicates a high degree of multicollinearity, as the standard error of the associated beta weight would be more than three times as large as it would be ($\sqrt{10} = 3.1623$) if the independent variable were uncorrelated with the other variables in the model (Gelman & Hill, 2006). Reported in Table 1, *VIFs* for multiple linear regressions revealed a relatively low degree of multicollinearity among the predictors of gene expression composite scores (range of *VIF* = 1.08 to 4.32)

In sum, multiple linear regression analysis of the Big Five domains and gene expression composite scores failed to replicate the associations documented by Vedhara et al. (2015). However, as previously noted, analysis of composite scores may be overly conservative (i.e., likely to produce a type-II error), as it does not adequately address potential heterogeneity and non-independence of RNA expression across different genes. Therefore, consistent with previous studies of differential gene expression in relation to psychosocial outcomes (Cole et al., 2015; Fredrickson et al., 2015; Vedhara et al., 2015), we estimated the models described above that treat RNA expression as a repeated measure and account explicitly for heterogeneity and non-independence of gene transcripts by estimating an unstructured residual variance-covariance matrix.

Reported in Table 2, estimates of fixed effects indicate that the results of secondary analyses were largely consistent with the results of primary analyses. A mixed model that did not include the Big Five domains as predictors of pro-inflammatory gene expression (AIC = 16,686.40, BIC = 17,366.10) was preferred over a mixed model that included the Big Five as predictors of pro-inflammatory gene expression (AIC = 16,693.1, BIC = 16,794.00). Moreover, none of the

ARTICLE IN PRESS

Table 2
Estimated Fixed Effects from a Repeated Measures Model of the Big Five Domains, Demographic and Health-Related Variables, and Technical Covariates Predicting a Pre-Specified Set of Pro-Inflammatory Indicator Genes.

Dependent Variable: Pro-Inflammatory Indicator Gene Expression Continuous Predictors:	*b*	*SE*	*t*	*p*
(Intercept)	.226	.050	4.48	<.001
(Lymphocyte Covariates)				
CD3E	.036	.020	1.74	.083
CD3D	.034	.017	1.99	.047
CD4	.085	.020	4.22	<.001
CD8A	.036	.011	3.21	.001
CD14	.197	.013	14.75	<.001
CD19	.015	.009	1.73	.084
FCGR3A	.028	.010	2.79	.005
NCAM1	.029	.009	3.10	.002
(Demographic Variables)				
Age	.002	.001	1.99	.047
Gender	.008	.020	.42	.676
Black/African American	.001	.026	.04	.969
Native American	.078	.069	1.14	.256
Asian American	−.116	.077	1.50	.134
Other Race/Ethnicity	−.009	.038	.24	.808
(Health-Related Variables)				
Body Mass Index (BMI)	.010	.010	1.07	.285
Smoking Status	.053	.032	1.66	.099
History of Alcohol Use	−.009	.010	.92	.359
# of Physical Health Conditions	−.009	.010	.90	.368
(Big Five Domains)				
Conscientious	.005	.010	.51	.607
Agreeableness	−.012	.011	1.11	.266
Neuroticism	−.009	.010	.89	.372
Openness	.009	.011	.86	.391
Extraversion	−.004	.011	.34	.732

Type-III Tests of Fixed Effects Nominal Predictors:	df_1	df_2	*F*	*p*
(Gene & Technical Covariates)				
Gene	14	465	.02	.999
Plate Batch	7	465	21.69	< .001

Notes. *b* = multiple regression weight. SE = standard error. *T* = test statistic. *p* = probability of the observed data assuming the null hypothesis (i.e., *b* = 0) is true. The F-statistics and p-values reported under Type-III Tests of Fixed Effects provide an omnibus test of whether mean levels of gene expression are equal across all groups of nominal covariates.

associations between the Big Five domains and pro-inflammatory gene expression were significantly different than zero (i.e., *p*s > 0.25). The fixed effect of extraversion on pro-inflammatory gene expression was in the opposite direction as predicted (*b* = −.004, SE = .011, *p* = .732), as was the fixed effect of conscientiousness on pro-inflammatory gene expression (*b* = .005, SE = .010, *p* = .607). Further, the size and precision of estimated effects were similar after omitting potentially over-controlling covariates (see Table S5).

Reported in the supplemental materials (Tables S3), a similar pattern of results was observed when RNA values for antiviral indicator genes were specified as the dependent variable. A mixed model that only included covariates of antiviral gene expression (AIC = 30,359.10, BIC = 32,185.80) was preferred over a mixed model that included covariates plus the Big Five as predictors of antiviral gene expression (AIC = 30,359.30, BIC = 32,207.20). Further, none of the observed associations between the Big Five domains were significantly different than zero (*p*s > .10), with one exception. Although not predicted, the fixed effect of openness on antiviral gene expression (sign-inverted) was significantly different than zero (*b* = −.032, SE = .011, *p* = .006), but did not meet a Bonferroni-corrected threshold for statistical significance. Plus, a sensitivity analysis indicated that this effect was not significantly different than zero after omitting potentially over-controlling covariates (*b* = −.012, SE - .012, *p* = .506), and the zero-order correlation between openness and antiviral gene expression composite scores was not significantly different than zero (*r* = .05, CI.95% = −.03, .13, *p* = .245).

Finally, mixed models that simultaneously predicted pro-inflammatory and antiviral (sign-inverted) indicator gene expression provided little evidence for associations with the Big Five domains. None of the observed associations between the Big Five domains and CTRA gene expression met conventional standards for statistical significance (*p*s > .05), even before adjusting p-value thresholds for multiple comparisons, and the size and precision of estimated effects remained largely unchanged after omitting potentially over-controlling covariates. Although the present study did not replicate associations between extraversion and conscientiousness with pro-inflammatory gene expression, the current study replicated the Big Five correlates of BMI that have been observed in past studies: conscientiousness negatively correlated with BMI (*r* = −.141, 95% C.I. = −.223, −.056, *p* = .001), neuroticism positively correlated with BMI (*r* = .115, 95% C.I. = .030, .199, *p* = .008), and zero-order correlations between BMI and the remaining Big Five domains were small in magnitude (correlations < .07) and not significantly different than zero (*p*s > .10).[2]

4. Discussion

The present study attempted to replicate previously observed associations between the Big Five domains of extraversion and conscientiousness with pro-inflammatory gene expression in a community sample of middle-aged adults. Consistent with Vedhara et al. (2015), associations between the expression of antiviral indicator genes and the Big Five were small in magnitude and not significantly different from zero. However, the present study found that associations between the expression of pro-inflammatory indicator genes and the Big Five were also not significantly different than zero. Moreover, the association between the expression of pro-inflammatory indicator genes and conscientiousness was in the opposite direction as predicted. These null findings were consistent across different analytic strategies, as well as before and after accounting for several covariates, including demographic variables, health-related outcomes, and other biological and technical variables.

Although we did not find any significant associations involving extraversion or conscientiousness as reported by Vedhara et al. (2015), there was mixed evidence that openness was associated with antiviral gene expression. Multiple linear regressions revealed a null association between openness and the antiviral gene expression composite score, both before and after omitting potentially over-controlling covariates. The zero-order correlation between openness and the antiviral gene expression composite score was also not significantly different than zero. The results of the mixed model that accounted for heterogeneity and non-independence of gene transcripts, however, indicated that openness was significantly associated with antiviral gene expression. However, this fixed effect was not robust in sensitivity analyses that omitted potentially over-controlling covariates. In sum, only one of five relevant tests resulted in rejecting the null hypothesis of no association between openness and antiviral gene expression. Neuroticism was also marginally associated with CTRA gene expression composite scores, but in the opposite direction as expected. Neuroticism is a personality risk factor for anxiety and depression and so might be expected to *positively* relate to the CTRA, but we observed a trend toward negative association.

[2] $r_{Agreeablness}$ = .014, 95% C.I. = −.071, .099, *p* = .750; $r_{Openness}$ = −.069, 95% C.I. = −.154, .016, *p* = .111; $r_{Extraversion}$ = −.029, 95% C.I. = −.115, .056, *p* = .492).

Given that primary and secondary analyses failed to replicate associations between Big Five personality and CTRA gene expression, one may question the reliability of the MIDI measures of the Big Five implemented in the current study, especially given that internal consistencies of adjectival scales did not meet traditional psychometric standards for "high" reliability (e.g., $\alpha > .90$). However, when a brief measure is used to index a broad content space, such as measuring a broad domain of personality with a short list of adjectives, modest internal consistency is not only expected but desired (Boyle, 1991; Kline, 1979; Little, Lindenberger, & Nesselroade, 1999), as high internal consistency in such cases may indicate that a measure is too narrowly focused and, consequently, "antithetical to high validity" (Kline, 1986, p.118). Previous studies have also found that the MIDI scales used in the current study correlate with mortality and inflammatory markers in the expected directions (Chapman et al., 2009; Elliot, Turiano, & Chapman, 2017). The MIDI scales have additionally been linked to a myriad of other health related behaviors such as smoking, missing work for health-related reasons, and alcohol consumption (Chapman et al., 2009; Graham, Mroczek, & Elleman, 2015; Hakulinen et al., 2015). Ancillary correlational analyses were also conducted to provide a test of predictive validity for the MIDI measure of the Big Five. Importantly, zero-order correlations between the Big Five domains and BMI were largely consistent with past studies. Despite this evidence supporting the validity of the Big Five scales used in the current study, as Vedhara et al. (2015) measured the Big Five domains using the NEO personality inventory, differences in personality measurement may nevertheless contribute to the lack of replication in the present study.

There are a number of additional limitations worth nothing. Most notably, blood samples were collected in different facilities from where they were processed into PBMC and subsequently stored, and transcriptome alterations may have occurred during shipment and processing (e.g., due to hemolysis and/or hypoxia). This may have resulted in both general and transcript-specific alterations relative to a "fresh capture" method. Other differences between this study and the one reported by Vedhara et al. (2015) include the age cohort, cultural/national setting, racial and ethnic composition of the sample, sample recruitment methods, and the specific gene expression measurement (e.g., microarray assays of stabilized whole blood in Vedhara et al., 2015 vs. RNA sequencing of shipped and stored archival PBMC in this study, which resulted in differences in the general abundance of some gene transcripts between studies as well as differences in the specific set of gene transcripts analyzed). The Vedhara et al. (2015) study excluded individuals with any acute or chronic illness, whereas the present study included ill individuals but sought to control for such effects by including them as a covariate. Finally, it is important to note that the results of this highly focused replication study do not rule out the possibility that personality dimensions might correlate with other aspects of gene expression that remain to be examined in future research (and potentially even with other sets of inflammatory or antiviral genes besides the specific gene sets examined here).

In sum, although the Big Five domains were not reliably associated with either the CTRA profile overall or with its pro-inflammatory or antiviral subcomponents in the current study, the Big Five domains were associated with BMI in the expected directions. This suggests that the failure to replicate associations between the Big Five and CTRA gene expression cannot be explained simply by poor personality measurement quality. As the present study was adequately powered to detect the associations reported by Vedhara et al. (2015), failure to replicate links between personality and gene expression in this sample suggests that previously documented associations are unreliable in the sense of being either context-dependent (e.g., limited to specific population characteristics in ways that remain to be elucidated) or type-I errors. At the very least, assuming validity of the current personality and gene expression measures, and the absence of any material bias or confounding within this analysis, these results imply that the previously reported links of extraversion and conscientiousness to pro-inflammatory gene expression are not globally reliable across all study contexts.

Author contributions

KAH and RFK conceptualized the replication attempt. FDM performed formal analysis. KAH and FDM wrote the original draft of the manuscript. RFK supervised KAH and FDM. RFK and SWC acquired funding for the research and SWC conducted gene expression assays. All authors provided feedback and critical revisions before approving a final version of the manuscript.

Acknowledgment

Data were drawn from the biomarker project of MIDUS, which is funded by the National Institute on Aging (P01-AG020166 and U19-AG051426).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.paid.2020.109908.

References

Bogg, T., & Roberts, B. W. (2004). Conscientiousness and health-related behaviors: A meta-analysis of the leading behavioral contributors to mortality. *Psychological Bulletin, 130*(6), 887–919.

Boyle, G. J. (1991). Does item homogeneity indicate internal consistency or item redundancy in psychometric scales? *Personality and Individual Differences, 12*(3), 291–294.

Brim, O. G., Ryff, C. D., & Kessler, R. C. (2004). *The MIDUS national survey: an An overview. How healthy are we.* 1–36.

Brown, N. J. L., & Nickerson, C. A. (2016). Reanalysis of Vedhara et al.'s (2015) study of personality and gene expression. *Personality and Individual Differences, 91*, 133–137.

Brummett, B. H., Babyak, M. A., Williams, R. B., Barefoot, J. C., Costa, P. T., & Siegler, I. C. (2006). NEO personality domains and gender predict levels and trends in body mass index over 14 years during midlife. *Journal of Research in Personality, 40*(3), 222–236.

Chapman, B. P., Roberts, B., & Duberstein, P. (2011). Personality and longevity: Knowns, unknowns, and implications for public health and personalized medicine. *Journal of Aging Research, 2011*, 1–24.

Chapman, B., Fiscella, K., Duberstein, P., & Kawachi, I. (2009). Education and smoking: Confounding or effect modification by phenotypic personality traits. *Annals of Behavioral Medicine, 38*(3), 237–248.

Charles, S. T., Gatz, M., Kato, K., & Pedersen, N. L. (2008). Physical health 25 years later: The predictive ability of neuroticism. *Health Psychology, 27*(3), 369–378.

Cole, S. W. (2013). Social regulation of human gene expression: Mechanisms and implications for public health. *American Journal of Public Health, 103*(S1), S84–S92.

Cole, S. W., Levine, M. E., Arevalo, J. M. G., Ma, J., Weir, D. R., & Crimmins, E. M. (2015). Loneliness, eudaimonia, and the human conserved transcriptional response to adversity. *Psychoneuroendocrinology, 63*, 11–17.

Elliot, A. J., Turiano, N. A., & Chapman, B. P. (2017). Socioeconomic status interacts with conscientiousness and neuroticism to predict circulating concentrations of inflammatory markers. *Annals of Behavioral Medicine, 51*(2), 240–250.

Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior Research Methods, Instruments, & Computers, 28*(1), 1–11.

Fredrickson, B. L., Grewen, K. M., Algoe, S. B., Firestine, A. M., Arevalo, J. M., Ma, J., et al. (2015). Psychological well-being and the human conserved transcriptional response to adversity. *PloS One, 10*(3), e0121839.

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment, 4*, 26–42.

Graham, E. K., Mroczek, D. K., & Elleman, L. G. (2015). Personality & earnings lost: The economic costs of work cut back days due to physical and mental health. *Personality Psychology, 1*(1), 1–7.

Hakulinen, C., Elovainio, M., Batty, G. D., Virtanen, M., Kivimaki, M., & Jokela, M. (2015). Personality and alcohol consumption: Pooled analysis of 72,949 adults from eight cohort studies. *Drug and Alcohol Dependence, 151*, 110–114.

John, O. P. (1990). The "Big Five" factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In L. A. Pervin (Ed.). *Handbook of personality: Theory and research* (pp. 66–100). New York, NY: Guilford Press.

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative big five trait taxonomy. *Handbook of Personality: Theory and Research, 3*, 114–158.

Jokela, M., Batty, G. D., Nyberg, S. T., Virtanen, M., Nabi, H., & Singh-Manoux, A. (2013a). Personality and all-cause mortality: Individual-participant meta-analysis of 3,947 deaths in 76,150 adults. *American Journal of Epidemiology, 178*(5), 667–675.

Jokela, M., Hintsanen, M., Hakulinen, C., Batty, G. D., Nabi, H., & Singh-Manoux, A. (2013b). Association of personality with the development and persistence of obesity:

ARTICLE IN PRESS

A meta-analysis based on individual-participant data. *Obesity Reviews, 14*, 315–323.
Kern, M. L., & Friedman, H. S. (2008). Do conscientious individuals live longer? A quantitative review. *Health Psychology, 27*, 505–512.
Kline, P. (1979). *Psychometrics and psychology.* London: Academic Press.
Kline, P. (1986). *A handbook of test construction: Introduction to psychometric design.* New York: Methuen.
Lachman, M. E., & Weaver, S. L. (1997). *The midlife development inventory (MIDI) personality scales: scale Scale construction and scoring.* Waltham, MA: Brandeis University, Department of Psychology. (Tech. Rep. No. 1) https://www.brandeis.edu/psychology/lachman/pdfs/midi-personality-scales.pdf.
Lachman, M. E. (2005). *Addendum for MIDI personality scales.* MIDUS II version. http://www.brandeis.edu/departments/psych/ lachman/pdfs/revised-midi-scales.pdf.
Little, T. D., Lindenberger, U., & Nesselroade, J. R. (1999). On selecting indicators for multivariate measurement and modeling with latent variables: When" good" indicators are bad and" bad" indicators are good. *Psychological Methods, 4*(2), 192–211.
Mõttus, R., McNeill, G., Jia, X., Craig, L. C., Starr, J. M., & Deary, I. J. (2013). The associations between personality, diet and body mass index in older people. *Health Psychology, 32*(4), 353–360.
Powell, N. D., Sloan, E. K., Bailey, M. T., Arevalo, J. M., Miller, G. E., Chen, E., et al. (2013). Social stress up-regulates inflammatory gene expression in the PBMC transcriptome via β-adrenergic induction of myelopoiesis. *Proceedings of the National Academy of Sciences, 110*(41), 16574–16579.
Revelle, W. (2019). *psych: Procedures for psychological, psychometric, and personality research.* Evanston, IL: Northwestern University. R package version 1.9.12 https://CRAN.R-project.org/package=psych.
Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspective on Psychological Science, 2*, 313–345.
Ryff, C. D., & Krueger, R. F. (Eds.). (2018). *The Oxford handbook of integrative health science.* Oxford University Press.
Smith, T. W., Williams, P. G., & Segerstrom, S. C. (2015). Personality and physical health. In M. Mikulincer, P. R. Shaver, M. L. Cooper, & R. J. Larsen (Eds.). *APA handbook of personality and social psychology, vol.4: Personality processes and individual differences* (pp. 639–661). Washington, DC: American Psychological Association.
Trapnell, P. D., & Wiggins, J. S. (1990). Extension of the interpersonal adjective scales to include the Big Five dimensions of personality. *Journal of Personality and Social Psychology, 59*, 781–790.
Turcotte, L., & Verneris, M. (2016). Is it better to be rich or relaxed? Sociobiology meets bone marrow transplant. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research, 22*(1), 6–8.
Vedhara, K., Gill, S., Eldesouky, L., Campbell, B. K., Arevalo, J. M. G., Ma, J., et al. (2015). Personality and gene expression: Do individual differences exist in the PBMC transcriptome. *Psychoneuroendocrinology, 52*, 72–82.