Type 2 Diabetes Susceptibility Variants Contribute to Breast Cancer Risk

Mary Helen Black, PhD; Shuwei Li, PhD; Brice Sarver, PhD; Aaron Elliott, PhD; Hsiao-Mei Lu, PhD

Ambry Genetics, Aliso Viejo, CA

BACKGROUND

- Type 2 diabetes mellitus (T2DM) is a known risk factor for breast cancer (BC). However, whether these conditions share a common genetic etiology remains to be elucidated.
- Analyses to date have focused on T2DM GWAs-identified variants, nearly all of which were intronic or intergenic (Chen F. *et al.* 2011; Hou N. *et al.* 2012; Zhao Z. *et al.* 2016).
- Recently, exome analysis in nearly 100,000 individuals detected 21 coding variants that predispose to T2DM at genome-wide significance (Fuchsberger C. *et al.* 2016).
- We hypothesized that these coding variants may be associated with BC and/or clinical features of BC.

METHODS

- Whole exome sequencing in 9,639 BC cases and 3,988 controls was performed among patients referred for genetic testing at a single diagnostic laboratory in 2014-2015 (AmbryShare™)
- Demographic and clinical history data (Table 1) were collected from test requisition forms and clinical notes provided by ordering clinicians.
- Sample preparation, data processing and bioinformatics pipeline for alignment and joint variant calling were previously described (Mu *et al.* 2016). Samples with >75% of Q30 bases, mean base quality >30 and >80% of bases >10x in target exon regions were retained for analysis. Evaluating coverage for each site across samples, positions for which >20% of individuals had coverage <10x were filtered out as low coverage sites. Mean coverage: 29.4x (412.9x).
- We also utilized 1000 Genomes Phase 3 (G1K3, n=2,455) and UK10K (n=3,568) with standard QC, as additional controls in order to confirm our findings.
- Kinship-based inference for genome-wide association studies (KING) was used to identify and exclude individuals based on cryptic relatedness. We also performed principal component analysis (PCA) to infer ancestry and categorize patients into G1K3-defined racial/ethnic groups (Fig 1).
- From all internal and external data sources, we extracted genotypes for 21 T2DM-associated SNPs (Table 2) and assessed BC phenotype associations using logistic regression models estimated within genetically inferred G1K3-defined populations and adjusted for sex and 3 ancestry PCs.
- Unweighted genetic risk scores (GRS) were tested for association with BC phenotypes, within ethnic populations and adjusted as described above. SNPs with 0% RAF were not included in GRS models.

AMBRYSHARE™ BREAST CANCER

Table 1. Characteristics of Breast Cancer Cases (n=9,639)

	EUR (n=7,235)	AFR (n=968)	AMR (n=809)	EAS (n=522)	SAS (n=105)
Age at diagnosis [years]					
Mean 4 SD	49.6411.6	46.9411.2	46.0410.7	45.5410.4	42.548.6
Gender (%)					
Female	98.6%	99.0%	99.3%	99.4%	96.2%
Male	1.4%	1.0%	0.7%	0.6%	3.8%
Breast Cancer histology (%)					
Ductal only	70.0%	74.8%	72.2%	73.9%	78.1%
Lobular only	5.6%	2.9%	4.9%	3.6%	1.9%
Ductal and lobular	1.7%	0.9%	0.5%	1.0%	1.0%
Not provided	22.7%	21.4%	22.4%	21.5%	19.0%
Breast Cancer hormone receptor status (%)					
$HR^+/HER2^-$	26.6%	20.6%	22.4%	28.5%	23.8%
$HR^-/HER2^+$	2.8%	3.0%	4.2%	1.9%	2.9%
$HR^+/HER2^+$ (TPBC)	5.2%	5.7%	7.8%	7.1%	7.6%
$HR^-/HER2^-$ (TNBC)	10.4%	21.7%	15.5%	9.4%	15.2%
Other/Not provided	55.0%	49.1%	50.2%	53.1%	50.5%
Personal history of other cancer (%)					
Yes	10.5%	5.3%	5.4%	5.0%	4.8%
No	89.5%	94.7%	94.6%	95.0%	95.2%
Personal history of other cancer, type (%)					
Ovarian	3.3%	0.9%	2.1%	1.3%	2.9%
Uterine/Endometrial	3.0%	2.1%	2.2%	2.1%	0.0%
Colorectal	1.7%	1.2%	0.5%	0.2%	0.0%
Other	3.3%	1.7%	1.2%	1.7%	1.9%
1^{st}-/2^{nd}-degree relative with any cancer (%)					
Yes	87.6%	80.9%	73.7%	70.1%	66.7%
No	2.4%	3.0%	5.4%	5.2%	7.6%
Not provided	10.0%	16.1%	20.9%	24.7%	25.7%
1^{st}-/2^{nd} -degree relative with breast or ovarian cancer (%)					
Yes	64.3%	56.9%	49.6%	45.2%	39.0%
No	25.7%	27.0%	29.5%	30.1%	35.2%
Not provided	10.0%	16.1%	20.9%	24.7%	25.7%
Carriers of mutations in 8 known breast cancer genes (%)					
≥1 mutation	6.1%	5.1%	5.3%	5.4%	4.8%
None	93.9%	94.9%	94.7%	94.6%	96.2%

RESULTS

Table 2. T2DM SNP Frequencies and Association with Breast Cancer among EUR (n=13,970)[a]

Chr	Gene	SNP	Variant		case RAF[b]	control RAF[b]	OR (95% CI)[c]	p-value[d]
2	*COBLL1*	rs7607980	c.2815A>G	p.N939D	86.9%	87.3%	0.99 (0.91 - 1.07)	0.73
2	*GCKR*	rs1260326	c.1337T>C	p.L446P	55.9%	58.4%	0.98 (0.93 - 1.03)	0.36
2	*THADA*	rs35720761	c.4814G>A	p.C1605Y	87.8%	87.6%	0.96 (0.89 - 1.04)	0.32
3	*PPARG*	rs1801282	c.34C>G	p.P12A	94.2%	89.3%	1.63 (1.48 - 1.79)	5.0×10^{-23}
4	*WFS1*	rs1801212	c.997G>A	p.V333I	72.1%	72.1%	0.95 (0.89 - 1.00)	0.06
4	*WFS1*	rs1801214	c.1500C>T	p.N500N	60.2%	60.4%	0.96 (0.91 - 1.02)	0.17
4	*WFS1*	rs734312	c.1832G>A	p.R611H	54.2%	54.2%	0.99 (0.94 - 1.05)	0.84
5	*PAM*	rs35658696	c.1688A>G	p.D563G	3.0%	5.0%	0.67 (0.58 - 0.76)	1.68×10^{-09}
5	*PPIP5K2*	rs36046591	c.3619A>G	p.S1207G	2.5%	4.9%	0.60 (0.53 - 0.69)	6.88×10^{-13}
6	*RREB1*	rs9379084	c.3511G>A	p.D1171N	87.3%	88.3%	0.94 (0.86 - 1.01)	0.11
7	*PAX4*	rs2233580	c.575G>A	p.R192H	0.0%	0.0%	-	-
8	*SLC30A8*	rs13266634	c.973C>T	p.R325W	71.2%	69.9%	1.03 (0.97 - 1.09)	0.39
9	*GPSM1*	rs60980157	c.1172C>T	p.S391L	76.0%	75.4%	1.01 (0.95 - 1.07)	0.74
11	*ABCC8*	rs757110	c.4105G>T	p.A1369S	36.3%	36.1%	1.03 (0.98 - 1.09)	0.29
11	*KCNJ11*	rs5215	c.1009G>A	p.V337I	36.0%	35.9%	1.03 (0.97 - 1.09)	0.31
11	*KCNJ11*	rs5219	c.67A>G	p.K23E	36.0%	35.9%	1.03 (0.98 - 1.09)	0.27
19	*TM6SF2*	rs58542926	c.499G>A	p.E167K	6.9%	7.1%	1.01 (0.91 - 1.11)	0.91
22	*ASCC2*	rs11549795	c.367G>A	p.V123I	93.8%	91.9%	1.23 (1.12 - 1.36)	3.61×10^{-05}
22	*ASCC2*	rs28265	c.1219G>C	p.D407H	8.2%	8.2%	1.06 (0.96 - 1.16)	0.26
22	*ASCC2*	rs36571	c.1267C>T	p.P423S	92.0%	91.9%	0.95 (0.86 - 1.04)	0.26
22	*MTMR3*	rs41278853	c.2879A>G	p.N960S	91.7%	91.8%	0.94 (0.85 - 1.03)	0.19

[a] Analysis among EUR included 7,235 cases and 6,735 controls (combined 2,666 internal controls, 501 G1K3, 3,568 UK10K).
[b] Risk allele for type 2 diabetes as defined by Fuschberger et al *Nature* 2016; RAF = Risk allele frequency.
[c] Odds Ratios (ORs) and 95% CI are computed under additive genetic models adjusted for sex and within-race/ethnicity PC1-3.
[d] p-values shown are nominal, and not corrected for multiple comparisons.

Figure 1. Ancestry PCs in EUR BC Cases and Controls

Figure 2. GRS Distribution in EUR Subjects

- *PPARG* P12A, *PAM* D563G, *PPIP5K2* S1207G are each significantly associated with BC at genome-wide significance in EUR (Table 2).
- Association with *PPARG* P12A was replicated in AMR and SAS only ($p=3.28 \times 10^{-6}$ and 6.28×10^{-3}, respectively; data not shown). No SNP associations were observed in EAS or AFR.
- Among EUR, *PPARG* P12A associated with increased risk for all clinical BC subtypes. *PAM* D563G and *PPIP5K2* S1207G associated with decreased risk for HR^-, $HER2^-$ and TNBC only.
- Results were similar in sensitivity analyses excluding males, females >50 years of age, and/or those carrying other high- and moderate-penetrance cancer gene mutations.
- Unweighted GRS mean4SD among cases and controls was 26.943.5 and 26.643.4, and was significantly associated with BC (Fig 2; p=0.006 after adjustment for sex and PC1-3). Patients with a GRS in Q4 were 1.12 (95% CI: 0.99-1.26) times as likely to have BC as those in Q1.

CONCLUSIONS

- *PPARG* P12A confers risk for BC and its clinical subtypes in individuals of European, Latino and South Asian ancestry. *PAM* D563G and *PPIP5K2* S1207G may have protective roles in Europeans.
- Further studies are needed to understand the underlying mechanisms, and whether these variants modulate risk for moderate penetrance gene carriers and/or influence response to BC treatment.

AmbryShare™ https://share.ambrygen.com/ 15 Argonaut, Aliso Viejo, CA 92656 Toll Free 866 262 7943 Fax 949 900 5501 ambrygen.com Email: mblack@ambrygen.com