

ORIGINAL ARTICLE

Admixture mapping of genetic variants for uterine fibroids

Kui Zhang¹, Howard Wiener² and Brahim Aissani²

Uterine leiomyoma (UL) are benign neoplasms arising from the smooth muscle cells of the uterus. One of the established risk factors for UL is African American ethnicity. Studies have consistently shown that African Americans have two to three times higher risk compared with that of non-Hispanic Whites. However, there is still no adequate explanation for the higher risk among African Americans. To investigate the genetic contribution to the observed difference between the African American and European American populations, we conducted an admixture scan in 525 eligible African American women participants to the NIEHS uterine fibroid study (NIEHS-UFS). In models with no stratification, we found multiple genomic regions showing significant and suggestive evidence of association, with chromosomal band 2q32.2 at rs256552 showing the highest score (Z -score = 7.86, Bonferroni adjusted P -value = 5.5×10^{-12}) consistent with the suggestive evidence reported for this genomic region in the Black Women's Health Study. However, in models stratified by the body mass index (BMI) covariate, chromosome 1q42.2 was the sole genomic region that consistently showed suggestive associations across the BMI categories tested (Z -scores ≤ -3.96 , Bonferroni adjusted P -values ≤ 0.107). In age-stratified models, a significant association was observed in the older category (age > 40) reaching a Z -score of 6.44 (Bonferroni-adjusted P -value = 1.64×10^{-7}) at rs256552. The mean percentage of European ancestry among cases was lower than that among controls in the NIEHS-UFS study. However, our study did not show a significant association between mean percentage of European ancestry and UL.

Journal of Human Genetics (2015) 60, 533–538; doi:10.1038/jhg.2015.60; published online 4 June 2015

INTRODUCTION

Uterine leiomyoma (UL), or fibroids, are benign neoplasms arising from the smooth muscle cells of the uterus. It is believed that these tumors develop in the majority of American women by the time they reach menopause and become symptomatic in ~25%.¹ Despite their benign nature, UL are responsible for significant gynecologic morbidities including excessive bleeding, pelvic pain, urinary incontinence, infertility and pregnancy complications.^{2,3} Cumulative exposure to estrogen is believed to be a major etiologic factor,⁴ and factors that may influence the hormonal milieu, such as obesity, are also believed to be associated with the risk.⁵ However, the clearly established non-genetic risk factors are age (increasing risk with increasing premenopausal age), menopause (risk decreases with menopause) and African American ethnicity (higher risk compared with that of non-Hispanic Whites).⁶ There is still no adequate explanation for the higher risk among African Americans. The estimated cumulative incidence remained significantly higher for African Americans after controlling for parity and body mass index (BMI) in an ultrasound screening study.⁶

The genetic basis of susceptibility to UL is not well understood. Compelling evidence for a genetic liability in the development of

these tumors have been demonstrated in familial aggregation studies and twin studies;^{7–10} however, findings from genome-wide genome scans (including linkage and admixture mapping scans) and genome-wide association studies reported to date^{11–14} have not converged on a plausible mechanism of pathogenesis. Epidemiologic studies of UL in American women enrolled in the Right From The Start cohort and the BioVU DNA repository replicated the results of genome-wide association studies in a Japanese population for two out of the three most significantly associated genes encoding *BET1L* (blocked early in transport 1 homolog) and *TNRC6B* (trinucleotide repeat containing 6B).^{15,16} Recent studies in NIEHS-UFS (National Institute of Environmental and Health Sciences-Uterine Fibroid Study) has implicated gene variants encoding components of the extracellular matrix¹⁷ and replicated the association with *TNRC6B*.

To investigate whether the African American ethnicity confers a higher risk for UL, we performed admixture mapping in a sample of 393 UL cases of African American descent enrolled in NIEHS-UFS. Admixture mapping provides an efficient way to identify genetic risk variants that underlie the differences in population disease

¹Departments of Biostatistics, University of Alabama at Birmingham, Birmingham, AL, USA and ²Departments of Epidemiology, University of Alabama at Birmingham, Birmingham, AL, USA

Correspondence: Dr K Zhang, Department of Biostatistics, University of Alabama at Birmingham, Ryals Bldg. 327H, School of Public Health, 1665 University Blvd, Birmingham, AL 35294-0022, USA.

E-mail: kzhang@uab.edu

or Dr B Aissani, Department of Epidemiology, University of Alabama at Birmingham, R217J, School of Public Health, 1665 University Blvd, Birmingham, AL 35294-0022, USA.

E-mail: baissani@uab.edu

Received 9 February 2015; revised 28 April 2015; accepted 7 May 2015; published online 4 June 2015

prevalence.^{18,19} The idea behind admixture mapping is simple: genetic risk factors for diseases will be more frequent in chromosome regions derived from the ancestral population with the higher disease rate. For recently admixed populations, such as African Americans, admixture-generated linkage disequilibrium extends many millions of base pairs.¹⁹ Therefore, only a few thousands of ancestry-informative markers that differ in frequency between the parental populations of the admixed population are sufficient to carry out admixture mapping studies.¹⁸ Recently, with the availability of high-density admixture marker maps for several admixed populations, including African Americans and Mexican Americans^{18,20} and the development of efficient statistical and computational methods for admixture mapping,^{19,21–26} this approach has drawn considerable attention and successfully identified multiple genetic risk variants for human complex traits (for example, hypertension,^{27,28} multiple sclerosis²⁹ and so on). With an additional set of 132 clinically ascertained UL-free controls of African American descent enrolled in NIEHS-UFS, we also estimated the percentage of European ancestry for each case and control and investigated if there was an association between mean percentage of European ancestry and UL.

Despite the success of admixture mapping in gene disease mapping, only one study has used this approach to identify genetic risk variants for UL.¹³ Wise *et al.*¹³ carried out a genome-wide admixture mapping using DNA samples from 2453 UL cases and 2102 controls in the Black Women's Health Study (BWHS). The mean percentage of European ancestry was found to be significantly associated with the risk for UL in BWHS and genomic regions at chromosomes 2q32–33, 4p16 and 10q26 showed suggestive evidence for association with UL; however, these findings did not overlap those of recent genome-wide association studies in the Japanese¹² and North American¹⁴ populations. Our study did not show a significant association between mean percentage of European ancestry and UL. After stratifying by the BMI, our admixture mapping showed suggestive-to-significant associations at chromosome 1q42.2.

MATERIALS AND METHODS

Study population

Detailed characteristics of the study population have been reported,^{30,31} only those relevant to the present study will be described. Briefly, a random sample of women, aged 35–51 years, was selected from a computerized list of members of a prepaid urban health plan for enrollment in the NIEHS-UFS.⁶ Demographic data were collected by self-administered questionnaire. Reproductive and gynecologic history data were collected during a telephonic interview. About 92% of the study population was self-identified as African Americans or non-Hispanic European Americans. Of the enrolled women who were premenopausal and were screened for UL ($n = 1119$), 1045 (93%) had available DNA specimens and among these 574 (54.9%) self-identified as African Americans. The NIEHS-UFS and the present sub-study were approved by the Human Subject's Review boards at the NIEHS, George Washington University and University of Alabama at Birmingham, respectively. Participants gave written informed consents.

We used Discriminant Analysis of Principal Components³² of genetic data generated from a set of 2682 single-nucleotide polymorphism (SNPs) that were common to our data and to the majority of the reference populations of the International HapMapIII Project to assess population membership for each of the 558 eligible NIEHS-UFS individuals (417 cases and 141 controls, the only eligible AA controls) described in a previous report.³¹

Covariates

The covariates included age, parity after age 25 (earlier births were not significantly related to fibroid development in the NIEHS-UFS),^{33,34} BMI and physical activity.³⁰

Ascertainment

Fibroid status was assessed by ultrasound screening at baseline or by medical record review in 80% of the participants. For those women who had a pelvic ultrasound examination recently at the health plan, the radiology records from that examination were used to assess fibroid status. The remaining premenopausal participants were asked to have a pelvic ultrasound examination at the primary care site. Women for whom neither ultrasound nor medical record review could be conducted were excluded from the analyses. Both a transabdominal and a transvaginal ultrasound examination were performed. The abdominal portion evaluated fibroid change arising from the upper uterus that would not be readily seen with the transvaginal approach alone.

Ancestry informative markers and quality control

Blood DNA was available for 558 eligible participants. Illumina (San Diego, CA, USA) iSelect assay was used to genotype the selected SNPs at the Hudson Alpha Institute for Biotechnology, Huntsville, AL, USA as described.³¹ We selected and genotyped 1583 MALD (mapping by admixture linkage disequilibrium) SNPs. In selecting this subset of SNPs from the published set of 3011 MALD SNPs,¹⁸ we prioritized those SNPs overlapping the candidate regions (at chromosomes 2q37, 3p21, 5p13, 10p11, 11p15, 12q14 and 17q25) identified via linkage scan in the affected sister study,¹⁴ or by admixture mapping (2q32–33, 4p16 and 10q26) in BWHS.¹³ We extended this selected set with MALD SNPs from randomly selected chromosomes (chromosomes 1, 15 and 21) that we were able to include in the iSelect assay.

Admixture mapping

We used ADMIXMAP and assessed individual ancestry proportions to scan the genomic regions of African ancestry that differed significantly from the genome average.²³ ADMIXMAP (<http://homepages.ed.ac.uk/pmckeigu/admixmap/>) is a free program that implements a Markov chain Monte Carlo algorithm for modeling admixture using multilocus genotype data and individuals from admixed populations.¹⁹ The applications of ADMIXMAP include inferring distribution of individual admixture proportions and identifying genetic risk variants by admixture mapping. The genetic markers used in ADMIXMAP have been chosen to have different frequencies between the parental populations of the admixed African American population under study.¹⁹ The program was downloaded and run using 1000 burn-in iterations and 9000 follow-on iterations and outputting sampled values in every 25 iterations. The allele frequency of each MALD SNP in African and European populations was calculated using allele counts provided in the Supplementary Table of Smith *et al.*,¹⁸ and used as input for ADMIXMAP.¹⁸ The HapMap Phase II recombination rate on GRCh37 was used to calculate the genetic distance between the eligible MALD SNPs. These recombination rates are the leftover of HapMap phase II genetic map from human genome build b35 to GRCh37 and can be downloaded from <http://hapmap.ncbi.nlm.nih.gov/>. The score test implemented in ADMIXMAP was used to test for association between UL and each locus by comparing the observed and expected ancestry proportions at the studied locus in cases only. This test also is similar to that used in Wise *et al.*¹³ Statistical significance was assessed using Z -scores and Bonferroni adjusted P -values. A negative Z -score indicates that lower European ancestry at a particular locus compared with the genome-wide average is associated with higher UL risk, whereas a positive Z -score indicates that higher European ancestry is associated with higher UL risk. In Wise *et al.*,¹³ a $|Z\text{-score}| \geq 3$ indicated suggestive evidence, while a $|Z\text{-score}| \geq 4$ indicated a statistically significant result. Here, a $|Z\text{-score}| \geq 3$ indicates a suggestive result while a Bonferroni-adjusted $P\text{-value} \leq 0.05$ which corresponds to a $|Z\text{-score}| \geq 4.133$ indicates a statistically significant result. Since BMI is a confounding factor and ADMIXMAP does not allow for covariates to be modeled, we also performed BMI-stratified analyses in ADMIXMAP (BMI ≤ 25 , BMI > 25 and BMI ≤ 30 , BMI > 30 and BMI ≤ 35 or BMI > 35).

To test the association between the European heritage and the risk of UL, we estimated the proportions of ancestral populations for each individual and used logistic regression with or without age as a covariate to estimate the odds ratio and its 95% confidence interval. Logistic regression models were also used to perform subgroup analyses based on age (age ≤ 40 or age > 40), parity, BMI

(BMI ≤ 25 , BMI > 25 and BMI ≤ 30 , BMI > 30 and BMI ≤ 35 or BMI > 35), or physical activity (low, moderate, high or very high).

RESULTS

Thirty-three of the 558 typed individuals did not pass the genotyping quality control filter (call rate $> 90\%$); this resulted in a study population of 525 individuals (393 UL cases and 132 controls) of African American descent available for the analysis.

The characteristics of the 525 African American participants available for admixture mapping are summarized in Table 1. Mean age was 42.5 for UL cases and 40.5 for controls. The mean age of the studied participants in the cross-sectional NIEHS study differed from the mean age at diagnosis in the prospective BWHS, which was 38.3 and 42.9 years for cases and controls, respectively. Because age is an established risk factor (increased risk with increasing premenopausal age), only age-adjusted results will be compared between the two studies. About 47.6% of cases (187 out of 393) and 54.6% (72 out of 132) of controls had given birth after age 25. The mean BMI was 30.9 for UL cases and 29.5 for controls. Among the UL cases, 45.8% had BMI > 30 , while only 37.9% controls had BMI > 30 . Also 29.4% of UL cases and 41.2% control had high or very-high physical activity. Although the mean BMI difference between cases and controls was not significant, admixture mapping models with and without stratification by the BMI covariate were tested because (i) admixture mapping was conducted in a case-only design and (ii) both BWHS³⁵ and NIEHS-UFS⁶ reported an inverse J-shaped association (non-dose response) between BMI and UL. Also, even though we conducted admixture mapping in a case-only design, models with and without stratification by age were tested to allow comparisons of our results with those reported in BWHS.¹³

Marker quality control

Of the 1583 SNPs selected for typing, 182 (11.4%) were excluded because the assay failed, the SNP was monomorphic or rare, or the SNP deviated from Hardy–Weinberg equilibrium in the controls; this yielded 1401 MALD SNPs available for analyses. Of note, a large

number of the MALD SNPs that we have selected for this study are positional markers spanning the linkage and admixture peaks and flanking regions. The other non-positional MALD SNPs were from a random selection of chromosomes.

Global ancestry analysis

The mean percentage of European ancestry obtained from ADMIXMAP for UL cases and controls were 20.7 and 21.3% (Table 2), respectively. The mean differences with and without adjustment for age from the logistic regression were -1.16 and -1.27% , respectively. Subgroup analysis of global ancestry showed that the mean percentage of European ancestry was lower among UL cases than controls in most comparisons. Although the mean percentage of ancestry among UL cases and controls obtained from our study were similar to those from Wise *et al.*,¹³ our study failed to show evidence for significant association between percentage of European ancestry and UL.

Admixture mapping analysis

The detailed results of the admixture mapping scan obtained from ADMIXMAP in a case-only design are shown in Supplementary Table S1. For each statistically significant or suggestive peak, the SNP with the highest Z-score is listed in Supplementary Table S2. We have found multiple statistically significant or suggestive peaks at chromosomes 1, 2, 3, 4, 5, 11, 12 and 17 in models with no control for BMI. However, in models stratified by the BMI covariate, most statistically significant or suggestive peaks disappeared. From the quantile–quantile plot (Q–Q plot) of the Z-scores obtained from ADMIXMAP using all 393 cases (Supplementary Figure S1), we observed an inflation of the test statistic at the top of the plot. The inflations of the test statistic can be explained by the confounding effect of BMI since no such inflations were observed in the Q–Q plots for models stratified by BMI (Supplementary Figure S2). Therefore, the models with BMI stratification are used for the interpretation of our study results. Furthermore, using discriminant analysis of principal components, we show that the entire case and control sample used in this study clusters only with the HapMap III reference populations of the African ancestry Yoruba in Ibadan, Nigeria (YRI), Luhya in Webuye, Kenya (LWK) and Maasai in Kinyawa, Kenya (MKK) or African ancestry in Southwest USA (ASW) and not with any other population tested, thus indicating that spurious results due to population structure are unlikely (Supplementary Figure S3).

Chromosome 1q42.2 was the sole genomic region that showed suggestive-to-significant associations at rs7546784 ($-6.81 \leq Z\text{-scores} \leq -3.96$, Bonferroni-adjusted $1.40 \times 10^{-8} \leq P\text{-value} \leq 0.107$) consistently across the four BMI categories and age categories (Table 3). In agreement with African ancestry being at risk, the association indicates that lower European ancestry at chromosome 1q42.2 increases the risk for UL. It is not clear if this region of chromosome 1q42.2 was screened in BWHS since that study too has not used the full set of available MALD SNPs.¹⁸ The signal on chromosome 2q32.2 peaked at rs256552 around position 190.7 Mb with a Z-score of 7.86 and Bonferroni-adjusted P-value of 5.5×10^{-12} in models with no control for BMI. Although the association was not significant across the four BMI categories, suggestive evidence of association (Z-score of 3.04 in the BMI category 30 to 35 and Z-scores around 2 for the other BMI categories) was observed. In age-stratified models, a significant association was observed in the older category (age > 40) reaching a Z-score of 6.44 (Bonferroni-adjusted P-value = 1.64×10^{-7}).

We noticed that candidate SNPs rs7546784 and rs256552 were both common SNPs in African Americans (Table 4). Further re-evaluation of the association by logistic regression with adjustment for BMI and

Table 1 Baseline characteristics of African American women with uterine fibroids (cases) and controls

	Cases (n = 393)		Controls (n = 132)		P ^a
	Mean	Count (%)	Mean	Count (%)	
Age, years	42.5		40.5		7.5×10^{-7}
Parity after age 25					0.45
0		206 (52.4)		60 (45.4)	
≥ 1		187 (47.6)		72 (54.6)	
BMI	30.9		29.4		0.07
BMI ≤ 25		88 (22.4)		44 (33.3)	
BMI > 25 and BMI ≤ 30		125 (31.8)		38 (28.8)	
BMI > 30 and BMI ≤ 35		80 (20.4)		26 (19.7)	
BMI > 35		100 (25.4)		24 (18.2)	
Physical activity					0.05
Low		143 (36.6)		46 (35.1)	
Moderate		133 (34.0)		31 (23.7)	
High		56 (14.3)		27 (20.6)	
Very High		59 (15.1)		27 (20.6)	

Abbreviation: BMI, body mass index.

^aP-values from Fisher's exact test for categorical covariates and from t-test for continuous covariates.

Table 2 Mean percentage of European ancestry^a among African American women with uterine fibroids (cases) and controls

	Cases		Controls		Unadjusted mean	Age-adjusted mean	Age-adjusted	P-value
	No.	mean %	No.	mean %	difference (β)	difference (β)	95% CI of (β)	
All cases and controls	393	20.7	132	21.3	-1.16	-1.27	(-4.02, 1.48)	0.362
Case with age <40, all controls	126	20.5	132	21.3	-1.72	-1.46	(-5.22, 2.31)	0.449
Case with age \geq 40, all controls	267	20.8	132	21.3	-0.86	-0.74	(-4.00, 2.52)	0.657
Cases and controls, parity = 0	206	21.0	60	20.6	0.81	0.44	(-3.64, 4.52)	0.832
Cases and controls, parity \geq 1	197	20.4	72	22.0	-2.95	-2.84	(-6.61, 0.93)	0.140
Cases and controls, BMI \leq 25	88	21.2	44	21.0	0.50	0.47	(-4.74, 5.67)	0.861
Cases and controls, BMI > 25 and BMI \leq 30	125	21.3	38	21.7	-0.60	-1.25	(-6.07, 3.57)	0.611
Cases and controls, BMI > 30 and BMI \leq 35	80	20.4	26	21.5	-2.22	-2.33	(-8.67, 4.00)	0.470
Cases and controls, BMI > 35	100	19.7	24	21.3	-3.31	-2.90	(-9.32, 3.52)	0.376
Cases controls, low physical activity	143	20.6	46	20.1	1.14	1.92	(-3.04, 6.88)	0.448
Cases controls, moderate physical activity	133	21.1	31	21.8	-1.25	-1.99	(-7.13, 3.17)	0.450
Cases controls, high physical activity	115	20.6	54	21.8	-2.32	-3.17	(-7.83, 1.49)	0.182

Abbreviations: BMI, body mass index; CI, confidence interval.
^aEstimated using the program ADMIXMAP.

Table 3 Top candidate SNP from admixture mapping^a in 393 African American women with uterine fibroids

SNP ID	Chr. ^b	Genetic position	Physical position ^c	Z-score							Gene	Function
				all cases	BMI 25	BMI > 25 and BMI \leq 30	BMI > 30 and BMI \leq 35	BMI > 35	Age			
								\leq 40	> 40			
rs7546784	1q42.2	252.46	233388281	-4.02	-3.96	-4.10	-4.70	-6.81	-4.87	-4.08	PCNXL2	Intron
rs256552	2q32.2	200.49	190742374	7.86	2.30	2.31	3.04	1.98	2.83	6.44	PMS1	Downstream

Abbreviations: BMI, body mass index; SNP, single-nucleotide polymorphism.

^aPerformed with the program ADMIXMAP.

^bChromosomal band.

^cAll physical positions in this table and in the paper are shown according to the Human Genome build GRCh37.

Table 4 Re-evaluation of risk associated with candidate admixture mapping SNPs for uterine fibroids in a case and control study

SNP ^a	Alleles ^b		Genotype frequency									Logistic regression			
	1	2	All (525)			Cases (393)			Controls (132)			11 vs 22		12 vs 22	
			11	12	22	11	12	22	11	12	22	OR (95% CI), P-value		OR (95% CI), P-value	
rs7546784	C	G	0.032	0.385	0.583	0.023	0.394	0.583	0.061	0.356	0.583	0.26 (0.086, 0.79), 0.017		1.03 (0.67, 1.58), 0.89	
rs256552	A	G	0.103	0.448	0.450	0.112	0.430	0.458	0.076	0.500	0.424	1.01 (0.46, 2.90), 0.98		0.78 (0.51, 1.20), 0.25	

Abbreviations: CI, confidence interval; OR, odds ratio; SNP, single-nucleotide polymorphism; UL, uterine leiomyoma.

Association of genotypes at candidate SNPs with the risk of UL, or fibroids, in a case and control sample of African American women participants to the NIEHS-UFS. The OR and the 95% CI associated with genotypes were estimated in logistic regression models adjusted for age and body mass index, with the homozygous genotype for the most frequent allele (2/2) in the controls used as the referent genotype. rs7546784 and rs256552 were the only candidate SNPs from the admixture mapping study that showed significant or suggestive evidence, respectively, for the association with UL in case-only models.

^aSNP.

^bMinor (1) and major (2) alleles.

age showed that only rs7546784 was associated with UL (Table 4). Removing the age and BMI from the logistic regression did not affect significantly the results (Supplementary Table S3). Finally, except for the chromosome 1q42.2 signal, no linear relationship was observed between the Z-score and BMI. Changes to the Z-score sign across the BMI categories were observed in only a few instances (at rs10473043 on chromosomal band 5p13.2 and at rs1894223 on chromosomal band 11q24.2).

DISCUSSION

We performed an admixture mapping study of UL in 393 African American UL cases using 1401 ancestry informative SNPs. To our knowledge, this is the second report of admixture mapping for UL after that of BWHS.¹³ Although we have found several statistically significant peaks on chromosomes 1, 2, 3, 4, 5, 11, 12, 15 and 17, after controlling for the BMI or the age covariate, a sub-region of chromosome 1q42.2 tagged by rs7546784 was the sole region that

consistently showed significant associations across the BMI or age categories. The negative *Z*-score observed at chromosome 1q42.2 is consistent with African ancestry being at higher risk.³⁶ To our knowledge, this locus was not highlighted in any previous genome-wide association studies or admixture study of UL. However, in models with no control for BMI, the strongest association peaked at chromosome 2q32.2 at a location close to the peak reported in BWHS for rs2271767 (*Z*-score=3.69) at position 192.3.¹³ Using ADMIX-MAP, Wise *et al.* have found a suggestive evidence for a 2q32-q33 region around rs7573626 (physical position=206.3; *Z*-score=3.42 using all UL cases), rs2271767 (physical position =192.3, *Z*-score=3.69 using cases aged <35 years), rs6710083 (physical position=192.3, *Z*-score=4.10 using surgical cases) and rs920249 (physical position=201.3, *Z*-score=5.10 using surgical cases aged<35 years).

We were unable to replicate the significant association between the mean percentage of European ancestry and UL reported in BWHS. As indicated by the results shown in Table 2, the mean percentage of European ancestry among cases was lower than that among controls in the NIEHS study, which is in the same direction to that reported in BWHS. However; the mean difference in the NIEHS (means of European ancestry of 20.7 and 21.3% in cases and controls, respectively) was smaller than that in the BWHS (means of 20 and 21.6% for cases and controls, respectively). This smaller mean difference, together with the smaller size of our study sample compared to that of BWHS, account for the lack of statistical significance in our study.

The comparison of the results between the two studies should be interpreted with cautions because: (i) all the participants to NIEHS-UFS were more than 35 years of age at enrollment and only the mean percentage of European ancestry was reported for this age category in BWHS; (ii) contrasting with our study which tested 1401 SNPs on a smaller sample of 525 individuals (393 cases and 132 controls), the data from the BWHS were based on the 30 most informative MALD SNPs³⁷ typed in a larger sample (2453 cases and 2102 controls); and (iii) NIEHS-UFS ascertained clinically the enrolled controls in a cross-sectional study whereas BWHS is a prospective cohort study that assessed the UL affection status in the controls through a questionnaire at the end of the follow-up period. It is clear; however, that given the estimated effect size of 0.086, our study based on 393 cases and 132 controls had limited power to detect difference in mean percentage of European ancestry.

Most of the MALD SNPs selected for the present study map within candidate genomic regions identified by admixture mapping in BWHS or by linkage scan in the affected sister study,¹⁴ thus, back replication of the present findings in these discovery studies is an important step toward their validation.

The interpretation of our results from the models with and without BMI stratification and their comparison with those of BWHS remains difficult. First, to enable comparisons between the two studies, we conducted stratified analyses based on the BMI categories used in BWHS; however, the association between BMI and UL peaked in the BMI range 30–34 in NIEHS-UFS³⁰ whereas in BWHS, the peak was observed in the BMI range 25–29.³⁵ Second, as a proxy measure for the steroid-producing body fat, BMI may affect the association with these hormone-dependent tumors through the causal pathway by decreasing the serum level of sex hormone-binding globulin and consequent increase in bioavailability of free steroids,³⁸ through pleiotropy or through linkage disequilibrium between a UL and an obesity genes.³⁹ The interpretation of these results depends on the validity of these alternative etiologic models. If BMI is a risk factor operating in the causal pathway then the data obtained from models

with no BMI-stratification are more robust because they are based on the full BMI variance. In the alternative that BMI rather confounds the association between the actual UL genes and the UL outcome, then confidence in the data derived from models with BMI stratification is greater. In either model, the locus on chromosome 1q42.2 remains an attractive candidate.

It is still early to draw conclusions about potential candidate UL genes. The SNP that captured the association at chromosome 1q42.2 is an intronic variant in *PCNXL2* (pecanex-like 2 gene), which encodes a transmembrane protein with uncharacterized function. The strongest signal on chromosome 2q32.2 was captured by a variant located in the immediate 3'UTR downstream region of *PMS1* (post-meiotic segregation increased 1) whose product belongs to the DNA mismatch repair mutL/hexB family. This locus and the surrounding loci within a 1–2 Mb interval from position 190.74 Mb on chromosome 2q, which is the average physical distance of the admixture mapping resolution, are target candidate loci for follow-up fine mapping. Importantly, with the suggestive association (*Z*-score=3.69) reported in BWHS at position 192.3 Mb, this is the sole candidate region that we were able to replicate, albeit with moderate association strength, in NIEHS-UFS.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We are indebted to the Principal Investigator of the NIEHS uterine fibroid study, Dr Donna Day Baird, for providing access to the NIEHS-UFS resources. The authors would like to express their gratitude to all participants in NIEHS-UFS. We thank Drs Cynthia Morton and Lauren Wise for disclosing the positions of the linkage and admixture peaks before publication of their studies. This study was supported by grant (R01-HD-064398) (to BA) from the Eunice Kennedy Shriver National Institute of Child Health and Human Development.

- 1 Cramer, S. F. & Patel, A. The frequency of uterine leiomyomas. *Am. J. Clin. Pathol.* **94**, 435–438 (1990).
- 2 Stewart, E. A. Uterine fibroids. *Lancet* **357**, 293–298 (2001).
- 3 Walker, C. L. & Stewart, E. A. Uterine fibroids: the elephant in the room. *Science* **308**, 1589–1592 (2005).
- 4 Andersen, J. Growth factors and cytokines in uterine leiomyomas. *Semin. Reprod. Endocrinol.* **14**, 269–282 (1996).
- 5 Schwartz, S. M., Marshall, L. M. & Baird, D. D. Epidemiologic contributions to understanding the etiology of uterine leiomyomata. *Environ. Health Perspect.* **108** (suppl 5), 821–827 (2000).
- 6 Baird, D. D., Dunson, D. B., Hill, M. C., Cousins, D. & Schectman, J. M. High cumulative incidence of uterine leiomyoma in black and white women: ultrasound evidence. *Am. J. Obstet. Gynecol.* **188**, 100–107 (2003).
- 7 Kurbanova, M., Koroleva, A. G. & Sergeev, A. S. Genetic analysis of the predisposition to uterine myoma. Prevalence and morbidity. *Genetika* **25**, 1122–1124 (1989).
- 8 Vikhlyayeva, E. M., Khodzhaeva, Z. S. & Fantschenko, N. D. Familial predisposition to uterine leiomyomas. *Int. J. Gynaecol. Obstet.* **51**, 127–131 (1995).
- 9 Treloar, S. A., Martin, N. G., Dennerstein, L., Raphael, B. & Heath, A. C. Pathways to hysterectomy: insights from longitudinal twin research. *Am. J. Obstet. Gynecol.* **167**, 82–88 (1992).
- 10 Luoto, R., Kaprio, J., Rutanen, E. M., Taipale, P., Perola, M. & Koskenvuo, M. Heritability and risk factors of uterine fibroids—the Finnish Twin Cohort study. *Maturitas* **37**, 15–26 (2000).
- 11 Gross, K. L., Panhuyzen, C. I., Kleinman, M. S., Goldhammer, H., Jones, E. S., Nassery, N. *et al.* Involvement of fumarate hydratase in nonsyndromic uterine leiomyomas: genetic linkage analysis and FISH studies. *Genes Chromosomes Cancer* **41**, 183–190 (2004).
- 12 Cha, P. C., Takahashi, A., Hosono, N., Low, S. K., Kamatani, N., Kubo, M. *et al.* A genome-wide association study identifies three loci associated with susceptibility to uterine fibroids. *Nat. Genet.* **43**, 447–450 (2011).
- 13 Wise, L. A., Ruiz-Narvaez, E. A., Palmer, J. R., Cozier, Y. C., Tandon, A., Patterson, N. *et al.* African ancestry and genetic risk for uterine leiomyomata. *Am. J. Epidemiol.* **176**, 1159–1168 (2012).
- 14 Eggert, S. L., Huyck, K. L., Somasundaram, P., Kavalla, R., Stewart, E. A., Lu, A. T. *et al.* Genome-wide linkage and association analyses implicate FASN in predisposition to uterine leiomyomata. *Am. J. Hum. Genet.* **91**, 621–628 (2012).

- 15 Edwards, T. L., Michels, K. A., Hartmann, K. E. & Velez Edwards, D. R. *BET1L* and *TNRC6B* associate with uterine fibroid risk among European Americans. *Hum. Genet.* **132**, 943–953 (2013).
- 16 Edwards, T. L., Hartmann, K. E. & Velez Edwards, D. R. Variants in *BET1L* and *TNRC6B* associate with increasing fibroid volume and fibroid type among European Americans. *Hum. Genet.* **132**, 1361–1369 (2013).
- 17 Aissani, B., Zhang, K. & Wiener, H. Follow-up to genome-wide linkage and admixture mapping studies implicates components of the extracellular matrix in susceptibility to and size of uterine fibroids. *Fertil. Steril.* **103**, 528–34 e13 (2015).
- 18 Smith, M. W., Patterson, N., Lautenberger, J. A., Truelove, A. L., McDonald, G. J., Waliszewska, A. *et al.* A high-density admixture map for disease gene discovery in African Americans. *Am. J. Hum. Genet.* **74**, 1001–1013 (2004).
- 19 Patterson, N., Hattangadi, N., Lane, B., Lohmueller, K. E., Hafler, D. A., Oksenberg, J. R. *et al.* Methods for high-density admixture mapping of disease genes. *Am. J. Hum. Genet.* **74**, 979–1000 (2004).
- 20 Collins-Schramm, H. E., Chima, B., Morii, T., Wah, K., Figueroa, Y., Criswell, L. A. *et al.* Mexican American ancestry-informative markers: examination of population structure and marker characteristics in European Americans, Mexican Americans, Amerindians and Asians. *Hum. Genet.* **114**, 263–271 (2004).
- 21 Falush, D., Stephens, M. & Pritchard, J. K. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* **164**, 1567–1587 (2003).
- 22 Montana, G. & Pritchard, J. K. Statistical tests for admixture mapping with case-control and cases-only data. *Am. J. Hum. Genet.* **75**, 771–789 (2004).
- 23 Hoggart, C. J., Shriver, M. D., Kittles, R. A., Clayton, D. G. & McKeigue, P. M. Design and analysis of admixture mapping studies. *Am. J. Hum. Genet.* **74**, 965–978 (2004).
- 24 Tang, H., Coram, M., Wang, P., Zhu, X. & Risch, N. Reconstructing genetic ancestry blocks in admixed individuals. *Am. J. Hum. Genet.* **79**, 1–12 (2006).
- 25 Price, A. L., Tandon, A., Patterson, N., Barnes, K. C., Rafaels, N., Ruczinski, I. *et al.* Sensitive detection of chromosomal segments of distinct ancestry in admixed populations. *PLoS Genet.* **5**, e1000519 (2009).
- 26 Sankararaman, S., Sridhar, S., Kimmel, G. & Halperin, E. Estimating local ancestry in admixed populations. *Am. J. Hum. Genet.* **82**, 290–303 (2008).
- 27 Zhu, X. & Cooper, R. S. Admixture mapping provides evidence of association of the *VNN1* gene with hypertension. *PLoS One* **2**, e1244 (2007).
- 28 Rao, W., Ma, Y., Ma, L., Zhao, J., Li, Q., Gu, W. *et al.* High-resolution whole-genome haplotyping using limited seed data. *Nat. Methods* **10**, 6–7 (2013).
- 29 Reich, D., Patterson, N., De Jager, P. L., McDonald, G. J., Waliszewska, A., Tandon, A. *et al.* A whole-genome admixture scan finds a candidate locus for multiple sclerosis susceptibility. *Nat. Genet.* **37**, 1113–1118 (2005).
- 30 Baird, D. D., Dunson, D. B., Hill, M. C., Cousins, D. & Schectman, J. M. Association of physical activity with development of uterine leiomyoma. *Am. J. Epidemiol.* **165**, 157–163 (2007).
- 31 Aissani, B., Wiener, H. & Zhang, K. Multiple hits for the association of uterine fibroids on human chromosome 1q43. *PLoS One* **8**, e58399 (2013).
- 32 Jombart, T., Devillard, S. & Balloux, F. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet.* **11**, 94 (2010).
- 33 Dunson, D. B. & Baird, D. D. A proportional hazards model for incidence and induced remission of disease. *Biometrics* **58**, 71–78 (2002).
- 34 Baird, D. D. & Dunson, D. B. Why is parity protective for uterine fibroids? *Epidemiology* **14**, 247–250 (2003).
- 35 Wise, L. A., Palmer, J. R., Spiegelman, D., Harlow, B. L., Stewart, E. A., Adams-Campbell, L. L. & Rosenberg, L. Influence of body size and body fat distribution on risk of uterine leiomyomata in U.S. black women. *Epidemiology* **16**, 346–354 (2005).
- 36 Baird, D. D., Dixon, D. & Sandler, D. P. African Americans at higher risk than whites for uterine fibroids: ultrasound evidence. *Am. J. Epidemiol.* **147**, S90 (1998).
- 37 Ruiz-Narvaez, E. A., Rosenberg, L., Wise, L. A., Reich, D. & Palmer, J. R. Validation of a small set of ancestral informative markers for control of population admixture in African Americans. *Am. J. Epidemiol.* **173**, 587–592 (2011).
- 38 Schwartz, S. M. Epidemiology of uterine leiomyomata. *Clin. Obstet. Gynecol.* **44**, 316–326 (2001).
- 39 Aissani, B. Confounding by linkage disequilibrium. *J. Hum. Genet.* **59**, 110–115 (2014).

Supplementary Information accompanies the paper on Journal of Human Genetics website (<http://www.nature.com/jhg>)